Cognition 119 (2011) 81–95

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Effects of practice on task architecture: Combined evidence from interference experiments and random-walk models of decision making

Juan E. Kamienkowski [a], Harold Pashler [b], Stanislas Dehaene [c,d,e], Mariano Sigman [a,*]

[a] Laboratorio de Neurociencia Integrativa, Departamento de Fisica, FCEyN UBA and IFIBA, Conicet; Pabellon 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina
[b] Department of Psychology 0109, University of California, San Diego, La Jolla, CA 92093, USA
[c] INSERM, U562, Cognitive Neuroimaging Unit, F-91191 Gif/Yvette, France
[d] CEA, DSV/12BM, NeuroSpin Center, F-91191 Gif/Yvette, France
[e] Collège de France, F-75005 Paris, France

ARTICLE INFO

Article history:
Received 5 April 2009
Revised 27 October 2010
Accepted 17 December 2010
Available online 4 February 2011

Keywords:
Learning
Task architecture
Dual-task performance
Psychological refractory period
Response time distributions
Cognitive processes

ABSTRACT

Does extensive practice reduce or eliminate central interference in dual-task processing? We explored the reorganization of task architecture with practice by combining interference analysis (delays in dual-task experiment) and random-walk models of decision making (measuring the decision and non-decision contributions to RT). The main delay observed in the Psychologically Refractory Period at short stimulus onset asynchronies (SOA) values was largely unaffected by training. However, the range of SOAs over which this interference regime held diminished with learning. This was consistent with an overall shift observed in single-task performance from a highly variable decision time to a reliable (non-decision time) contribution to response time. Executive components involved in coordinating dual-task performance decreased (and became more stable) after extensive practice. The results suggest that extensive practice reduces the duration of central decision stages, but that the qualitative property of central seriality remains a structural invariant.

1. Introduction

Several cognitive theories share the hypothesis that most mental and neural operations are modular and a dedicated architecture is required to establish flexible links amongst them (Baars, 1989; Chun & Potter, 1995; Dehaene, Kerszberg, & Changeux, 1998; Posner, 1994; Shallice, 1988). It has been proposed that this flexible architecture, capable of routing information according to any arbitrary program (task-setting) may result in serial information processing bottlenecks (Zylberberg, Fernandez Slezak, Roelfsema, Dehaene & Sigman, 2010). Processing bottlenecks are indeed ubiquitous in dual-task performance. For instance, when two tasks are presented simultaneously or sequentially at a short interval a systematic delay observed in the execution of the second task, a phenomenon referred to as the Psychological Refractory Period (Pashler & Johnston, 1989; Smith, 1967; Telford, 1931).

1.1. Mapping the PRP bottleneck

The exact nature of the processes causing the PRP bottleneck has been debated. A typical observation in the PRP design is that response time to the first task (RT1) is little affected while response time to the Task 2 (RT2) is greatly slowed as SOA is decrease (with a slope approaching -1). This can easily be explained in terms of a sequential processing scheme in which certain aspects of Task 2 cannot proceed until Task 1 is completed. Experiments investigating which aspects of Task 2 can proceed in parallel and which reflect serial queuing have mapped the bottleneck to the response selection process (Kamienkowski & Sigman, 2008; Pashler, 1984).

* Corresponding author.
E-mail address: sigman@df.uba.ar (M. Sigman).

0010-0277/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2010.12.010

However, while the response selection bottleneck is the principal source of the PRP, both psychophysical and physiological evidence have suggested systematic departures from the simple sequential bottleneck model (Allport, Styles, & Hsieh, 1994; De Jong, 1993, 1995; Jentzsch, Leuthold, & Ulrich, 2007; Logan & Gordon, 2001; Meiran, Chorev, & Sapir, 2000; Ruthruff, Pashler, & Klaassen, 2001; Sigman & Dehaene, 2006). In a classic PRP experiment, responses to Task 1 are independent of SOA, but they are slower than when performing the task in isolation (Jiang, Saxe, & Kanwisher, 2004; Sigman & Dehaene, 2005). We reasoned that this could be related to an executive control stage engaged before the execution of the first task. We hypothesized that in situations in which task order is unknown, this executive time should increase, reflecting a hierarchical decision processes: first, which task to respond to, and second, the specific decision involved in each task. This hypothesis was verified in a new series of experiments in which we concluded that in a situation of task uncertainty, executive components (engaging and disengaging in a task) had to be incorporated in order to account for a broad range of behavioral observations (Sigman & Dehaene, 2006).

