Psychological Review
2007, Vol. 114, No. 3, 577–598

0033-295X/07/$12.00 DOI: 10.1037/0033-295X.114.3.577

An Integrated Theory of Prospective Time Interval Estimation: The Role of Cognition, Attention, and Learning

Niels A. Taatgen
Carnegie Mellon University and University of Groningen

Hedderik van Rijn
University of Groningen

John Anderson
Carnegie Mellon University

A theory of prospective time perception is introduced and incorporated as a module in an integrated theory of cognition, thereby extending existing theories and allowing predictions about attention and learning. First, a time perception module is established by fitting existing datasets (interval estimation and bisection and impact of secondary tasks on attention). The authors subsequently used the module as a part of the adaptive control of thought—rational (ACT–R) architecture to model a new experiment that combines attention, learning, dual tasking, and time perception. Finally, the model predicts time estimation, learning, and attention in a new experiment. The model predictions and fits demonstrate that the proposed integrated theory of prospective time interval estimation explains detailed effects of attention and learning during time interval estimation.

Keywords: prospective time estimation, cognitive model, divided attention, instance learning, multi-tasking

The ability to estimate short time intervals routinely plays an important role in everyday life. Time estimates are important in situations in which we take an action and expect a response, for example, when we click on a link in a web browser or when we judge whether we should brake for a yellow traffic light. It also affects multi-tasking situations in which we have to switch between tasks after specific intervals, for example, when using a mobile phone in a car (Kushleyeva, Salvucci, & Lee, 2005; Salvucci, Taatgen, & Kushleyeva, 2006). This type of time interval estimation in real life is often implicit, automated, and tightly interwoven with other aspects of cognition such as perception, learning, and decision-making. All these examples concern *prospective time estimation*, because at the start of the interval it is known that an estimate will have to be made. This can be contrasted with *retrospective time estimation*, in which one is asked to estimate a duration after the time interval has passed. Prospective time estimation is often implicit in nature, because for most tasks the timing aspect is secondary to the real task being performed. For example, Grosjean, Rosenbaum, and Elsinger (2001) found that participants in a choice–reaction time experiment adapt to the interval between stimuli without being aware that they are doing so. The implicit aspect of prospective time estimation sets it apart from many other forms of reasoning about time that involve explicit reasoning and problem solving (see Michon & Jackson, 1985, for an overview). Thus, in retrospective time estimation, an explicit reasoning process might be used that involves recalling events that took place between the onset of the interval and its end (Zakay & Block, 2004). In our view, basic prospective time estimation is best explained as part of the human cognitive architecture in the same sense as visual perception is: Basic prospective time estimation processes are provided by a separate time module, and more complex and explicit forms of time estimation are explained by more general cognitive strategies that build on this basic capability.

Despite the fact that time estimation is, in general, only a component of complex task performance, it is usually studied in isolation. Zakay (1990) identified four paradigms to study interval estimation: (a) verbal estimation: after exposure to a time interval, reporting how much time has elapsed; (b) interval production: producing an interval of a certain duration, for example, 1 min; (c) interval reproduction: perceiving an interval of a certain duration and then reproducing it; and (d) interval comparison: comparing two intervals and reporting which is longer. In each of these paradigms time estimation is the explicit focus of the task. It is therefore quite possible that, analogous to the observation that memory studies using explicit recognition and recall do not necessarily tell the whole story of everyday implicit memory usage, explicit time estimation paradigms do not provide a complete picture of the role of time estimation within the cognitive system. As an example of a real-world task that involves implicit timing, consider sending a text message on a mobile phone. In order to

Niels A. Taatgen, Department of Psychology, Carnegie Mellon University, and Department of Artificial Intelligence, University of Groningen, Groningen, the Netherlands; Hedderik van Rijn, Department of Artificial Intelligence, University of Groningen; John Anderson, Department of Psychology, Carnegie Mellon University.

This research was supported by Office of Naval Research Grant N00014-06-1-0055. We thank Stefani Nellen and Simone Sprenger for their helpful comments on the manuscript and Daniel Dickison and Lisette Mol for collecting the data.

Correspondence concerning this article should be addressed to Niels A. Taatgen, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213. E-mail: taatgen@cmu.edu

send a text message with a numerical keyboard, multiple letters are mapped onto a single number key. For example, the letters *D*, *E*, and *F* are all under the number *3*. In order to type the letter *E*, one must press the *3* twice. In order to enter two consecutive letters from the same key (e.g., to type *DE*), one must insert a pause of about a second (*3*–pause–*3*–*3*) to disambiguate the two letters from three key-presses that signify the third letter (*F*). An important aspect in this task is learning: Even if the manual of the phone states that the interval is 1 s, learning the exact interval is partly characterized by trial-and-error. Only after sufficient practice does pressing the key at the right moment become fully automated.

One experiment that did study timing as secondary task was reported by Zakay (1989, Study 3). In this Stroop-type experiment, the experimenter instructed participants to focus on either the timing aspects of the task or on the word reading. This manipulation showed that emphasis on timing influences accuracy. However, the two tasks discussed by Zakay are two clearly separate and distinguishable tasks, instead of one single task in which timing is an integrated but secondary aspect.

In this article, we present a model of prospective time estimation as a module in a larger theory of cognition. We describe how this module interacts with other aspects of cognition to explain a wide variety of phenomena associated with time estimation. This embedded approach is necessary to fully understand the role of timing in both laboratory settings and in tasks like sending text messages, driving a car, and others involving complex skills. We embedded the timing model into the adaptive control of thought—rational (ACT–R; Anderson et al., 2004) model, a cognitive architecture that supplies mechanisms for learning, attention, perception, and motor behavior and which has been applied to many different tasks with a wide range of complexity. Before we present our own model, we review the existing models of time estimation.

Existing Theories of Prospective Time Estimation

Two theories address interval estimation: the *internal clock theory* and the *attentional gate theory*. Each of these theories has been formalized in one or more models: *internal clock models* and *attentional gate models*.

Figure 1 illustrates the *pacemaker–accumulator model*, of which there is both an internal clock and an attentional gate version. The *pacemaker–accumulator internal clock model*, as described by Matell and Meck (2000), does not have the attentional gate. Following Gibbon (1977), they identified a series of models that share the property that the internal clock itself is unaffected by outside processes. In one of these models, an internal pacemaker produces a steady stream of pulses. An accumulator counts these pulses, but only after a switch is opened by a start signal. After the time interval has ended, the accumulated value of pulses is stored in memory. When an interval of equal length has to be reproduced, a start signal is sent to the switch and pulses are counted until the same number of pulses has been reached as were stored in memory. In its most simple version, this model cannot account for differences in timing accuracy in tasks in which attention is (partly) directed away from the timing process. The attentional gate theory (see the attentional gate in Figure 1; see also Hicks, Miller, & Kinsbourne, 1976; Thomas & Weaver, 1975; Zakay & Block, 1997) was developed to explain that prospective time estimates tend to be longer if less attention can be paid to the estimation process, and vice versa. The model associated with this theory is an extension of the pacemaker–accumulator internal clock model. In addition to the components of that model, the attentional gate theory assumes that the accumulator is updated only when attention is being directed to the timing process, opening a gate. As soon as attention is directed elsewhere, the accumulator is not augmented until attention has returned. This way, attention determines the frequency by which the accumulator is being updated.

The amount of attention that can be paid to time estimation affects not only the mean of the estimate sometimes but also the variability. In a series of experiments by Brown (1997), participants had to repeatedly produce intervals of 2 or 5 s, either as single task or together with a second task. Brown found that in most experiments the estimates of the intervals were somewhat longer in the (more difficult) dual-task situations. However, the main effect he found was in the variability of the estimates. In the dual-task conditions, variability was increased by a factor of 2 to 3 compared with the single-task conditions.

The Internal Clock Theory

The internal clock theory is part of a long tradition of studying time perception in animals and in psychophysics. In some of Pavlov's experiments, the reinforcement was delayed by a particular time interval. When dogs were trained with the delay, they started salivating only at the end of the interval (Pavlov, 1927). Many other animal studies have shown rats, dogs, pigeons, and

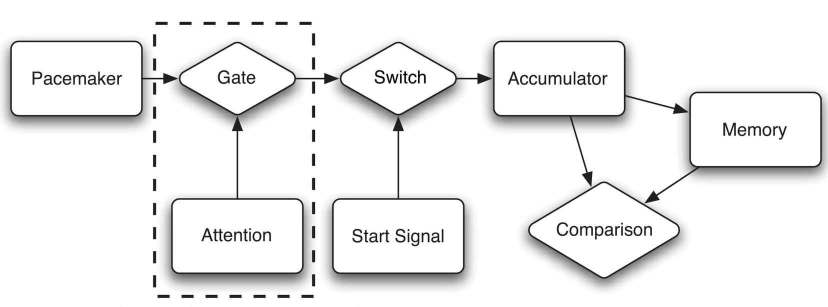

Figure 1. The pacemaker–accumulator model. The internal clock model does not have an attentional gate, whereas the attentional gate model does.

other animals to be capable of learning the temporal structure of tasks.

Studies in psychophysics have shown that time perception shares characteristics with other forms of perception, most notably Weber's law. The consequence of this law is that uncertainty in a time estimate scales with the magnitude of the interval, which is also called the *scalar property* of time estimation (Gibbon, 1977).

Matell and Meck (2000) gave an overview of three possible models of the internal clock theory: a *pacemaker–accumulator model*, a *process-decay model,* and an *oscillator–coincidence detection model*. All three models are based on an internal clock that is not affected by attention. Figure 1 depicts an example of a pacemaker–accumulator model in which an accumulator counts the pulses generated by a pacemaker. In a process-decay account, decay of activation in memory is used to estimate elapsed time (Staddon, 2005; Staddon & Higa, 1999, 2006). In the oscillator–coincidence detection account, which is favored by Matell and Meck because of its neurobiological feasibility, stimuli can synchronize neurons in a certain area of the cortex, effectively acting as a starting sign. As each of the neurons produces its own particular pattern of activation over time, each interval is associated with a unique pattern of activation, which can serve as a basis for later comparison. Although Matell and Meck's three models differ in their neurobiological plausibility, they are equivalent with respect to time-estimation-related predictions, even though the pacemaker–accumulator model has to make some additional assumptions. Because we are, in the context of this article, primarily interested in the behavioral characteristics of time estimation, we consider these implementations as belonging to one family.

In contrast to the attentional gate theory, internal clock models do not require any attention, and errors in time estimation are due to noise in the system. In a typical interval timing experiment, participants were trained on an interval of either 8, 12, or 21 s by being exposed to it multiple times, which they then had to reproduce (Rakitin et al., 1998). Each participant produced 80 estimations and was given feedback about the true durations every few trials. Figure 2 shows the distributions of the responses. Although the variance increased for larger intervals, the peak of each of the distributions aligns with the duration of the target interval. Consistent with Weber's law, these distributions exhibit the scalar property: The standard deviation in the estimation increases approximately linearly with the length of the interval. If we were to divide the times on the *x*-axis in Figure 2 by 8, 12, and 21, respectively, and readjust the proportions on the *y*-axis, the three curves would fall on top of each other (see Rakitin et al., 1998, Figure 2). The pacemaker–accumulator model can reproduce this scalar property if equality judgments are based on the ratio of the current interval and the target interval (Gibbon, 1977; Matell & Meck, 2000).

The Attentional Gate Theory

In experiments associated with time estimation, the estimation task itself is almost always accompanied by an unrelated secondary task. The purpose of the secondary task is to prevent explicit counting, because counting makes time estimation much more accurate (e.g., Rakitin et al., 1998, Experiment 2). However, the nature of this secondary task turns out to have an influence on the estimation of the interval. If the secondary task is very demanding, people's estimation of duration tends to be shorter than when the secondary task is less demanding (Block & Zakay, 1997, offer a meta-analysis of 20 experiments). The attentional gate theory accounts for this by assuming that fewer pulses accumulate when another task demands attention, leading to a shorter estimate.

Here, we focus on an experiment by Zakay (1993), in which participants had to estimate and reproduce a single interval (12 s) once. Each participant was required to perform a second task. This task had to be performed either during the presentation of the interval, when participants had to determine the duration, or during the reproduction of the interval, when participants had to use their

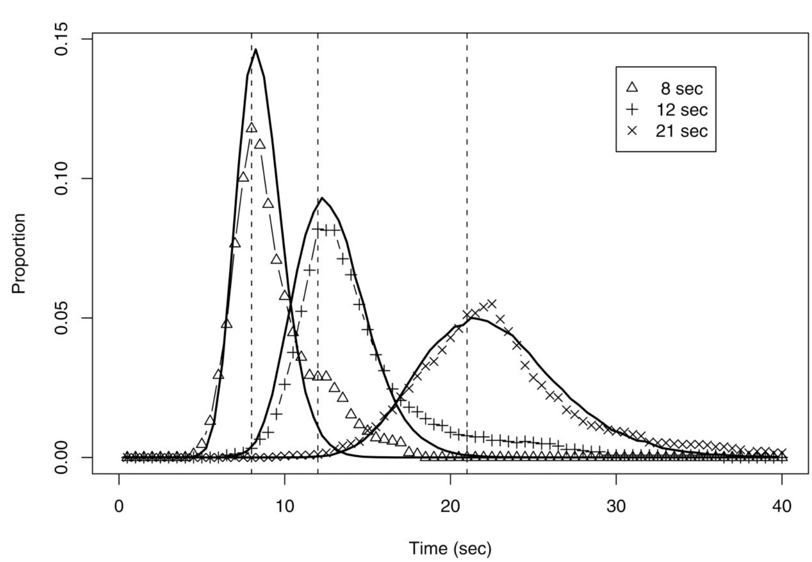

Figure 2. Distribution of estimates of intervals of 8, 12, and 21 s by participants in the Rakitin et al. (1998) study. Vertical lines indicate 8, 12, and 21 s. The solid lines are adaptive control of thought—rational (ACT–R) model fits discussed in the section *A Model of Perception and Reproduction of Time Intervals*. sec = seconds.

perceived duration estimation for reproduction. The attentional gate theory predicts that a demanding task during the reproduction leads to an estimate that is too long, because during the reproduction the timer is slowed down. In contrast, a demanding task during the presentation leads to an estimate that is too short, because fewer pulses have been counted during presentation (see Figure 3, right). The secondary tasks were, in increasing complexity, (a) empty time (ET): no secondary task; (b) words (W): reading color words printed in black; (c) color words (CW): naming the color of color words printed in incongruent ink (the Stroop task); and (d) color–word associations (CWA): like the Stroop task, but participants had to name a word associated with the ink color. In the relatively easy ET and W conditions there is no effect of the secondary task on time estimates, but in the more demanding CW and CWA tasks, time estimates are affected in the way the attentional gate theory predicts (see Figure 3, left).