Evidence for the involvement of such executive components could also be derived from human electrophysiological studies of the PRP. In an event-related potential (ERP) study in which a visual number comparison task was performed as Task 1 and an auditory pitch discrimination task was performed as Task 2, it was found that the peak of an early sensory component of Task 2 (Auditory N1 wave) occurred at a fixed delay after S2 presentation, indicating that certain perceptual stages of Task 2 can occur in parallel with Task 1. By contrast, the peak of the P3 wave, another ERP component which relates mostly to distributed parietal, temporal and frontal sources and thought to be involved in working memory, flexible routing of information and conscious perception (Donchin & Coles, 1998), showed a strictly serial delay. While this was in very good accordance with the predictions of the bottleneck model (Sigman & Dehaene, 2008), several other observations deviated from this simple model. First, the amplitude of the sensory N1 component of the second task decreased slightly during the interference regime. Second, the temporal course of the N1 component of Task 2 started prior to stimulus presentation, probably reflecting task expectation and preparation. Finally, a Task 2 related P3 component emerged at long SOAs, even before the Task 2 stimulus (auditory tone) was presented. This anticipatory component peaked around 500 ms, thus coinciding closely with the end of the visual P3 evoked by Task 1 (Sigman & Dehaene, 2008). This ERP sequence is compatible with the hypothesis that as soon as Task 1 was completed, subjects re-oriented their attention to prepare for Task 2, reflecting an executive component of task engagement (De Jong, 1993; Logan & Gordon, 2001; Meiran et al., 2000; Ruthruff et al., 2001; Sigman & Dehaene, 2006). In addition, it suggests that the absence of attentional top-down control may explain the amplitude attenuations observed during interference (Gilbert & Sigman, 2007). Overall, these data indicate that PRP experiments involve both a central bottleneck and an active process of task-oriented attention.

1.2. Can the PRP bottleneck be bypassed? Effects of practice on dual-task interference

Another unsolved matter which has attracted the attention of many scientists in cognitive psychology is whether central resources can be bypassed with extensive practice or in very "natural" stimulus–response mappings (McLeod, 1977; Posner & McLeod, 1982) such as responding with the right-hand to a right pointing arrow (Greenwald & Shulman, 1973; Lien, McCann, Ruthruff, & Proctor, 2005; Pashler, Carrier, & Hoffman, 1993; Schumacher, Seymour, Glass, Kieras, & Meyer, 2001). Recent results suggest that even under conditions of high ideomotor compatibility, the locus of the central processing bottleneck may be reduced but not completely eliminated (Lien et al., 2005). This suggests that establishing a temporary mapping between otherwise independent processors involves the engagement of a strictly serial processing stage which can be drastically reduced for highly practiced or non-arbitrary tasks (Greenwald, 2003; Lien, Proctor, & Allen, 2002; Lien et al., 2005).

Logan and colleagues have extensively studied the process of automatization, using an alphabet arithmetic task (e.g. H + 3 = K) (Compton & Logan, 1991). Based on subjective reports and on an analysis of the time-course of the response time variability during the course of learning, they provided substantial evidence in favor of a race model. According to this model, different strategies to solve the task co-occur: an algorithmic computation and a memory retrieval process. These two mechanisms operate simultaneously and the selection process is determined by a race. During the course of learning, memory retrieval is consolidated and becomes faster than the slow algorithmic computation, thus dominating the race and leading to automatic performance (Compton & Logan, 1991). An important assumption of such model is that practice does not affect the qualitative organization of the system, but rather changes the parameters of an invariant architecture. Evidence for such continuous progression in the automaticity process with practice came from a study in which an alphabet arithmetic task, at different stages of practice, was performed concurrently with a speech task (Klapp, Boches, Trabert, & Logan, 1991a, 1991b).

1.3. Random-walk models can decompose processing stages in a cognitive task

Virtually all PRP research – including the study of the effects of practice – has focused exclusively on mean RTs. It is possible however, that certain effects of practice do not directly affect the mean response time, but rather result in a change of the relative contributions of distinct processing stages to RT. Alternatively, of course, learning could result in a combination of both effects. How can one parse a task, simply relying on response time information, into different processing stages and understand the relative contribution of each processing stage to response time?

A separate psychological research tradition seeks to answer these questions, investigating how the decision to respond is achieved. The decision-making process has been modeled as a noisy integrator that accumulates evidence provided by the sensory system (Gold & Shadlen, 2001; Link & Heath, 1975; Luce, 1986; Machens, Romo, & Brody, 2005; Ratcliff, 1988; Reddi & Carpenter, 2000; Schall, 2000; Schwarz, 2001; Shadlen & Newsome, 1996; Usher & McClelland, 2001; Vickers, 1970; Wong & Wang, 2006).

Although many variants have been proposed, the basic idea is that perceptual evidence is stochastically accumulated in time. A decision results from a random walk of an internal abstract variable. Indeed, in many circumstances, such a decision mechanism can be optimal in the sense that it maximizes the overall likelihood of a correct classification of the stimuli (Edwards, 1965; Laming, 1968).

Within this framework, one can choose to model RT distributions with elaborated or with comparatively simpler models. Detailed reviews of the use of diffusion models to account for RT distributions can be found in (Luce, 1986; Ratcliff & Smith, 2004; Ratcliff & Tuerlinckx, 2002). As a brief summary, we present here some variants which are relevant for the present work.

Two-barrier diffusion process models (Ratcliff & McKoon, 2008; Ratcliff & Rouder, 1998) have been very successful in explaining in full detail two-choice forced tasks, including many observations of RT in error trials. One of the caveats of the full model with two-barriers is that it has too many parameters requiring a complex fitting-procedure (Diederich & Busemeyer, 2003; Ratcliff & Tuerlinckx, 2002) and sufficient information both in correct and incorrect responses (which in turn requires sufficient error trials to be fitted). Different simplifications of the full model have been proposed:

Wagenmakers and collaborators – in the EZ-diffusion model – simplified the full Two-Barrier Diffusion Process, including only what they considered to be the most psychologically relevant parameters of the Ratcliff model: drift rate v (i.e., quality of information), boundary separation a (i.e., response conservativeness), and non-decision time T_0. Under these simplifications, they could derive an analytical solution for the resulting distribution which can be calculated from the mean and the variance of RT and percentage of correct responses. (Grasman, Wagenmakers, & van der Maas, 2009; Wagenmakers, van der Maas, Dolan, & Grasman, 2008; Wagenmakers, van der Maas, & Grasman, 2007);

An even more simplified model postulates a single-barrier Diffusion Process (Gold & Shadlen, 2002; Heathcote, 2004; Link, 1992; Link & Heath, 1975; Luce, 1986; Schwarz, 2001; Sigman & Dehaene, 2005). This modeling strategy ignores the possibility that the second barrier can absorb trajectories and is only valid if error rates are very low. It is broadly used in one-choice alternatives although it has also been used in two-choice alternatives as in go/no-go procedures (Schwarz, 2001). The fitting parameters of this model are the same as in the EZ-diffusion model (i.e. v, a and *NDT*). A considerable advantage of this model is that it has an analytical solution and thus the parameters of the model can be fitted without explicitly simulating the random walk.

As for any scientific investigation, choosing the adequate model poses a compromise between Occam's razor and accurate descriptions. Here we opted for an iterative modeling procedure, starting from the simplest model capable of describing the key observables of these experiments and progressing towards more realistic models (Fig. 1). As in other empirical studies (Gold & Shadlen, 2002; Heathcote, 2004; Link, 1992; Link & Heath, 1975; Luce, 1986; Reddi & Carpenter, 2000; Schwarz, 2001; Sigman & Dehaene, 2005) we started by using the simplest modeling scheme, fitting RT distributions to a single barrier decision model. This model assumes that all the variance comes from decision variable and it also makes the implicit assumption that – given that errors are very rare – a single-barrier model is sufficient. Since these assumptions are not frequently made in RT distribution studies we have progressively expanded the model to see if under broader and more realistic assumptions our main observations remained stable. We thus extended the one-barrier model to Wagenmakers EZ model, the main difference (Fig. 1) being that it has two-barriers and takes into account the errors and possible differences in decisions thresholds. The EZ model does not assume variability in non-decision time and hence the next progression in the space of models was to extend our results to a Ratcliff model which includes variability in non-decision time and a possible initial bias in the decision.

In this study we explore the robustness and plasticity of different contributions to processing bottlenecks by training subjects on a PRP experiment involving a visual (number comparison), and an auditory task (pitch comparison). We examine changes in RT distributions, measuring the dynamics of RT models of decision time. We investigate the evolution of the parameter models throughout the course of learning to determine which components of processing bottlenecks are robust and which are plastic.

2. Methods

2.1. Participants

Three females and 1 male (ages between 18 and 24) participated in this study. Participants were all native English speakers. All subjects gave written informed consent and were naïve about the aims of the experiment.

2.2. General procedure

Participants performed a total of 16 experimental sessions of about 40 min each. The experiment was divided into three phases. In the first phase, participants performed six sessions in different days. Each session comprised nine independent blocks: three blocks of *Number Comparison Task* (80 trials per block, numbers were presented in Arabic Digits and Spelled Words), 3 blocks of *Tone Task* (40 trials $\times$ 3 blocks), and 3 *Double Task* blocks (120 trials per blocks) with numbers presented in Arabic Digits and Spelled Words. Each session comprises a total of 720 trials. The blocks were presented in random order.