However, a weakness of the attentional gate model is that the impact of attention has not been quantified precisely. The model has no basis on which the proportion of time spent with the attentional gate opened versus closed can be assessed. In addition, the attentional gate model assumes that the value of the accumulator can be stored in memory for later comparisons, but it has no detailed account on how multiple experiences interact. The model generally does not need such an account because in most experiments supporting the attentional gate theory, participants make only a single estimate of a time interval they have perceived once. (A single estimate is one of the criteria that Block & Zakay, 1997, used for inclusion in their meta-analysis.)

Recent Additions to the Internal Clock Theory

Lejeune (1998) has argued that the attentional gate is not necessary because its functional characteristics can be modeled by the switch that starts the estimation process, leading to a debate of *switching* versus *gating* (Lejeune, 1998, 2000; Zakay, 2000). According to Lejeune (1998), the detection of the onset of a time interval depends on how much attention is being paid to the external signal that signifies the onset. If less attention is being paid, the start of the timing process can be delayed. In addition to that, Lejeune discussed the possibility that the switch can be opened and closed during the time estimation process, making the model's functional characteristics potentially very similar to the gating theory.

Buhusi and Meck (2006) proposed a model in which the values in an accumulator are subject to decay and showed that this model is able to account for multiple phenomena associated with animal and human time estimation. In this model, the amount of decay is assumed to be in proportion to secondary events' salience, explaining the increased time estimation in settings in which participants are confronted with a more demanding secondary task. Buhusi and Meck presented a computational implementation of their theory, in which the amount of decay was estimated separately for different conditions. In this way, the model could be used for explaining observed time estimations but provides only limited information when predicting time estimation distributions in the context of new secondary tasks.

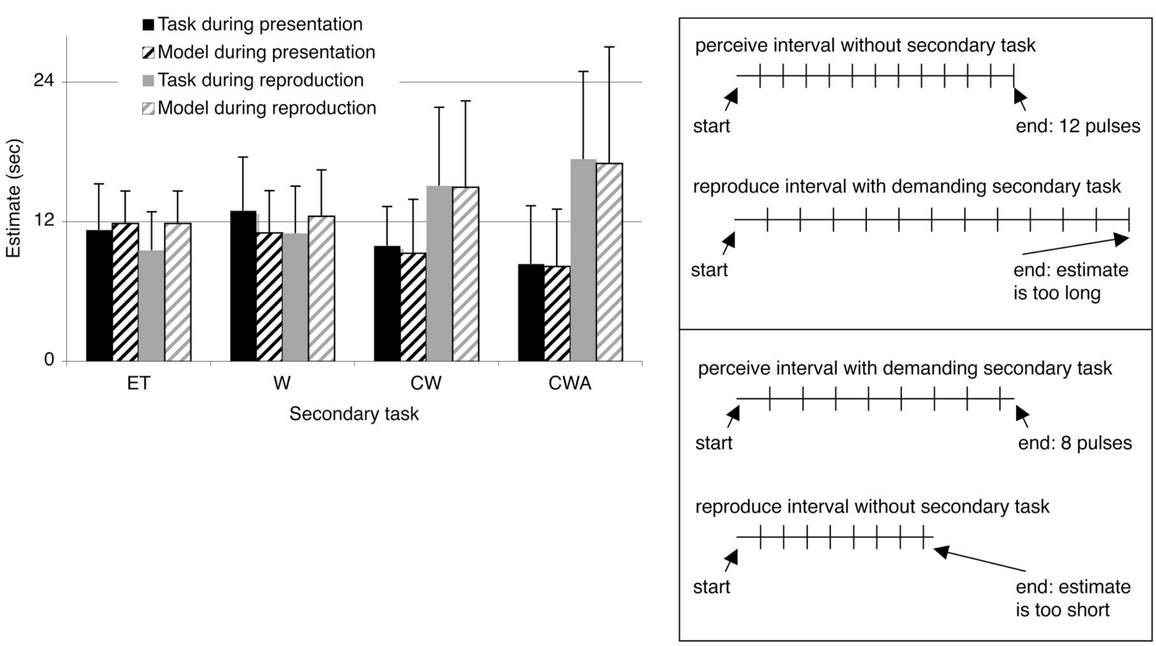

Figure 3. Left: Results of Zakay's (1993) experiment and the adaptive control of thought—rational (ACT–R) model fits discussed in the section *A Model of Perception and Reproduction of Time Intervals*. Right: A depiction of the attentional gate theory explanation. Error bars are standard deviations. ET = empty time: no secondary task; W = words: reading color words printed in black; CW = color words: naming the color of color words printed in incongruent ink (the Stroop task); CWA = color–word associations: like the Stroop task, but participants had to name a word associated with the ink color.

Overview of the Article

Internal clock models focus mainly on the estimation of the interval itself and have only an approximate theory on how this connects to other aspects of cognition. The attentional gate model incorporates one of these aspects, attention, but the extent of the model is limited primarily to explicit, one-shot time estimates. Even in those situations the model's predictions are mainly qualitative. We take the approach of considering time estimation as a module in a more extensive cognitive system. In this approach, the timing module accounts for all primary aspects of time estimation. Secondary and more complex aspects of time estimation, such as dividing attention with other tasks and how time intervals are learned, are explained by how timing interacts with the rest of cognition. The impact of attention can be explained by general models of divided attention. In addition, existing theories of learning and skill acquisition account for specific learning effects found in interval estimation.

In order to be able to understand interval estimation in a broader cognitive context, we have embedded the time estimation module within the ACT–R cognitive architecture (Anderson et al., 2004). The basis for the module is a pacemaker-based internal clock, with functional characteristics similar to the internal clock accounts proposed by Matell and Meck (2000). Interaction with the rest of the system allows explanations for the role of attention and learning, without the need to incorporate the latter into the timing module itself.

We build our case as follows: On the basis of the distributions of time responses found by Rakitin et al. (1998), we constructed an internal clock module that reproduces these distributions. We then validated this module by reproducing the results of the bisection experiments of Penney, Gibbon, and Meck (2000). To test the module in a context in which attention, learning, perception, and motor actions interact, we constructed a task in which keeping track of time intervals is only a single aspect of what participants had to do. The goal of this experiment was not only to test the quantitative accuracy of the model but also to find support for the role of attention and learning as they are hypothesized by the incorporation of the temporal module in ACT–R. To this end, we first conducted the experiment and then constructed the model to fit the data. To test whether our account also holds when the to-be-explained data are not known beforehand, we changed the setup of the experiment to make attending to the time harder than in the first experiment. Before conducting the experiment itself, we applied the model to this new task to predict the outcome. The experiment then confirmed the predictions made by the model.[1]

The Basic Internal Clock Module

In this section we establish the internal clock module and make only minor assumptions about processing in the rest of the architecture. The larger architecture and its impact on time perception are discussed in The Dual-Task Timing Task section.

The internal clock module is based on the pacemaker–accumulator model. This facilitates comparisons with existing models, most of which also use a pacemaker and an accumulator. According to the pacemaker–accumulator model, a pacemaker generates pulses at certain intervals, which are counted by an accumulator. A reset event sets the accumulator to zero, after which it starts counting pulses anew. Instead of assuming a pacemaker with a constant rate, we chose to increase the interval between the pulses as the interval progresses, like a metronome that ticks more and more slowly with time. The interval estimate is based on the number of pulses the accumulator has counted. This process implicitly creates a logarithmic time scale in which longer intervals have counts that are closer together than those of shorter intervals. Because the gradual slowing of the pacemaker occurs in both perception and reproduction, it will lead to the same estimate, although less precise with longer time intervals as the time between pulses gets longer. In our approach, the temporal system is considered a module with an internal process that runs independently of other cognitive processes. The cognitive system as a whole (i.e., the ACT–R architecture) has access only to the result of this process, that is, the current value of the accumulator. Because the temporal module is encapsulated, any other model with the same behavioral characteristics can in principle replace it and produce the same behavior.[2]

The duration of the first pulse is set to some start value: $t_0 = startpulse$. Each pulse is separated from the previous pulse by an interval that is a times the interval between the previous two pulses. Noise from a logistic distribution is added to each pulse. The distribution of this noise is determined by the current pulse length, modified by a parameter b:

$$t_{n+1} = at_n + \text{noise}(M = 0, SD = b \cdot at_n).$$

These equations have three parameters: *startpulse*, a, and b. As the behavior of the timing module is assumed to be independent from the task that the architecture currently executes, these parameters should be estimated to fit a single "benchmark task" and then be left untouched. As a benchmark task we used the Rakitin et al. (1998) experiment whose results are presented in Figure 3. We have left the parameters at the estimated values for all other models presented here.

A Model of Perception and Reproduction of Time Intervals

The basic model for perceiving and reproducing intervals is simple. During the perception of an interval, the accumulator will first be reset and the pacemaker starts at the beginning of the interval. The value of the accumulator is read at the end of the interval and stored in memory. If there are multiple presentations of the interval, the stored values are averaged to obtain a more

[1] For the implementation details of the models in this article, we refer to the complete model code available on the Internet at http://act-r.psy.cmu.edu/models

[2] For example, the process-decay model by Staddon and Higa (1999) assumes time is represented by a function that decreases logarithmically with time. Our model has a representation of time that increases logarithmically with time. However, because the representations are used only for comparisons, the fact that they increase or decrease is functionally irrelevant. Therefore, using a value that decreases logarithmically with time, instead of an accumulator, would lead to the same model results.

accurate estimate.[3] Reproducing an interval means starting the timer, waiting until the accumulator has reached the stored value, then making the response. Figure 4 illustrates this process. When the stimulus appears, the counter is started. As soon as the stimulus disappears, the value of the counter, equaling six pulses, will be read out and stored in memory. When the interval has to be reproduced, the timer is started again, and when the counter reaches the stored value, the model assumes that the same amount of time has passed.

In Experiment 3 of Rakitin et al. (1998), participants were first trained on a certain time interval (8, 12, or 21 s). Training consisted of 10 trials in which a blue rectangle appeared on the screen and changed to magenta when the time interval had elapsed. In the 80 test trials, participants had to predict the interval by pressing a key when they expected the rectangle to change color. To make sure that participants' representation of the interval would not drift, the rectangle changed color when the interval had elapsed in 25% of the test trials. The results are based on the remaining 75%, in which the rectangle stayed blue. Participants were instructed not to count during the experiment.

The model of this experiment closely resembles the example in Figure 4. In the learning phase the number of pulses in the interval was estimated (the model took the average of the 10 presentations), and during the testing phase the model waited until the appropriate number of pulses had passed and then made a response. Based on least-square estimations of fit between model and data, we estimated the following values for three model parameters: 11 ms for *startpulse*, 1.1 for *a*, and 0.015 for *b*. Figure 2 shows the fit between this model and the three conditions of the Rakitin et al. (1998) experiment. The fit between model and data is overall very good. This is no surprise considering there are three parameters to fit the data. The only aspect of the data the model did not predict perfectly is the shape of the tail of the distribution for the 8- and 12-s conditions. However, in similar experiments (e.g., Experiment 1 in Rakitin et al., 1998) the tails of the distributions were much shorter. We therefore decided against complicating the temporal module with a mechanism to simulate this aspect of the data and made the assumption that these tails were produced by factors outside of the temporal module itself.

Application of the Module to Bisection Experiments

Another paradigm in time perception concerns so-called bisection experiments. In these experiments, participants are trained on two time intervals: one short interval and one long interval. After this learning phase, they are exposed to new time intervals that are either equal to the short or the long interval or somewhere in between. Participants are then asked to judge whether the presented interval is closer to the short or to the long interval. To test whether the estimated parameters fit this timing paradigm equally well, we modeled Experiment 2 from Penney et al. (2000). In this experiment, three short–long pairs of intervals were used: 3 and 6 s, 2 and 8 s, and 4 and 12 s. During the training phase, 10 tones of either the short or the long duration were presented to the participant. After that, participants were tested for 100 trials, 30% of which were anchor point intervals (short or long), and 70% were tones of different durations in between the anchor points.

The model used the training phase to determine the timing for both the short and the long interval. During testing, it counted the number of pulses during the presented interval then compared the value to the number of pulses associated with both anchor intervals. If the value was closer to that of the short interval, it answered "short"; if it was closer to the long interval, it answered "long." The parameters for the model were identical to those used to fit the interval estimation experiment earlier. Figure 5 shows the results of the experiment and the model. A typical result in bisection experiments is that an interval exactly in between short and long is judged to be long considerably more often than chance. For example, in the 2–8-s version of the task (Figure 5, top), 5 s is judged to be long 80% of the time. According to the model, this happens because 2 s corresponds to an average of 33 pulses, 5 s to an average of 42 pulses, and 8 s to an average of 47 pulses. In terms of pulses, 5 s (42 pulses) is much closer to 8 s (47 pulses) than to 2 s (33 pulses). As can be seen in Figure 5, this mechanism yields results very similar to those observed by Penney et al. (2000). The fact that the model predicted that the mid-point between long and short intervals is before the actual mid-point is directly attributable to the logarithmic scale of perceived time. However, the fact that all the other points on the graph are fit well by the model indicates that noise in the estimate is also modeled correctly. For example, the model correctly predicted that in the 3–6-s experiment, in 9% of the cases a pure 3-s interval was judged as long, and a 6-s interval as short in 5% of the cases.

The interval estimation and bisection models demonstrate that the theory can accurately fit existing timing data, but nothing beyond what existing models can already account for. We now proceed to cases in which attention and learning play a role and show how an integrated approach is needed to account for them.

An Alternative Model of Zakay's (1993) Results

In the introduction we discussed the experiment by Zakay (1993) in which time estimates were influenced by the difficulty of a secondary task. This effect was reason for Zakay and others to propose the attentional gate theory, introducing a modulating effect of attention on the timing mechanism (Hicks, Miller, & Kinsbourne, 1976; Thomas & Weaver, 1975; Zakay & Block,

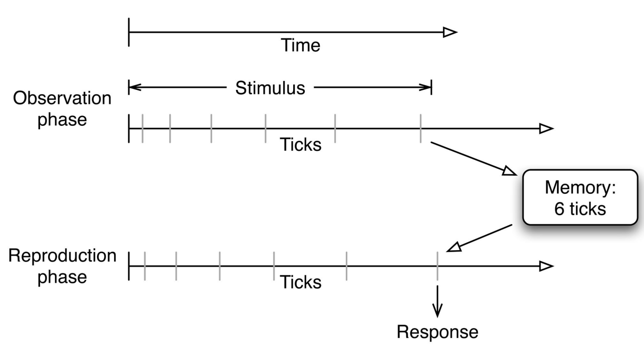

Figure 4. Illustration of perceiving and reproducing a time interval.

[3] For the models discussed in the *Learning the Time Interval Through Instance Learning* section below, a more sophisticated learning technique, instance learning, is used. However, for this experiment and the bisection experiment discussed next, instance learning and averaging produce similar results.

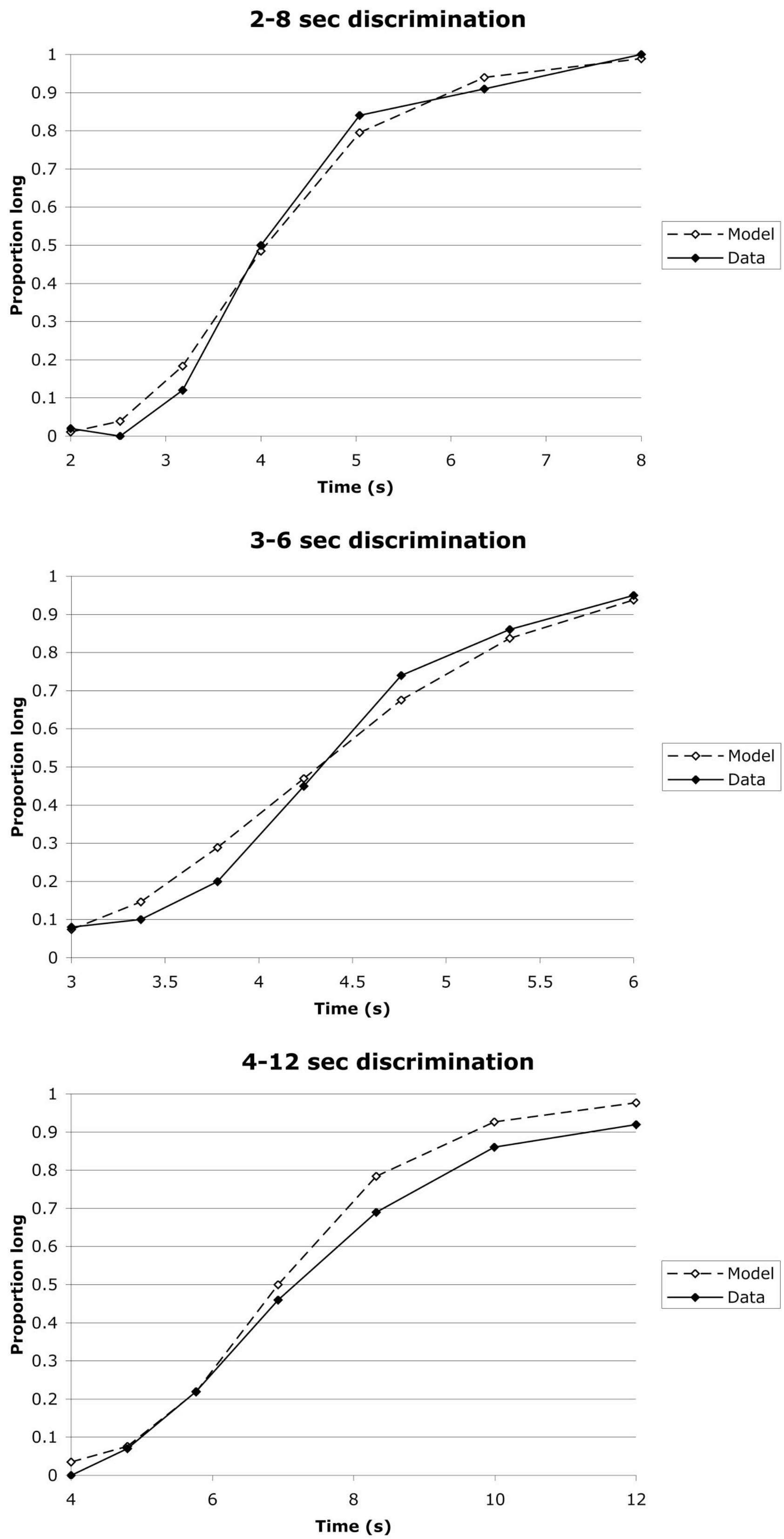

Figure 5. Comparison between the adaptive control of thought—rational (ACT–R) model and data in three bisection experiments by Penney et al. (2000). sec = seconds.

1997). However, an important aspect in the experiment is that all participants made only a single estimation of an interval. This might have made them prone to all sorts of "startup" mistakes. One possible mistake is that the temporal module is accidentally used for one of the secondary tasks. The temporal module could be used implicitly to estimate the inter-stimulus intervals in the secondary task (of the type manipulated in the Grosjean, Rosenbaum, & Elsinger, 2001, experiment). This means the value of the timer is lost for the primary task. The probability for this becomes larger as the secondary task becomes more demanding. Because the task demands that some response is made at some point, the model's strategy is to reset the counter and proceed from there. A reset during the presentation means fewer pulses are counted, leading to shorter estimates, whereas a reset during reproduction leads to a restart of the estimate somewhere partway and therefore longer estimates. In order to reproduce the Zakay (1993) data, we set the interruption probability for each pulse to 0% for empty time (ET), to 1% for word reading (W), to 3.5% for the Stroop task (CW), and to 5.5% for Stroop with association (CWA). On the basis of these estimates, the model produced the results in Figure 3, essentially producing a fit that is very similar to what the attentional gate theory would predict using the same number of free parameters. However, our alternative model can also fit the standard deviations of the experiment. Although Brown (1997) showed that manipulations in attention influence variability of time estimates, it is not clear to what extent the attentional gate theory can explain the magnitude of this effect.

Although the model based on the temporal module presented here requires the estimation of disruption parameters for each of the four tasks, it is similar to an attentional gate model in explanatory power, as the attentional gate model would require an estimate of the proportion of attention available for interval estimation for each task. The model based on the temporal module can be considered slightly better, because it also fits the variability in the estimates. The two explanations are remarkably similar on the surface, despite the fact that the internal structures that produce them are quite different. The alternative model can therefore be easily adapted to model other experiments in which participants have to make a single time estimate. The fact that two conflicting models can both fit the data shows that the validity of both theories (attentional gate and the temporal module) cannot be decided on the basis of experiments with single estimates.

Summary

We have described a temporal module that was designed to reproduce the distribution of estimates in a simple timing experiment. A model based on this module proved to be capable of explaining a second class of timing experiments, the bisection experiments, without adjusting any parameters of the underlying module. To fit the model to the Zakay (1993) data, we made some assumptions about the possibility that timing is disrupted. On the basis of these cases, we can conclude that the temporal module is quite successful in explaining data from existing time-estimation tasks.

The simplicity of the tasks in this section allowed explanations in terms of the temporal module and a few additional assumptions. Although the temporal module is successful in the sense that it can offer explanations for experiments that have been studied in the context of both the internal clock theory and the attentional gate theory, it has not yet offered any details of larger integration with other aspects of cognition. Both the attentional gate theory and our temporal module can explain Zakay's (1993) data, but neither provides an account in terms of attention as a general cognitive process. Furthermore, most studies on timing have neglected the effects of learning. Do time estimations get better over time? And if so, how does attention modulate this effect? To answer these questions, we designed an experiment that incorporates both learning and attention. To account for the data from this experiment, we have incorporated the temporal module into the ACT–R architecture. ACT–R already provides mechanisms for learning and attention. This provides us with an appropriate test-bed for assessing our claim that timing is an integral aspect of cognition and that the interplay of different cognitive mechanisms results in the observed timing effects.

The Dual-Task Timing Task

The purpose of the dual-task timing task (DTT) is to study the effects of attention and learning on interval estimation in a fairly complex task, in which time estimation, at least from the participant's perspective, is just one of the many prerequisites to achieving accurate performance. The task is supposed to mirror real-life situations in which people have to discover the temporal structure of a situation or a device.

In the DTT, participants worked on two simultaneous subtasks that were either both hard (verify additions) or both easy (recognize letters). Points were awarded for each correct response. A time interval had to be estimated as part of one of the tasks. The participant had to determine the duration of this interval by trial-and-error while doing the other task. One aspect of the task is that the primary goal from the perspective of the participant is to respond to the stimuli (because that scores points directly), and estimating the interval is only secondary (because it helps in scoring more points). To be successful at the task, it is necessary to spread attention over all the subtasks. To study the specific effects of attention, we switched the difficulty of the task at some point from easy to hard in one of the conditions and vice versa in one of the other conditions. This change in task difficulty modified the amount of attention that could be spent on estimating the interval. Because the task involved many repetitions, it also allowed us the study of the effects of learning.

Figure 6 shows the task display and an example of a single trial. The display was divided into two halves. The left half contained a high-profit box, and the right half a low-profit box. Stimuli, to which participants could respond, appeared in each of these boxes. Stimuli were buttons with either an addition (with one-digit numbers) or a letter, depending on the condition. Additions were either correct or wrong by one, and letters were either *A* or *B*. Participants had to respond to correct additions and to *A*s by clicking on them with the mouse when they were in the right box or by pressing the space key on the keyboard when they were in the left box. Incorrect additions and *B*s had to be ignored. Stimuli in the left box did not appear automatically: They were available only during certain time periods and after the participant had pushed the *Test High* button. Each trial was 13 s long and built up as depicted in Figure 6, bottom. To indicate the previous trial had ended, the text *end of high profit* appeared in the left box. During the next 7 s the