In the second phase (day 7) participants performed seven blocks (80 trials each) in which they had to name a number presented in a new notation, a string of

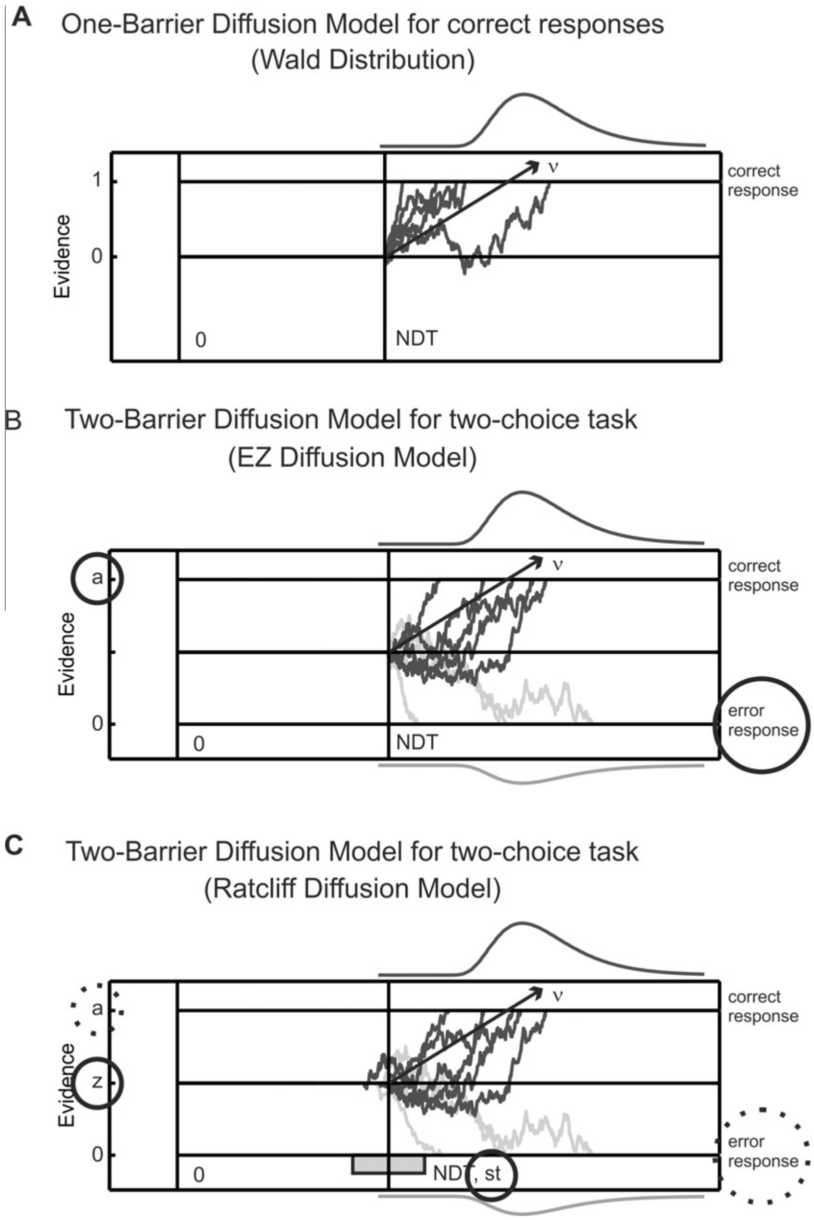

Fig. 1. RT Models of decision making (A) Single-Barrier (Wald) model. We used a fixed boundary separation ($a = 1$) and a fixed diffusion constant σ across all conditions, and fitted the distributions of RTs for correct responses with two parameters: {v, NDT}. (B) EZ-diffusion model. This model incorporates a second absorbing barrier which can account for errors. Parameters do not vary in a trial-by-trial basis and there is no original decision bias. Under these assumptions parameters {a, v, NDT}. can be obtained analytically from measured variables: {MRT, VRT, Pc}. (C) Simplified Ratcliff diffusion model (RDM): We incorporate to further parameters: the initial bias (z) and the variability in the NDT (st).

consonants. Subjects were given feedback on a trial-by-trial basis. We refer to this task as the "Naming task".

The third phase involved nine sessions. Each session of phase 3 of the experiment was identical to sessions of phase 1 except that it also included, in addition, a single Naming Task block, and that the number comparison task was also performed using the consonant-string notation.

The trials with the consonant-string notation are not included in any of the analyses presented in this paper.

2.3. Stimuli and tasks

(1) In the *Tone Task* participants heard a tone lasting 200 ms (either 440 Hz or 880 Hz) and responded with the middle and index finger of the right-hand, indicating whether it was high or low pitch. Participants had 2.5 s to respond.
(2) In the *Number Comparison Task* a number between 1 and 9 (excluding 5) was presented on the center of the screen, and participants responded whether it

was larger or smaller than 5 with the middle and index fingers of the right-hand. The numbers were randomly presented as either *digits* or *words*, for 250 ms. At the end of each block feedback was provided indicating the percentage of correct trials.

(3) In the *Dual-Task* subjects had to perform the tone and number comparison task in a classic PRP design. In each trial the sound was presented first followed by the number at 4 possible *Stimulus Onset Asynchronies* (SOA = {100, 250, 1100, 1250} milliseconds). Participants had to respond with the left-hand to Tone Task and with the right-hand to the Number Comparison Task.

Participants had 3.5 s to respond both tasks. To discourage a grouping strategy (i.e. wait for both stimuli to be presented and respond to them simultaneously) participants were explicitly instructed to respond accurately but as fast as possible to each stimuli.

2.4. Data analysis

Dual-Task trials were considered correct if responses to Task 1 and to Task 2 were correct. For RT analysis, we excluded RTs slower than 1500 ms for single-task trials, and for the First Task; and RTs slower than 2000 msec for the Second Task (in Dual-Task trials). Less than 5% of the trials were excluded using this criterion. Performance in the experiment was very accurate throughout all sessions and was independent of SOA (see Fig. 2 GH and Supplementary Fig. 1; $p = 0.1016$ ($df = 3$, $F = 2.08$) ANOVA with Pc as independent variable and {SOA, Task, Learning} as dependent variables, and Subject as Random Variable).

To study the effect of learning (Fig. 2) we concatenated all the trials acquired throughout the successive sessions of the experiment in a single sequence. We then performed a running average of 20 trials for each subject × condition (in Fig. 2 CDEGH) and 40 trials in Fig. 2F. The resulting running average was fitted to an exponential function to obtain a time-scale of learning.

2.5. Specifications, choice of models and parameters of RT distribution analysis

The analyses of the single-barrier model were based solely on correct trials. Analysis of the EZ (Wagenmakers et al., 2007, 2008) and the RDM models (Ratcliff, 1978; Ratcliff & Rouder, 1998; Vandekerckhove & Tuerlinckx, 2007) included error trials In all cases, all reported parameter values were calculated fitting the model for each individual subject, condition (Notation or SOA) and learning bin, and then averaging. Since we had four participants in this study and those were not sufficient to perform reliable statistics on the parameters, many of the statistical tests on the regressions were performed on the average data. In Figs. 3–5 and Supplementary Fig. 2 we presented the (mean ± std error) of estimated parameters for the different models. In Supplementary Table 1, we report the fitted values of the model and their corresponding linear regressions for each individual participant.

The single-barrier model is based on a random walk with a drift described by the following Markov process (see Fig. 1A):

$$d(evidence) = v \cdot d(time) + \sigma \cdot d(random\ \ variable)$$

where v is the drift rate, σ is the standard deviation of a gaussian noise with zero mean. The resulting distribution for first hitting times to a fixed threshold a could be estimated analytically and is determined by the Wald (also referred as Inverse Gaussian) distribution:

$$pdf(t) = \frac{a}{\sigma \cdot \sqrt{2\pi \cdot t^3}} \cdot \exp\left(-\frac{(a - v \cdot t)^2}{2 \cdot \sigma^2 \cdot t}\right)$$

Response Times are then modeled by this distribution plus a fixed delay (fix here refers to the fact that it does not vary from trial to trial) T_{ER}, which accounts for contributions to RT not related to the decision as motor execution of the response.

The analytical form of the resulting RT distribution is referred as the shifted-Wald distribution (see Fig. 2A):

$$pdf(t) = \frac{a}{\sigma \cdot \sqrt{2\pi \cdot (t - T_0)^3}} \cdot \exp\left(-\frac{(a - v \cdot (t - T_0))^2}{2 \cdot \sigma^2 \cdot (t - T_0)}\right)$$

The threshold was set to a fix value (arbitrarily taken as one) and we obtain the distribution that we used to fit RT distributions:

$$pdf(t) = \frac{1}{\sigma \cdot \sqrt{2\pi \cdot (t - T_0)^3}} \cdot \exp\left(-\frac{(1 - v \cdot (t - T_0))^2}{2 \cdot \sigma^2 \cdot (t - T_0)}\right)$$

We also used a fixed value of σ across subjects and conditions for each task (number or tone). We found this optimal value as a global minima for all subjects and conditions. The results reported here were very stable regarding changes in the value of σ.

Fits were based on the Mean Square Distance (MSD) of the cumulative distributions because this fitting procedure turned out to be the more stable. To fit the parameters of the model we used the Nelder–Mead Simplex Method (Lagarias, Reeds, Wright, & Wright, 1998), implemented in Matlab using the *fminsearch*() function. This is the same algorithm used in the DMAT toolbox (Vandekerckhove & Tuerlinckx, 2007; Vandekerckhove & Tuerlinckx, 2008; http://ppw.kuleuven.be/okp/software/dmat/) which we used to fit the Ratcliff Diffusion Model.

The EZ-diffusion model is the analytical resolution of the simplest version of a two-barrier model. Wagenmakers and collaborators showed that using a reduction of the full Ratcliff model (parameters do not have trial-to-trial variance and there is no decision bias) could map analytically the mean RT (MRT), the variance of RT (VRT) and Probability of Correct Response (Pc) to the parameters of the Ratcliff Diffusion Model (see Wagenmakers et al., 2007 for a rationale on this reduction from the larger Ratcliff Diffusion Model). Under these assumptions, the EZ-diffusion model considers only three free parameters (Fig. 1B): the drift rate (v), the non-decision time (T_{ER}) and the boundary separation (a). These parameters are calculated analytically from MRT, VRT and Pc (Wagenmakers et al., 2007).