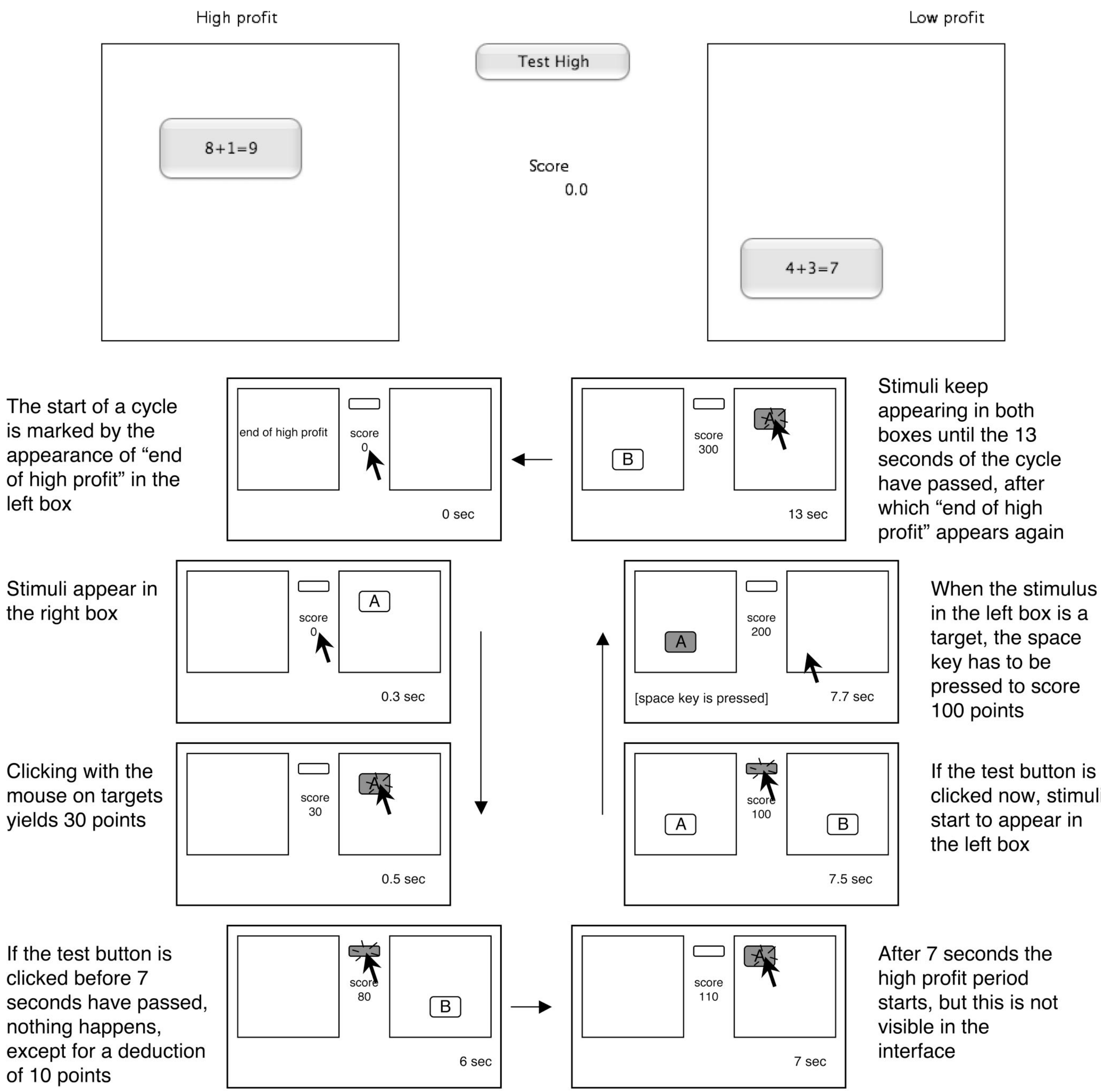

Figure 6. The dual-task timing task. Top: A screenshot of the Addition task. Bottom: Example of a single trial with the Letter task. sec = seconds.

interface would present stimuli only in the right box. Correct responses (clicking on *A*s or correct additions) in the right box scored 30 points. Clicking on the *Test High* button (we call this the *test button* from here on) during this interval had no effect. When the test button was clicked after 7 s had passed (but before the end of the 13-s trial), stimuli started appearing on the left side. Testing for the high-profit period by clicking on the test button always cost 10 points. Participants had to respond to targets in the left box by pressing the space key on the keyboard. Correct responses in the left box scored 100 points. Because stimuli also kept appearing in the right box, it was possible to work on both boxes at the same time. After a total of 13 s had passed, the text *end of high profit* appeared in the left box and a new 13-s trial was started. The text *end of high profit* also appeared when the participant had not clicked the test button at all. Optimal behavior was to click the test button exactly 7 s after the word *end* appeared in the left box. Participants were informed only that high-profit periods would appear at fixed durations but not of the length of the interval, which they had to discover themselves.

Experiment 1

The experiment had four between-subject conditions. Each condition consisted of three phases. In each phase the task was either Letter (easy) or Addition (hard). The four conditions were as follows: three phases with the Letter task (LLL), three phases with the Addition task (AAA), two phases with the Letter task followed by one phase with the Addition task (LLA), and two phases with the Addition task followed by one phase with the Letter task (AAL). Each phase consisted of five 120-s blocks of nine 13-s trials. Additional details on the experiment can be found in the Appendix.

Although the model was constructed and fitted to the data after the first experiment with the DTT, we did have some expectations of the results on the basis of the timing module and general ACT–R characteristics. A first expectation was a learning effect in the sense that participants would become increasingly better at estimating the interval and improving their score. The reason for this is that ACT–R generally learns from experience by storing and retrieving examples of past behavior. Thus, as the model gains experience it is able to approximate the time interval with increasing accuracy.

A second expectation, contrary to the attentional gate theory's prediction, is that the interval transfers perfectly from one task to the other. That is, when the duration of the interval is learned during one task, this knowledge can be used to estimate the interval for a different version of the task that places higher or lower demands on attention. As a consequence, we expected that the effects of changing task difficulty on time estimates would be small. As we show, the main impact of task difficulty is similar to that in our model of Zakay's (1993) experiment (i.e., losing track of timing). But the behavioral manifestation is different. In the Zakay experiment, a response had to be made at some point, even if the participant had lost track of time. In the DTT task, participants can just skip a trial and try again on the next one. Instead of leading to late responses, timing errors lead to non-responses in the DTT task. And whereas in our model of the Zakay experiment the "resetting" of the timer had to be explicitly modeled, missing an estimate in the DTT task is a side effect of performing the main task.

Time Estimation

The two solid lines in each side of Figure 7, top, plot the distributions of the moments at which participants first clicked the test button within a trial. These time points are defined as the deviation from the optimal time, that is, the time at which new high-profit stimuli became available (which is 7 s into the trial, so −7 is the beginning of the trial). A negative value represents a click that is too early, and a positive value a click that is too late. The data are averaged over the two conditions that start with the Addition task (AAA and AAL) and the two conditions that start with the Letter task (LLL and LLA) and are plotted separately for Phase 1 and Phase 2. The higher peaks for Phase 2 suggest that participants were more accurate in Phase 2, indicating that a more accurate estimate had been learned. The proportions plotted in these and all subsequent histograms are based on the total number of trials in the phase (instead of the total number of first clicks). This means that trials in which no attempt at all was made to make a time estimate also weigh into the proportions (we discuss these non-response trials below). The dotted lines plot the distributions that would be expected if this were a pure interval estimation experiment like the Rakitin et al. (1998) experiment (see Figure 2). We derived this expectation by scaling the distribution according to the scalar property for the 8-s interval from the Rakitin et al. experiment to the 7-s interval of this experiment. The wider empirical distributions indicate that participants performed worse than ideally—that is, they deviated more from the optimal time—which could be expected because participants first had to discover the duration of the interval. Analyses of variance (see the Appendix) of the moment of the first click revealed that the only significant factor was phase, indicating that participants improved their time estimates with practice. No effect was found for condition or the interaction between condition and phase.

In the discussion of the Zakay (1993) model, we assumed that inaccuracies in time estimation were caused partly by participants forgetting about time estimation and restarting it at some point later. In the present task, forgetting to estimate results in making no estimate at all. This can be assessed by analyzing how often participants failed to make any estimate in a given trial. Figure 8 shows the proportions of non-responses by phase and condition. A logistic regression of the non-response proportions (see the Appendix) did show an effect of condition (together with an effect of phase). Combined with the analyses of the estimation accuracy, this suggests that the task difficulty did not affect the accuracy of the time estimate itself, but it does cause people to increasingly omit making an estimate at all.

Changes in the Accuracy of Time Estimation Due to Changes in Task Difficulty

An interesting question is what happens to the distributions of time responses when the task difficulty changes. According to the attentional gate theory, a major change in estimate should occur after the shift. This follows from the idea that fewer ticks reach the accumulator when the task is more difficult. One can compare this with a slower or faster ticking clock (see Figure 3, right). If the estimate of the interval is based on the clock that ticks quickly, and the clock is slowed due to a more difficult task, the estimate should be too long. On the other hand, if the estimate is based on a slow clock that speeds up due to an easier task, the estimate should be short. Figure 9, top, shows what this theory would predict if the clock ticks 25% more slowly on the Addition task than on the Letter task.

Figure 9, middle, compares the empirical distributions of the first clicks before and after a switch in difficulty. Although the changes in distribution are slight, there is a significant shift to the left in the AAL condition, indicating that participants clicked earlier after the change to an easier task, whereas there is a significant shift to the right in the LLA condition, indicating that participants clicked later after a change to a harder task. Although the shift in the AAL condition can still be attributed to a learning effect, the shift towards later responses in the LLA condition seems to support the attentional gate theory. In order to obtain a better idea of the nature of the shift, we compared

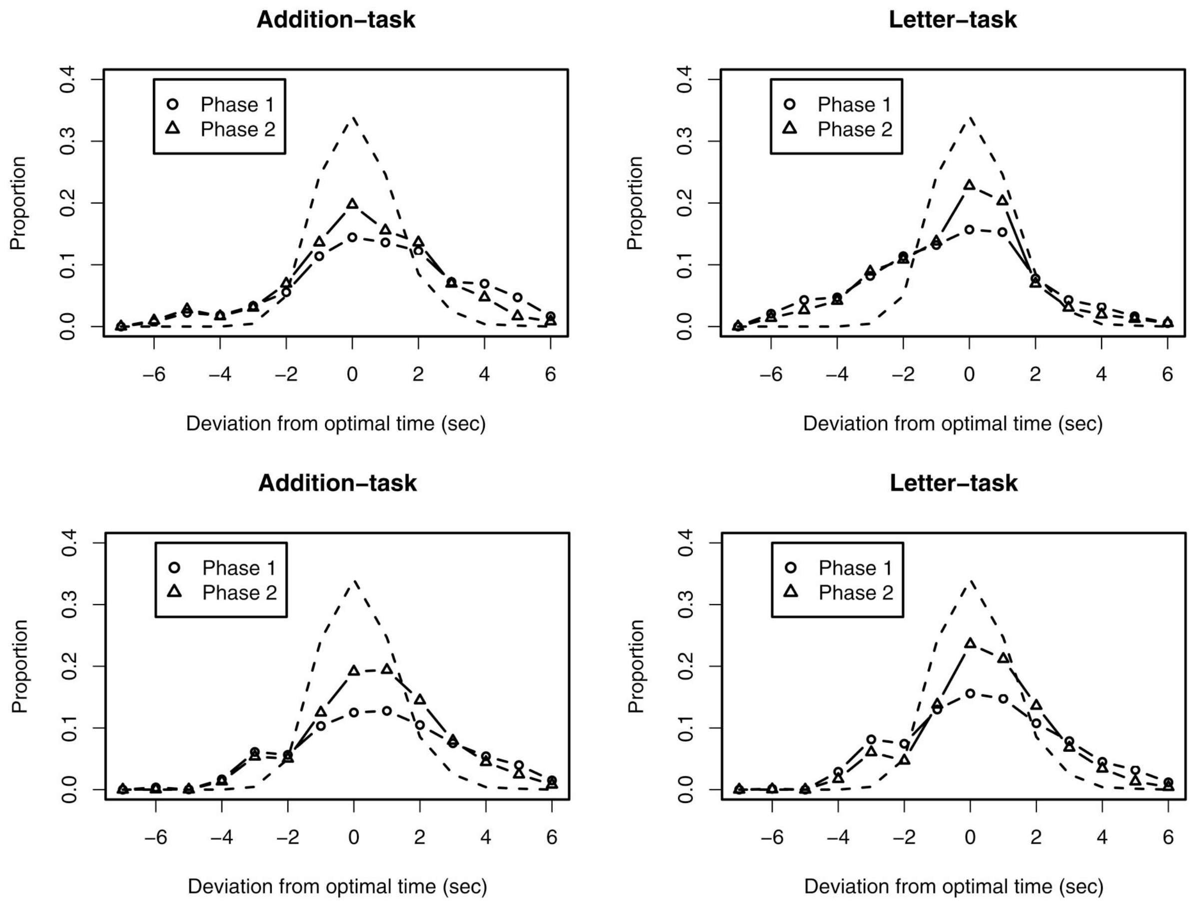

Figure 7. Distribution of first-click times in two subsequent phases for both tasks. Top: Empirical data of Experiment 1. Bottom: Adaptive control of thought—rational (ACT–R) model fit. The dashed line is the expected distribution for a pure interval estimation experiment. sec = seconds.

the average response times before and after the shift in task difficulty (see Figure 10). The attentional gate theory would predict that the effect of the change would be most pronounced on Trial 91, right after the shift, because that is where the update rate of the accumulator suddenly changed and the participant had no opportunity to adjust. Instead of a peak in click time that leveled off afterwards, the click time increased gradually after the shift. Moreover, the average click time con-

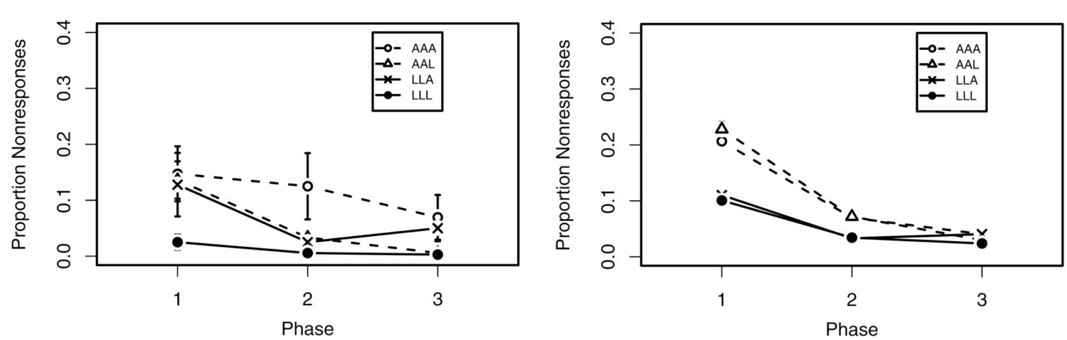

Figure 8. Proportions of non-responses. Left: Empirical data of Experiment 1. Right: Adaptive control of thought—rational (ACT–R) model fit. Error bars are standard errors. AAA condition = three phases with the Addition task; AAL condition = two phases with the Addition task followed by one phase with the Letter task; LLA condition = two phases with the Letter task followed by one phase with the Addition task; LLL condition = three phases with the Letter task.

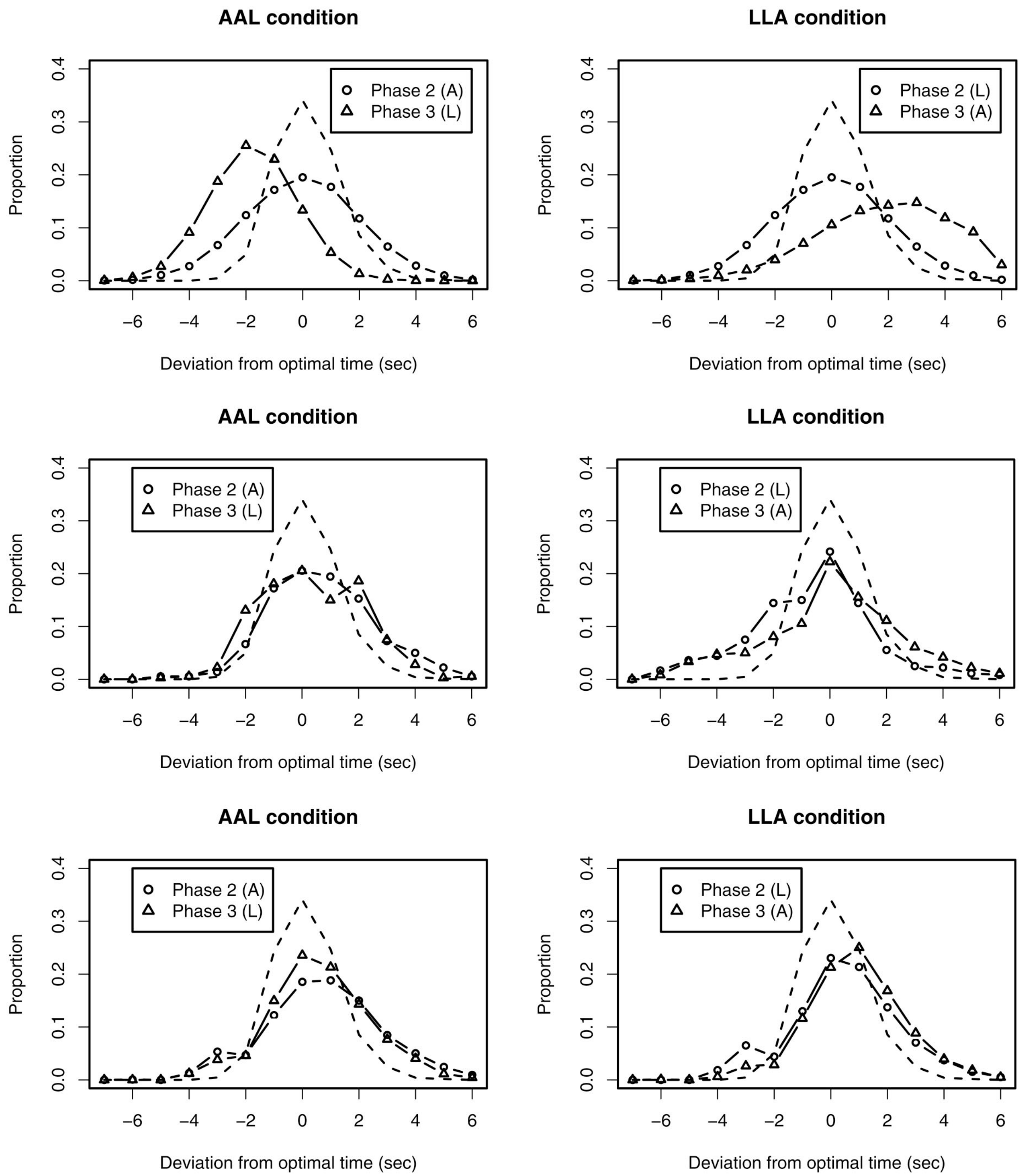

Figure 9. Distribution of first-click times in Phases 2 and 3 for the AAL and LLA conditions. Top: Attentional-gate-theory-based prediction. Middle: Empirical data of Experiment 1. Bottom: Adaptive control of thought—rational (ACT–R) model fit. The dashed line represents the distribution that would be expected if this was a pure interval estimation experiment (cf. Figure 2). AAL condition = two phases with the Addition (A) task followed by one phase with the Letter (L) task; LLA condition = two phases with the Letter task followed by one phase with the Addition task. sec = seconds.

verged to that in the AAA condition but did not exceed it. This indicates that responses in the Addition task were all slightly later than those in the Letter condition and independent of switches in the task. A comparison of the average response moment in Phases 1 and 2 for the Letter and the Addition tasks confirmed this observation (mean response time for the Letter task was −472 ms; mean response time for the Addition task was 594 ms).

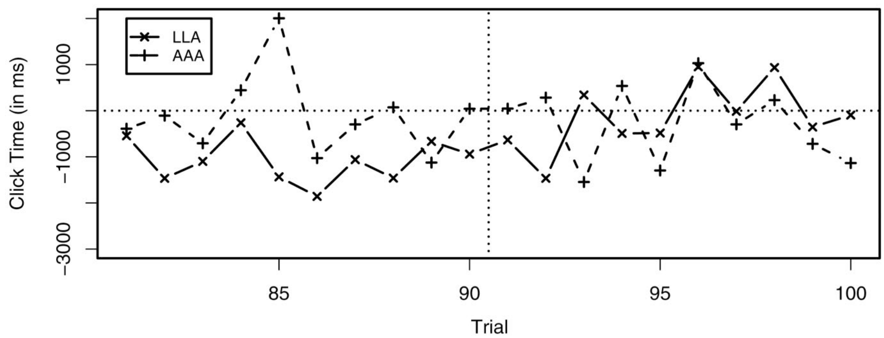

Figure 10. Average click time in the LLA condition (solid line) 10 trials before and after the change in task difficulty (indicated by the vertical dashed line). For comparison the AAA condition is plotted with a dashed line. LLA condition = two phases with the Letter task followed by one phase with the Addition task; AAA condition = three phases with the Addition task.

To summarize, task difficulty had hardly any impact on the accuracy of time estimation as evidenced in the timing of the first click. Instead it had an effect on how often a time estimate was missed, which is similar to our account of Zakay's (1993) results. Time estimates for the Addition task were all slightly later than those for the Letter task, without affecting the absolute accuracy of the estimate.

The results up to here imply that task difficulty and shifts in task difficulty have an impact on the task performance and the role of interval estimation, but not in the way the attentional gate theory would predict. The attentional gate theory would predict significant and immediate shifts in timing after a task change (as in the AAL and LLA conditions). The shift in the LLA condition was, however, small and seemed to be more related to global properties of the tasks than to changes between tasks.

Dual Tasking

A possible explanation for the small impacts of task on time estimation is that the Addition task is too easy: Zakay (1993) found attention effects only in the more difficult secondary tasks. If both the Addition and the Letter tasks are easy, enough processing time is left to keep track of the time. But if that were the case, participants would also have enough processing time left to do a secondary task when they do not have to attend to the time. A measure of dual tasking can be obtained by looking at these high-profit periods. We assume that as stimuli in the left box produce higher scores, people will react only to stimuli in the right box if they have spare capacity to do so. We therefore took as a measure of dual tasking the proportion of target stimuli in the right box to which the participant responded while there were also stimuli in the left box. Participants turned out to be able to achieve a level of 86% dual tasking in the Letter task but only 47% in the Addition task (see Figure 11). This shows that the Addition task does indeed require much more attention than the Letter task. According to the attentional gate theory, this difference should have an impact on time estimation. The quality of the estimate was, however, unaffected by task difficulty, whereas a shift in task difficulty (in the LLA and AAL conditions) caused only relatively small changes in the estimate without affecting its quality.

Taken together, our results are not consistent with the attentional gate theory, as this theory would predict a larger and different impact of the task-difficulty and switch manipulations. On the other hand, the internal clock theory does not cover this experiment because it has no explicit theory on how the clock interacts

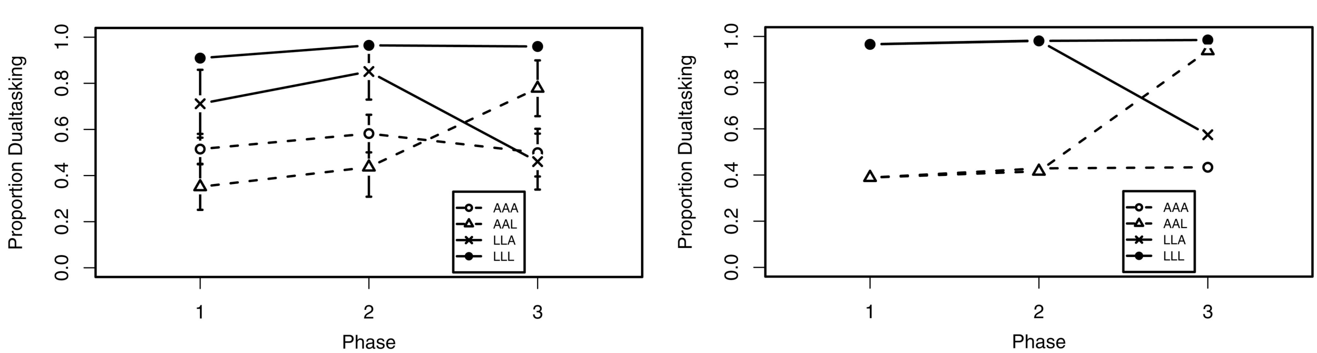

Figure 11. Proportion dual tasking in the three phases for four conditions. Left: Empirical data of Experiment 1. Right: Adaptive control of thought—rational (ACT–R) model fit. Bars are standard errors. AAA condition = three phases with the Addition task; AAL condition = two phases with the Addition task followed by one phase with the Letter task; LLA condition = two phases with the Letter task followed by one phase with the Addition task; LLL condition = three phases with the Letter task.

with aspects of cognition outside time management. The results also show pronounced effects of learning that are covered by neither theory. The results are, however, consistent with the expectations that we formulated at the beginning of this section: a clear learning effect and an unbiased transfer of the time estimate between tasks of different difficulty, as can be seen in Figure 9, middle.

The Integrated Model of Prospective Time Interval Estimation

In this section we discuss how the temporal module fits into a cognitive architecture and how it allows a fit of the data from Experiment 1. In addition to the temporal module, the cognitive part of the model builds on the ACT–R cognitive architecture, or more specifically, on earlier models of instance learning (Lebiere, Wallach, & Taatgen, 1998; Logan, 1988) and of central bottleneck theories of divided attention (Anderson, Taatgen, & Byrne, 2005; Pashler, 1994).

The ACT–R Architecture and the Role of Attention

Figure 12 shows a general overview of the ACT–R architecture, including the temporal module (Anderson et al., 2004). The center of the architecture is *procedural memory* (the production system), shown in the middle of the diagram. The production system has access to all the other modules in the system through *buffers*, each of which can hold only a single item of information. For example, the *temporal buffer* holds the current value of the accumulator, the *visual buffer* holds the currently attended visual stimulus, and the *retrieval buffer* holds the last element retrieved from *declarative memory*. The basic cycle of the central production system consists of the contents of all the buffers being matched against the rules stored in procedural memory. A single rule is then chosen on the basis of its *utility*, and this rule carries out its set of actions, which it communicates to the other modules through their respective buffers.

In the discussion of the Zakay (1993) model, we proposed one possible impact of attention: The secondary task also accesses the temporal module and disrupts timing in the primary task. However, we assume this lack of coordination between the tasks is a property of initial novice behavior that does not play much of a role in the DTT, in which the two tasks have to be done over an extended period of time.

A different impact of attention occurs when the contents of the temporal buffer are accurate but are not used by the rest of the system at the appropriate time because the system is too occupied with other tasks. The temporal module's output is only one of many buffers that the production system can match, and if it is busy with another subtask in a multi-tasking situation, it may fail to integrate the information from the temporal module with other processing. More specifically, during dual tasking, the model might be busy attending visual stimuli and responding to them with motor responses. Part of this process involves declarative memory to determine whether the stimulus is a target or a foil. As a consequence, attempts at reasoning about time (which also involve declarative memory) can be postponed or disrupted, acting like a system with a central bottleneck (Anderson et al., 2005; Pashler, 1994). In summary, in the Zakay (1993) model the temporal module itself is the contended resource, whereas in the present model the contended resource is declarative memory.