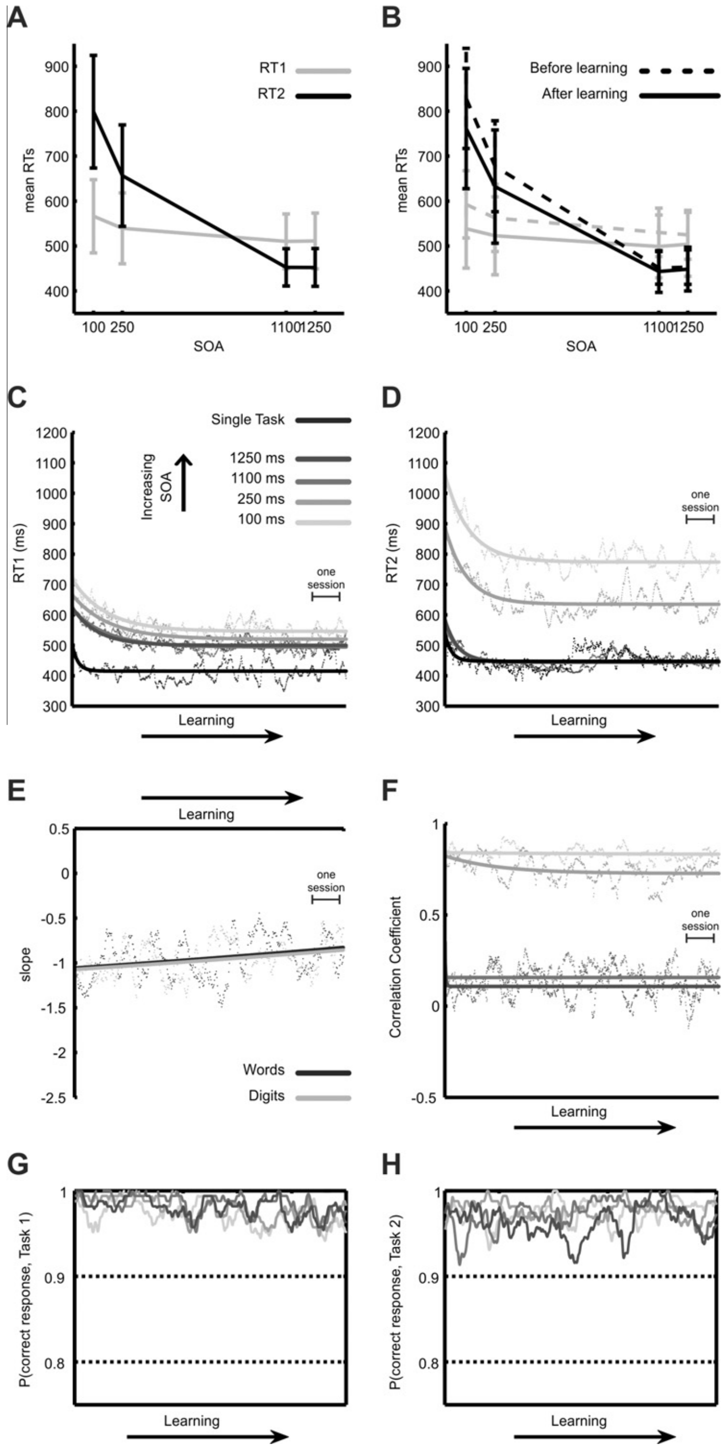

Fig. 2. What aspects of dual-task performance change with extensive practice? (A) Main PRP effect: mean RT2 (black line) decreases with SOA with a slope of −1. for short SOA values. Mean RT1 (grey line) remains almost constant. (B) mean RT as a function of SOA at different stages of learning. Trials were grouped in two categories: before learning (the first 20% of trials in the course of learning) and after learning (the last 20% of trials in the course of learning). (C and D) RT1 (C) and RT2 (D) for different SOA values and RT of the single-task as a function of learning. We observed an effect of SOA on RT1 (compare darkest and lightest) which decreased with learning. The bulk of the learning effect on mean response times occurred during the first training sessions. (E) The PRP effect, defined as the slope of RT2 between the first SOA values (RT2(SOA = 100 ms) − RT2(SOA = 250 ms))/150 ms shows a small, but progressive change with practice. (F) The correlation between RT1 and RT2. (C–F: Dotted lines: Experimental data, Solid Lines: The best exponential fit ($RT(n) = A + B * \exp(-n/\tau)$). (G and H) Probability of Correct Responses (Pc) for Task 1 (G) and Task 2 (H), for different SOA values as a function of learning.