In this experiment the model has to divide its attention between three tasks: attending and responding to the left box, attending and responding to the right box, and attending to the time. Only two of these tasks are relevant at the same time: Either both the left and the right box have to be attended, or both the right box and the time. The model is mainly event-driven and responds to changes on the screen. When a new stimulus appears on the screen, the model attends to it and initiates a response. The only exception is

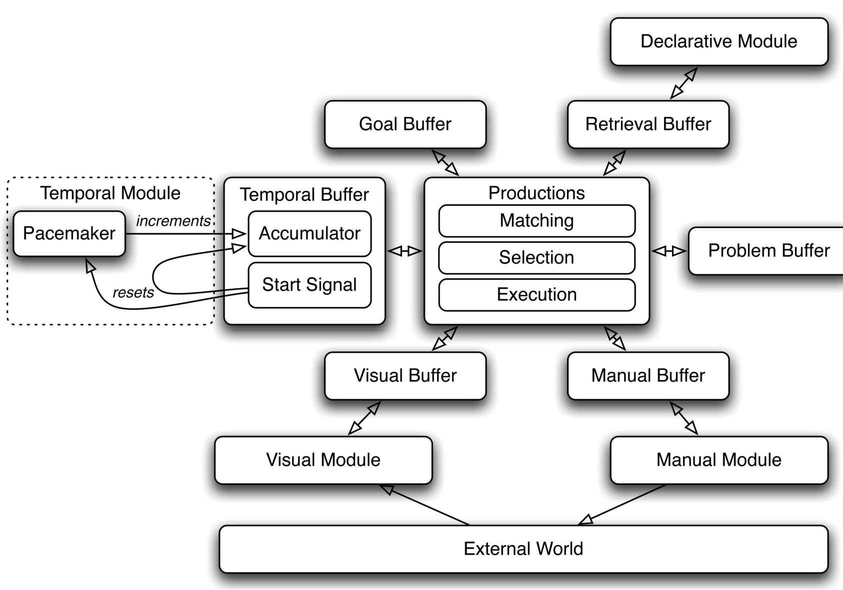

Figure 12. Interval timing as part of the adaptive control of thought—rational (ACT–R) architecture.

when the model is busy with a stimulus in the left box. In that case it will ignore stimuli in the right box until it is done with the stimulus in the left box. A stimulus in the left box can, on the other hand, interrupt processing in the right box. This reflects the fact that the score for the left box is 100 points and for the right box only 30 points. Attending to the time interval is initiated whenever the model has no stimulus to process. Because retrieving a past experience takes time, especially when these experiences are relatively new and still have a low activation, attending to the time will be interrupted when a new stimulus appears on the screen, making it necessary to restart the time estimation process once the stimulus has been processed.

Learning the Time Interval Through Instance Learning

Because the duration of the interval is initially unknown, the model has to determine it by trying out various intervals. One of the currently dominant theories of skill acquisition is *instance theory* (Logan, 1988). Within ACT–R, instance theory—or *instance learning*—is generally used to model situations of implicit learning, in which people are not directly aware they are learning something new but performance gradually increases with practice, such as in sequence learning (Lebiere & Wallach, 2001). This seems to fit the situation in this experiment, in which participants gradually built up a representation of the duration of the interval. Instance learning in ACT–R assumes that previous experiences, in this case of a specific time interval, are stored in (declarative) memory. For example, memory could contain an experience that 30 pulses is too short and another that 50 pulses is correct.

When the model sees *end of high period* in the left box, signaling the start of the interval, it starts generating time pulses, as illustrated in Figure 4. Whenever the model has time in between processing stimuli, it will attempt to retrieve a previous experience of pressing the test button at approximately the present time. If a successful experience is retrieved, the model will initiate a click on the test button. If a "too short" experience is retrieved, the model will do nothing. Finally, if no experience at all is retrieved for the present time, the model randomly decides to click the button or not. Experiences are then stored in ACT–R's declarative memory, which has an activation-based mechanism to model forgetting. However, if two experiences are identical (i.e., concerning the same judgment for the same number of pulses), their activations are combined. Because of the decay in memory, a particular experience sometimes has to be repeated a number of times before it can be retrieved at all, and retrieval will become faster with frequent use (see Anderson et al., 2004, for the details of declarative retrieval).

After the button has been clicked, the model judges whether the button click was successful. If stimuli appear in the left box, the present time is stored as successful, but if nothing happens, the present time is stored as "too short." Note that because of the nature of the task, late test clicks are judged as successful, even if they are 4 s late. However, the model will not tend to wait 4 s too long because it will tend to retrieve a successful time earlier. Early presses, however, are judged as failures ("too short"), even if they are early by only 100 ms. As the model accumulates more experiences, it becomes more accurate at estimating the right interval, but only within the boundaries of the accuracy of the temporal module itself (i.e., what is depicted by the dotted pure interval estimation distribution in Figure 7). In addition to that, experiences of button clicks around the 7-s mark are a mixed set of successes and failures, adding noise to the timing process.

Instance learning mainly captures the long-term effects of learning, a gradual accumulation of experience that slowly improves performance, as witnessed in the timing accuracy and the reduction of non-response trials over the experience. An alternative model would be to respond more directly to the previous trial by responding later if the previous trial was too early, or earlier if the previous trial was late. Such a strategy would result in short-term adaptation instead of long-term learning. However, there is little evidence in the data for short-term adaptation: When an estimate was made too early, the next estimate was also early in 64% (1,134 out of 1,764) of the cases, and when an estimate was late by 0.5 s or more, the next estimate would be late by at least 0.5 s in 53% (813 out of 1,544) of the cases.

Model Results

We used the same parameters for the temporal module as in all earlier models. The parameters that control the timing of the interaction with the interface (time to attend stimuli on the screen, timing of mouse actions) were left at their ACT–R default values. We estimated the parameter that controls the threshold at which elements in declarative memory are forgotten and the probability that the test button is clicked when no previous experience could be found to fit the score and time estimation data. The model was run 100 times for each of the four conditions. The results of the model were already shown in Figures 7–9 and 11 to make comparisons with the data easier.

Figures 7 and 9 show that the distribution of the time estimates of the model was very similar to that found in the data. The qualitative effects found in the analysis are also present in the model fits: There is an effect of learning on timing accuracy in terms of a 370-ms improvement in the absolute values of the deviations from optimal time between Phases 1 and 2 (compared to 356 ms in the data) but only a 120-ms improvement between Phases 2 and 3 (there was no improvement at all in the data). The differences in time estimation accuracy between the two tasks were very small in the model: On average the deviation on the Letter task was 149 ms shorter than on the Addition task; this corresponds well to the non-significant 94-ms difference in the data.

The shifts in average click time due to changes in the task were also produced by the model: After the task changed from Letter to Addition, the average first-click time was 402 ms later (726 ms in the data), and after the task changed from Addition to Letter, the average first-click time was 158 ms earlier (330 ms in the data). The model's explanation for these shifts is that processing an Addition stimulus takes more time than a Letter stimulus. When the moment arrives to make a click, the model still has to complete its response to the current stimulus, resulting in slightly earlier clicks for the Letter task and slightly later clicks for the Addition task. The main impact of the difficulty manipulation is on the proportion of trials in which no response is made at all. The model captured this phenomenon. Although the graphs in Figure 8 are hard to compare due to the noisiness of the data, the model did exhibit the two main effects of condition and learning, without an interaction, that are present in the data. Finally, Figure 11 shows that the model also correctly captures the dual-tasking results, confirming that accuracy of time estimation and the amount of attention that can be devoted to it are relatively independent.

The two expectations we formulated before the experiment were confirmed: There is a clear learning effect on the time estimates that participants make, and the accuracy of the time estimate transfers very well from one task to the other. The task difficulty manipulation had very little impact on the accuracy of time estimation: The distributions of the first-click times are very similar for all conditions. Instead, the impact of task difficulty is an increase in the proportion of non-response trials.

The key difference between the attentional gate theory and the ACT–R model is that the former predicts that the main effect of a harder secondary task is a shift in the time estimate, whereas the latter predicts that the main effect is an increased probability that there is no response at all. According to the attentional gate theory, a shift to an easier secondary task would produce responses that are too early, and vice versa, whereas the ACT–R model predicts that there is a change in non-response trials instead. The ACT–R model also predicts a small shift in time estimate due to the longer processing time of Addition versus Letter tasks. Both ACT–R predictions are confirmed by the data. Although the empirical results are not consistent with the attentional gate theory, it can be argued that it is not a strong test of the ACT–R model yet, because the task difficulty did not have a large impact on time perception and the model was fitted to the data. To build a stronger case, we conducted a second experiment such that a strong effect of task difficulty could be expected on the basis of the model, but one that was different from what the attentional gate theory would predict. Moreover, instead of fitting the model to the data, we made a model prediction, thereby avoiding the criticism that insufficiently constrained cognitive models can be made to fit any dataset (Roberts & Pashler, 2000).

Experiment 2

The second experiment was identical to Experiment 1 with one major change: There were now always stimuli in the left box, instead of only during a high-profit period. Correctly identifying the high-profit period now increased the score only for each hit in the left box. Because participants could work on the two tasks all the time, there were fewer opportunities to estimate the time interval. Nevertheless there was slack time to attend to interval estimation due to randomness in the task: The interval between subsequent stimuli was randomized, producing occasional gaps in presentation, and if both boxes displayed a foil, no response had to be made. The general expectation on the basis of the ACT–R model is that an increase in demand of the visual tasks produces an increase in trials in which no time estimation response is made. As we show in the model predictions, we expected that when participants started with the Addition task they would not be able to make enough estimates to determine the duration of the interval.

The Model

As indicated in the introduction, we used the model for Experiment 1 to make a prediction for Experiment 2 before doing the actual experiment,[4] with one slight modification: In Experiment 1 the feedback for a successful press on the test button was that stimuli started to appear in the left box, whereas in Experiment 2 it was the appearance of *HIGH* above the left box. We adjusted the model to be able to interpret the changed feedback correctly. Otherwise all parameters and production rules in the model were kept the same. The main qualitative predictions were as follows:

1. The Addition task is so hard that it is almost impossible to learn the correct interval. We expected the accuracy of the presses on the test button in Phase 1 to be at chance level in the conditions that start with the Addition task.

2. The Letter task leaves some time to attend and learn the interval, and therefore the model predicted that participants would be able to learn the interval and make reasonable estimates, although at a lower level of accuracy than in Experiment 1.

3. Time estimation has to compete with two other tasks (the left and the right box). The assumption in the model is that a new visual stimulus will interrupt any ongoing reasoning about time. Although reasoning about the time interval can be restarted as soon as the stimulus has been processed, time has passed in the meantime, making it necessary to start over again. In practice, this means that in many trials there will be no attempt to click the test button at all. This is especially true with the Addition task, in which the model predicts that in the majority of the cycles there will be no attempt at estimation, but also with the Letter task, in which the model predicts that no attempt will be made in approximately a quarter of the opportunities.

4. For the event in which the task shifts from Letter to Addition in the LLAA condition (see below), the model predicts no strong shift in the interval estimate (contrary to the attentional gate theory). The model even predicts that participants will do better on the Addition task after two blocks of the Letter task than after two blocks of the Addition task because the Letter task offers better learning opportunities. The attentional gate theory would predict the opposite, because it predicts that the estimate of the interval during the Letter task does not transfer to the Addition task.

Changes to the Task

The task was identical to the task in Experiment 1 with the following modifications. In Experiment 1, stimuli appeared in the left box only after the participant had clicked the test button in the high-profit period. In Experiment 2, stimuli appeared in the left box all the time but yielded 100 points only after the participant had clicked the test button in the high-profit period; otherwise they yielded 30 points, the same score as for the right box. Stimuli in both boxes appeared for 1,200 ms and were separated by a random interval between 0 and 2,000 ms. Feedback with respect to the high-profit period was given by the display of *HIGH* above the left box after the participant had clicked the test button in the high-profit period. The *HIGH* text was removed at the end of a high-profit period.

The experiment had four between-subject conditions: four phases of five blocks with the Letter task (LLLL), four phases of five blocks with the Addition task (AAAA), two phases of five blocks with the Letter task followed by two phases of five blocks

[4] On March 7, 2005, an email was sent to all the members of the ACT–R community with a web-link to the prediction. We started the experiment the week after that.

with the Addition task (LLAA), and two phases of five blocks with the Addition task followed by two phases of five blocks with the Letter task (AALL). To prevent participants from counting the stimuli in the left box to help their time estimate, the inter-stimulus intervals of the left box were randomized with the same randomization process as was used for the right box. Otherwise the procedure was identical to the procedure of Experiment 1. Further details on the experiment can be found in the Appendix.

Results

In this section we discuss the results of Experiment 2 alongside the predictions of the model. As the model predictions are overall similar to the experimental data, we highlight only the aspects of the model that are of particular interest. Figure 13 shows the distributions of first-click times for the first phase of the experiment. Compared to those of Experiment 1, the distributions are much flatter, indicating that on many trials no response at all was made. This lack of response is made explicit in Figure 14, which shows the non-responses over the whole experiment. Consistent with the prediction of the model, participants had great trouble making accurate time estimates at all in Phase 1 of the Addition condition. This clearly shows that the task manipulation in this modified experiment had a major impact on time estimation, which was correctly predicted by the model. The explanation the model offers is that in the Addition task, the time estimation process is interrupted so often (much more often than in Experiment 1) that the few experiences it gets are spaced apart too far to produce a stable representation of the duration of the interval (in ACT–R terms, the activation of the experiences has dropped below the retrieval threshold at the time they are needed). When the model does push the test button it is therefore a blind guess.

The non-response data concur with these findings: Both task difficulty and learning have a significant impact on the non-response proportion (see Figure 14). These results confirm what was found in Experiment 1: The difficulty of the task influences how often participants make a response at all. However, the magnitude of the effect is much larger in this experiment. In addition, here there is an interaction between phase and condition (it was absent in Experiment 1 because by Phase 3 the non-response rates were all low). The interaction in this experiment is due to the fact that in two of the conditions the task changed halfway, and this affects the non-response rate (fewer responses when the task changes to Addition and more responses when the task switches to Letter). The model predicted both of these effects correctly, even though the exact fit between model and data is hard to assess because the data are quite noisy. Instead of predicting a shift in non-response trials, the attentional gate theory would predict a shift in the estimate with a change in task difficulty. In Experiment 1 we saw a small shift in the estimate, so according to the attentional gate theory this shift should be much larger in Experiment 2. However, as can be seen in the results (see Figure 15), there was no shift at all.

Discussion

In the introduction to Experiment 2, we stated that the ACT–R model makes four general predictions. The first prediction was that participants would not be able to make accurate time predictions in the Addition task and that the estimates they make would be at chance level. This prediction was confirmed by the data: In the first phase with the Addition task, participants often failed to make a prediction at all, and if they made one it was at chance level. The second prediction was that participants would be able to make proper time estimations in the Letter task, although at a lower level of accuracy than in Experiment 1. This prediction was also confirmed by the data. The third prediction of the model was that the main impact of the increased difficulty is an increase in the number of non-responses, trials in which participants make no attempt at all to give an estimate. This prediction was also confirmed by the data. The fourth prediction concerned the shift from Letter to Addition tasks in the LLAA condition, in which the model predicted no change in the time estimation after the shift. Moreover, it stated that participants would perform better in Phases 3 and 4 of the LLAA condition than in the AAAA condition. This prediction was only partially confirmed: There was indeed no change in the estimated interval after the shift, but participants' performance turned out to be the same in the AAAA and LLAA conditions with respect to Phases 3 and 4. For all four predictions (apart from the mispredicted part of Prediction 4), the model was able to make not only a correct qualitative prediction but also an accurate quantitative prediction.

Although the four predictions may not sound particularly counter-intuitive, they either cannot be explained by or are incon-

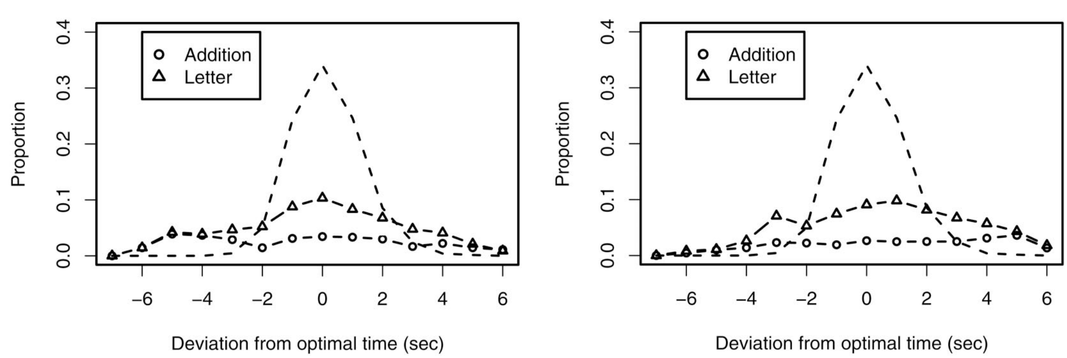

Figure 13. Distributions of the moment of the first click on the test button in the first phase of the experiment. Left: Empirical data of Experiment 2. Right: Adaptive control of thought—rational (ACT–R) model predictions. sec = seconds.

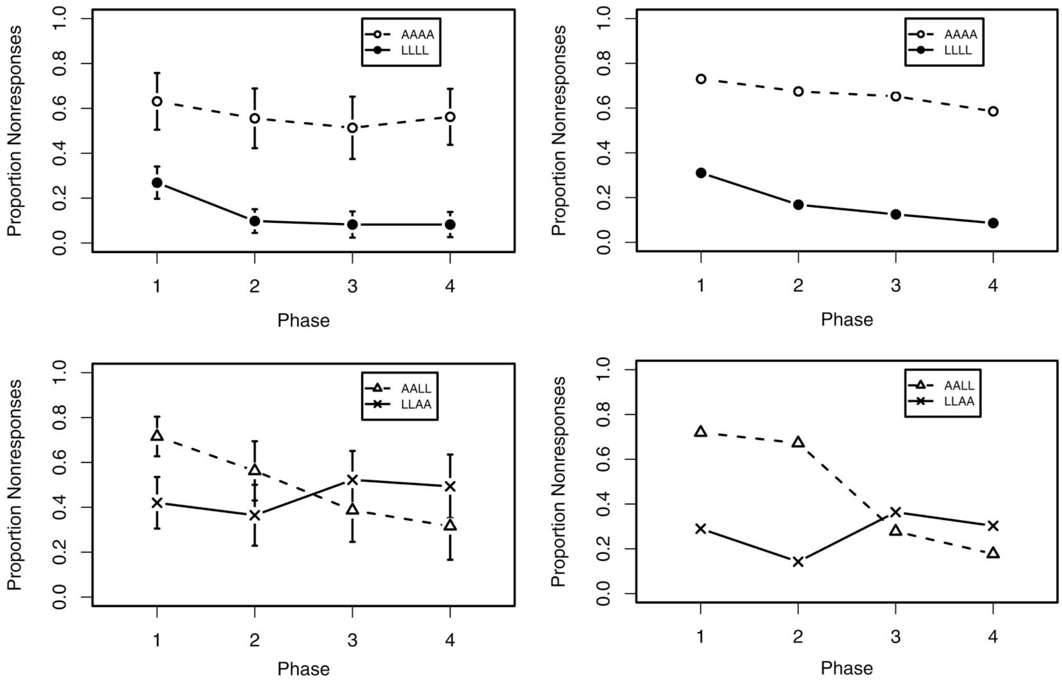

Figure 14. Proportion of non-response by condition and block. Left: Empirical data of Experiment 2. Right: Adaptive control of thought—rational (ACT–R) model predictions. AAAA condition = four phases of five blocks with the Addition task; LLLL condition = four phases of five blocks with the Letter task; AALL condition = two phases of five blocks with the Addition task followed by two phases of five blocks with the Letter task; LLAA condition = two phases of five blocks with the Letter task followed by two phases of five blocks with the Addition task.

sistent with the attentional gate theory. Consistent with the ACT–R prediction, the main impact of the difficulty manipulation was on the proportion of non-response trials. In the trial that started with Addition, the model correctly predicted that only in about 30% of the trials an estimate would be made and that this would not be enough for instance learning to form a stable representation within the first phase of the experiment. Although an absence of accurate responses in the conditions that start with Addition is not inconsistent with the attentional gate theory, it is also not predicted by it. However, the absence of a shift in mean response time in the LLAA condition, in which participants do make reasonable time estimations even after the change to the Addition task, is not compatible with the attentional gate theory. Though we found small effects in those directions in Experiment 1, they were absent in Experiment 2, whereas according to the attentional gate theory they should have been larger. We believe that with the four difficulties of the task (Letter and Addition in Experiment 1 and Letter and Addition in Experiment 2), we had good coverage of the various levels of difficulty, none of which produced any large shifts in time estimates. Instead the shifts had large impacts on the number of non-response trials, something that is not predicted by the attentional gate theory.

General Discussion

In this article we have presented a theory of time estimation that is integrated into a larger theory of cognition with a focus on attention and learning. The core of the theory is a simple pacemaker–accumulator module that counts pulses as time passes, similar to the theory described by Matell and Meck (2000). The main twist in the mechanism is that the duration of the pulses increases with the interval, producing a logarithmic scale, thereby allowing the module to produce the scalar property of the variance in the estimate. The behavior of the module is controlled by three parameters that we estimated on the basis of the Rakitin et al. (1998) data and which were confirmed in accurately predicting bisection data (Penney et al., 2000).

In most time estimation studies, interval estimation is the main task. However, time estimation should also be studied in contexts in which time estimation itself is secondary to a main task, because this corresponds to the natural role time estimation plays in everyday life. The success of the model presented here does not depend on the actual mechanism of time estimation itself but on the way it interacts with other aspects of cognition.

The variant of the pacemaker–accumulator mechanism we have chosen accurately models the scalar property in the variance in time estimation. Any mechanism with the same properties can in principle replace it, for example, an oscillator–coincidence mechanism (Matell & Meck, 2000) or a process-decay mechanism (Staddon & Higa, 1999; see our footnote 2). Once integrated into a larger framework, it can be used to model complex tasks in which time estimation is only a component and to make accurate predictions on the outcome of these complex tasks.

The ACT–R architecture models attention, or more specifically, divided attention, by having a control structure of the task that

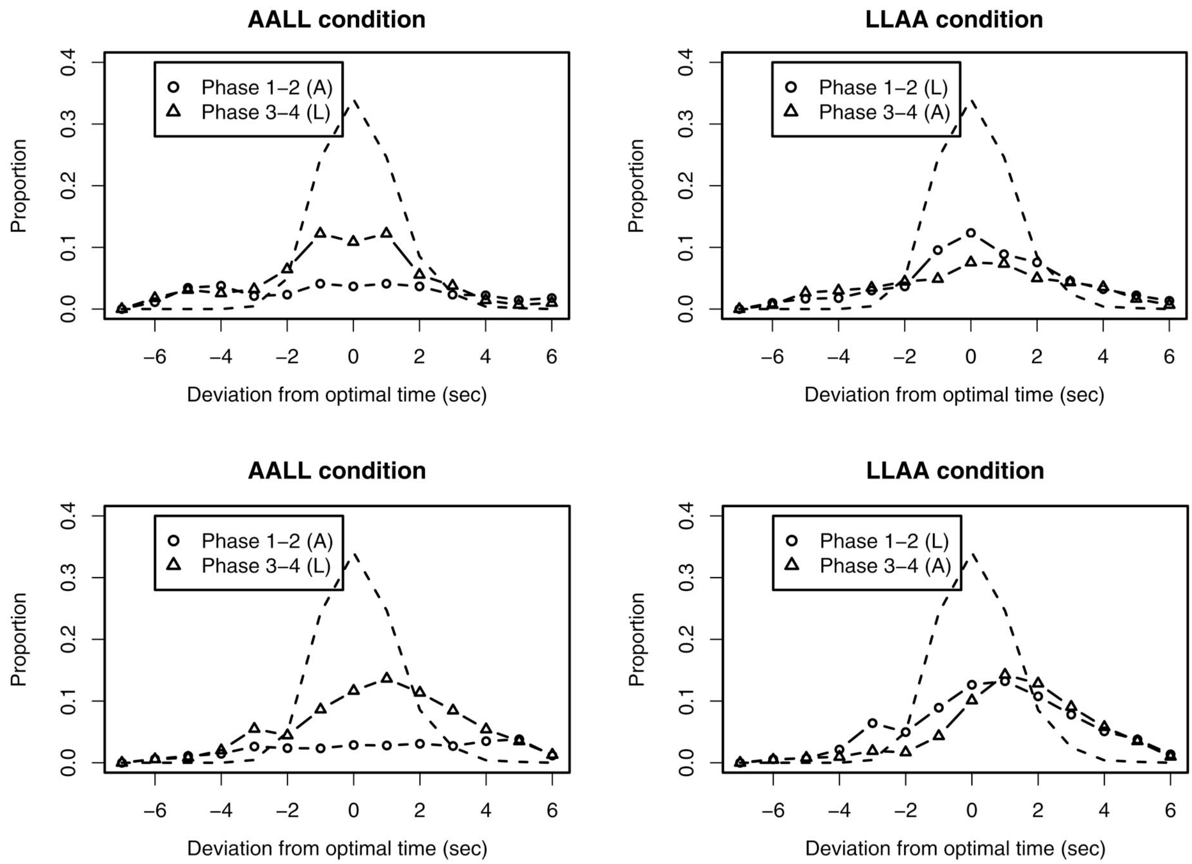

Figure 15. Changes in click time distributions for the conditions in which there was a change in task. Top: Empirical data of Experiment 2. Bottom: Adaptive control of thought—rational (ACT–R) model predictions. AALL condition = two phases of five blocks with the Addition (A) task followed by two phases of five blocks with the Letter (L) task; LLAA condition = two phases of five blocks with the Letter task followed by two phases of five blocks with the Addition task; sec = seconds.

determines which cognitive modules participate in determining the next action. This control structure is necessary for prioritization of the subtasks and prevention of interference. It produces behavior similar to central bottleneck theories of attention (Pashler, 1994). In the model of Zakay's experiment (Zakay, 1993), we assumed that an improper task structure led to a situation in which both tasks—time estimation and the secondary task—had access to the timing module, giving the secondary task the opportunity to cause shifts in the estimate. In the DTT task, in which there is enough opportunity to acquire the right task structure, timing difficulties occur because operations on the time module have to compete with other operations. Both time estimation and processing of visual stimuli need access to declarative memory, and if we assume the latter takes priority, the model produces the effects of attention we found in the experiments: an increase in variability of the estimate. This is consistent with other experiments in which the same estimate has to be made multiple times, like Brown (1997) and Rakitin (2005). The model we have presented here manages to capture all these aspects using general mechanisms of attention from the ACT–R architecture. In addition, in the case of Experiment 2, in which the task in the first phase was Addition, it correctly predicted a situation in which timing attempts were made but were too infrequent to obtain a clear representation of the duration of the interval.

Learning of time estimates was modeled with an instance retrieval strategy: Experiences with a certain time interval were stored in declarative memory and could be retrieved for future decisions. Accumulating experiences improved the estimate and increased its activation in memory, which sped up its retrieval. This learning process played a role in all models discussed in this article, even though we have simplified the process in the first two models. Although the learning aspect of time estimation is not as contended in the literature as the role of attention, we consider it to be a vital part of a theory of time perception.