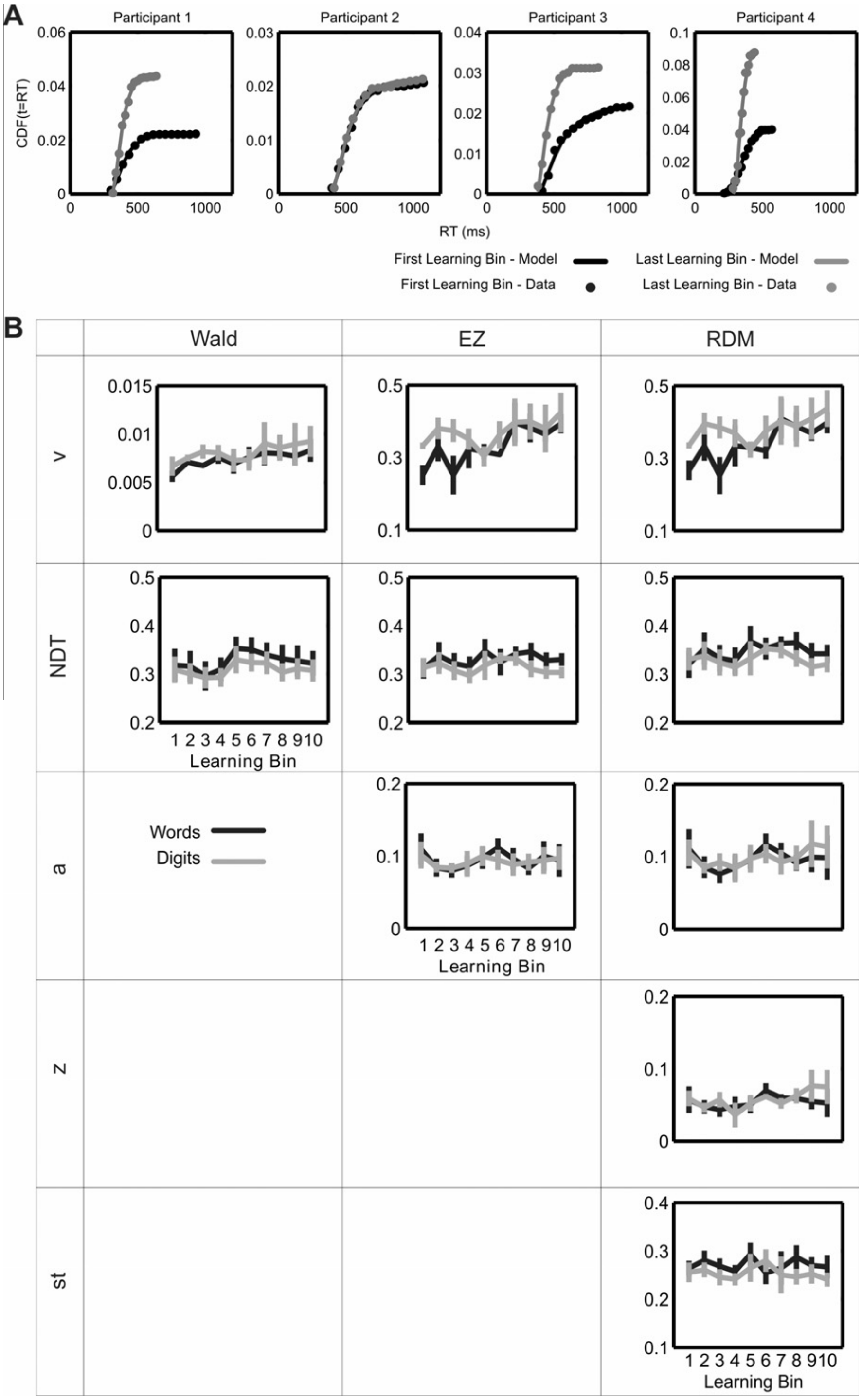

Fig. 3. What aspects of single-task architecture change with extensive practice? (A) Response time distributions of the number comparison task (Spelled words) at the beginning (black lines) and at the end (grey lines) of learning. (B) Evolution of estimated parameters from the three models: Wald distribution, EZ-diffusion model and RDM model, for both Notations in the Number task as presented as single-task. We observed, consistently, an increase in the drift v with learning.

All parameters of the EZ model were calculated with the *ezdiff*( ) function of the DMAT toolbox.