Although the model for the DTT was specifically designed for one task, the principles of learning and attention are general enough to be extended to other tasks. We have already successfully modeled operating a typing device while driving a car based on the temporal module introduced here and the same principles of instance learning and dual tasking (Salvucci et al., 2006). This model adapted the time it was willing to look away from the road to the changing demands of the driving task. Other potential situations in which timing is relevant include discovering how to interact with new devices, for example, determining how long to turn the key before the car engine starts, how long to wait after pushing the power button on a camera before it is ready to make a photo, and how long to wait before putting the meat in the pan while the oil heats.

An open question that remains is whether there is only a single timer or whether our cognitive system can time multiple things at the same time. For example, in some of the conditions in the bisection experiments (Penney et al., 2000), participants had to estimate two slightly staggered intervals at the same time and performed almost identically to when they had to estimate only a single interval. Although it is possible that separate timers track both intervals, it is also possible a single timer is used to track all the intervals in between events and that explicit reasoning is used to find the answer. Such an explicit process cannot simply subtract two time estimation counts, at least according to our model, because of the logarithmic scale of the time representation. Studies from the animal literature support the notion that animals, at least, can track multiple intervals (e.g., Jozefowiez, Cerutti, & Staddon, 2006), although it is unclear whether multiple timers are involved or other processes that reflect on time-related behavior.

Experiments by Rakitin (2005) suggest that explicit reasoning about time intervals can play a role even in simple time estimation experiments. Rakitin found that when participants had to do a choice-reaction task together with estimating a time interval after they had been trained on the interval first, they tended to serialize the two tasks (first the choice-reaction task and then the time estimation task). Because of the serialization, the internal estimation process had to estimate a shorter interval (the original interval minus the time to do the choice-reaction task). However, only the variability of the time estimate was affected, and not the mean, suggesting participants strategically shortened the interval using a process Rakitin called *temporal discounting*. Temporal discounting proved to be inaccurate in two of Rakitin's final experiments in which the start of the interval and the presentation of the choice-reaction stimulus were separated by a variable stimulus-onset asynchrony. Although participants in Rakitin's experiments serialized the tasks in a dual-task paradigm, participants in the DTT task did not, suggesting that strategic choice is possible depending on the task. It should be possible to capture Rakitin's results with an ACT–R model, which would adapt a serial control structure for the task and an explicit strategy for temporal discounting.

A related question is whether the timer can be stopped and started again later. Fortin, Bédard, and Champagne (2005) found that estimations of time intervals that were interrupted depended on where in the interval the interruption was placed. This makes it unlikely that the timer can just be stopped and restarted. Instead, an explanation in which explicit reasoning about time determines the estimate may be more appropriate to explain these results. Although we do not offer a theory of explicit reasoning about time in this article, the ACT–R architecture in general has many mechanisms that can help in building such theories. However, we consider it unlikely that a single theory of explicit reasoning about time can cover all phenomena. Instead, each particular phenomenon will have to be explained by assuming task-specific strategies, for example, retrograde time estimation for which explicit reasoning instead of an internal mechanism is responsible (Block & Zakay, 1997).

As a final note, a criticism of fitting cognitive models is that given enough free parameters, any dataset can be modeled (Roberts & Pashler, 2000). We have tried to counter this criticism by estimating the three parameters of the temporal module for the very first task only and using those parameters for all other models. Complex tasks require complex models, which makes it necessary to estimate additional parameters not related to time estimation. For the DTT model, we used the data from Experiment 1 to estimate the non-temporal parameters. This model was then used to make a true prediction for Experiment 2. Although Experiment 1 guided this prediction, the predictions for Experiment 2 were novel, for example, that time estimation in the Addition task completely breaks down. Given the success of the model, we are confident that the temporal module can be used in all situations in which intervals on the order of 1–30 s have to be estimated.

References

Anderson, J. R., Bothell, D., Byrne, M., Douglass, D., Lebiere, C., & Qin, Y. (2004). An integrated theory of mind. *Psychological Review, 111,* 1036–1060.

Anderson, J. R., Taatgen, N. A., & Byrne, M. (2005). Learning to achieve perfect timesharing: Architectural implications of Hazeltine, Teague, and Ivry (2002). *Journal of Experimental Psychology: Human Perception and Performance, 31,* 749–761.

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

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

Buhusi, C. V., & Meck, W. H. (2006). Internal timing with gaps and distractors: Evaluation of the ambiguity, switch, and time-sharing hypotheses. *Journal of Experimental Psychology: Animal Behavior Processes, 32,* 329–338.

Fortin, C., Bédard, M. C., & Champagne, J. (2005). Timing during interruptions in timing. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 276–288.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84,* 279–335.

Grosjean, M., Rosenbaum, D. A., & Elsinger, C. (2001). Timing and reaction time. *Journal of Experimental Psychology: General, 2,* 256–272.

Harrell, F. E., Jr. (2001). *Regression modeling strategies, with applications to linear models, logistic regression, and survival analysis.* New York: Springer.

Hicks, R. E., Miller, G. W., & Kinsbourne, M. (1976). Prospective and retrospective judgments of time as a function of amount of information processed. *American Journal of Psychology, 89,* 719–730.

Jozefowiez, J., Cerutti, D. T., & Staddon, J. E. R. (2006). Timescale invariance and Weber's law in choice. *Journal of Experimental Psychology: Animal Behavior Processes, 32,* 229–238.

Kushleyeva, Y., Salvucci, D. D., & Lee, F. J. (2005). Deciding when to switch tasks in time-critical multitasking. *Cognitive Systems Research, 6,* 41–49.

Lebiere, C., & Wallach, D. (2001). Sequence learning in the ACT-R cognitive architecture: Empirical analysis of a hybrid model. In R. Sun & C. L. Gilles (Eds.), *Sequence learning: Paradigms, algorithms, and applications* (pp. 188–212). Berlin: Springer.

Lebiere, C., Wallach, D., & Taatgen, N. A. (1998). Implicit and explicit learning in ACT-R. In F. Ritter & R. Young (Eds.), *Proceedings of the Second European Conference on Cognitive Modelling* (pp. 183–189). Nottingham, UK: Nottingham University Press.

Lejeune, H. (1998). Switching or gating? The attentional challenge in cognitive models of psychological time. *Behavioural Processes, 44,* 127–145.

Lejeune, H. (2000). Prospective timing, attention and the switch: A response to "Gating or Switching? Gating is a better model of prospective timing" by Zakay. *Behavioural Processes, 52,* 71–76.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Matell, M. S., & Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *BioEssays, 22,* 94–103.

Michon, J. A., & Jackson, J. L. (1985). Introduction: The psychology of time. In J. A. Michon & J. L. Jackson (Eds.), *Time, mind, and behavior* (pp. 2–17). Berlin: Springer.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 16,* 220–244.

Pavlov, I. P. (1927). *Conditioned reflexes.* Oxford, UK: Oxford University Press.

Penney, T. B., Gibbon, J., & Meck, W. H. (2000). Differential effects of auditory and visual signals on clock speed and temporal memory. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1770–1787.

Rakitin, B. C. (2005). The effects of spatial stimulus-response compatibility on choice time production accuracy and variability. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 685–702.

Rakitin, B. C., Gibbon, J., Penney, T. B., Malapani, C., Hinton, S. C., & Meck, W. H. (1998). Scalar expectancy theory and peak-interval timing in humans. *Journal of Experimental Psychology: Animal Behavior Processes, 24,* 15–33.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review, 107,* 358–367.

Salvucci, D. D., Taatgen, N. A., & Kushleyeva, Y. (2006). Learning when to switch tasks in a dynamic multitasking environment. *Proceedings of the seventh international conference on cognitive modeling* (pp. 268–273). Trieste, Italy: Edizioni Goliardiche.

Staddon, J. E. R. (2005). Interval timing: Memory, not a clock. *Trends in Cognitive Sciences, 9,* 312–313.

Staddon, J. E. R., & Higa, J. J. (1999). Time and memory: Towards a pacemaker-free theory of interval timing. *Journal of the Experimental Analysis of Behavior, 71,* 215–251.

Staddon, J. E. R., & Higa, J. J. (2006). Interval timing. *Nature Reviews Neuroscience, 7,* 1764.

Thomas, E. A., & Weaver, W. B. (1975). Cognitive processing and time perception. *Perception & Psychophysics, 17,* 363–367.

Zakay, D. (1989). Subjective time and attentional resource allocation: An integrated model of time estimation. In I. Levin & D. Zakay (Eds.), *Time and human cognition* (pp. 365–397). Amsterdam: Elsevier.

Zakay, D. (1990). The evasive art of subjective time measurement: Some methodological dilemmas. In R. A. Block (Ed.), *Cognitive models of psychological time* (pp. 59–84). Hillsdale, NJ: Erlbaum.

Zakay, D. (1993). Time estimation methods: Do they influence prospective duration estimates? *Perception, 22,* 91–101.

Zakay, D. (2000). Gating or switching? Gating is a better model of prospective timing (a response to "switching or gating" by Lejeune). *Behavioural Processes, 50,* 1–7.

Zakay, D., & Block, R. A. (1997). Temporal cognition. *Current Directions in Psychological Science, 6,* 12–16.

Zakay, D., & Block, R. A. (2004). Prospective and retrospective duration judgments: An executive-control perspective. *Acta Neurobiologiae Experimentalis, 64,* 319–328.

Appendix

Details and Analyses of the Two Experiments

Experiment 1

Participants

Thirty-two students from Carnegie Mellon University (17 men and 15 women) volunteered to participate in the experiment. Volunteers were paid for their participation.

Analysis of Variance of the Effect of Learning and Task Difficulty on Quality of the Time Estimate

To determine the effects of learning and task difficulty on the quality of the time estimates, we analyzed the absolute values of the deviations from optimal time for the trials in which the participant responded before the end of the trial was indicated by the word *end* appearing above the left box. As taking the absolute value introduced skew in the distribution of the data, we log-transformed the absolute deviations. Analyses of variance with phase as within-subject factor, condition as between-subject factor, and subjects as random factors revealed only a main effect of phase, $F(28, 3) = 7.88$, $MSE = 0.91$, $p = .009$. Paired t tests showed that this effect is due to learning between Phases 1 and 2 (an improvement of 356 ms), $t(31) = 3.16$, $p = .004$, but not between Phases 2 and 3 ($t < 1$). An alternative indication of the quality of the estimate is the variability in the time estimate (Brown, 1997). Analyses of variance of the standard deviation with phase as within-subject factor, condition as between-subject factor, and subjects as random factors also revealed only a main effect of phase, $F(28, 3) = 30.4$, $MSE = 1.86$, $p < .001$. Paired t tests showed that this effect is due to learning between Phases 1 and 2 (a decrease in standard deviation from 229 to 177), $t(31) = 6.13$, $p < .001$, but not between Phases 2 and 3 ($t < 1$).

Logistic Regression of Non-Response Proportions

The data were subjected to a logistic regression (Harrell, 2001) with proportion non-responses as response variable and phase and

(*Appendix continues*)

condition as predictors. In addition to these main predictors, an interaction between phase and condition was included so that we could test for differential effects of task in different phases of the experiment. A significant main effect was found for phase, $\chi^2(4) = 22.96$, $p < .001$, indicating a decrease in non-responses due to learning. A main effect was also found for condition, $\chi^2(6) = 21.33$, $p = .002$, but no effect was found for the interaction.

Changes in the Accuracy of Time Estimation Due to Changes in Task

A paired *t* test of the mean click-time in Phases 2 and 3 in the AAL condition revealed that the shift in click moment was significant, from 701 ms to 371 ms, $t(7) = -2.61$, $p = .035$, as was the shift between Phases 2 and 3 in the LLA condition, from −624 ms to 102 ms, $t(7) = 3.63$, $p = .008$.

Dual Tasking

A Welch two-sample *t* test between the average dual-tasking scores in Phases 1 and 2 for the two tasks showed significantly more dual tasking in the Letter task than in the Addition task, $t(29.954) = 4.252$, $p < .001$.

Experiment 2

Participants

Forty students from Carnegie Mellon University (21 men and 19 women) volunteered to participate in the experiment. Volunteers were paid for their participation.

The Effect of Task Difficulty on the Quality of the Time Estimate

Of the 294 estimates that were made with the Addition task in the first phase, 148 responses were within 3 s of the optimal time and 146 responses were more than 3 s early or late, suggesting that accuracy was indeed at chance level, because according to the internal-clock distribution it is possible to make almost all estimates within these 3 s. In the Letter condition, participants were able to make slightly better estimates but performed at a lower level than in Experiment 1. Nevertheless, 408 of the 590 responses made were within 3 s of the optimum, and only 182 were outside of it.

Logistic Regression of Non-Response Proportions

A logistic regression of the non-response data with proportion non-responses as response variable and condition, phase, and the interaction between condition and phase as predictors revealed a significant interaction between condition and phase, $\chi^2(3) = 8.89$, $p = .03$, as well as significant main effects of phase, $\chi^2(4) = 14.48$, $p = .006$, and of condition, $\chi^2(6) = 33.81$, $p < .001$.

Changes in the Accuracy of Time Estimation Due to Changes in Task

A paired *t* test on the mean first-click time in Phases 2 and 3 in the AALL condition revealed that there was no significant shift in first-click time, from −503 ms to −965 ms ($t < 1$), nor between Phases 2 and 3 in the LLAA condition, from 245 ms to −345 ms, $t(7) = 1.66$, *ns*.

Received June 8, 2006
Revision received January 22, 2007
Accepted January 23, 2007 ■