The Ratcliff Diffusion model (RDM). Parameters of the RDM were obtained using the DMAT toolbox (Vandekerckhove & Tuerlinckx, 2008). To verify whether changes in the variance with learning might be accounted by variability in non-decision time, we added an additional parameter determining inter-trial variability on non-decision

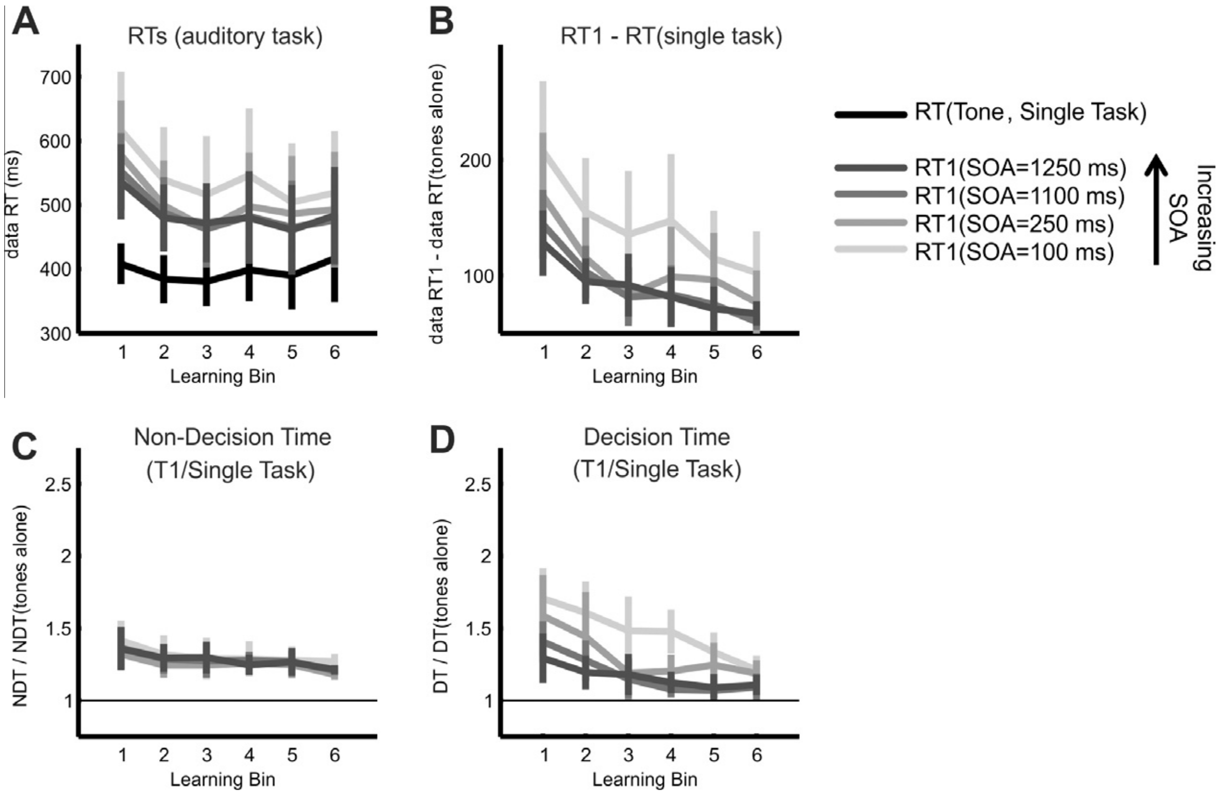

Fig. 4. Departure from strictly sequential model of dual-task: Effect of interference on RT1. (A) RT of tone task (single-task: black line, and as first task of dual-task for different SOA) as function of learning. (B) Difference between RT1 and RT (tone-task), as a function of learning. (C) Fraction of NDT for the Tone Task when performed as a single-task or as the first task of a PRP experiment. The fraction is larger than one indicating that NDT is slower when the task is performed as the first task of a PRP experiment. The effect does not depend on SOA and decreases with learning. (D) Same as (C), for DT. The effect increases for shorter SOA and decreases with learning.

(*st*) (Fig. 1C). The bias in the decision was also included as an additional free parameter to verify the hypothesis of the EZ model. The model – a simplified version of the full-RDM, has five free parameters (which are modeled for every participant, condition and learning bin) {*a, NDT, v, z, st*}. Since multi-dimensional fitting is more sensitive to initial conditions, we used an iterative procedure, taking the resulting values from the EZ-diffusion model as initial conditions for the shared parameters {*a, NDT, v*}, and *a*/*2* (no bias) as the initial condition for *z*.

3. Results

3.1. What aspects of dual-task performance change with extensive practice?

We first grouped the data across all learning sessions and observed a classic PRP effect: RT2 decreased with SOA with a slope close to -1 for short SOA values ($(RT2_{SOA=100}-RT2_{SOA=250})/(250-100)$, averaged across subjects = -0.96 ± 0.12) consistent with serial processing (during T1 processing, each acceleration in T2 presentation time is translated into a corresponding slowing of RT2). RT2 was constant for long SOA values (($(RT2_{SOA=1100}-RT2_{SOA=1250})/(1250-1100)$-, averaged value across subjects = 0.00 ± 0.01). We observed a small, yet significant slowing of RT1 at short SOA (Fig. 2A and C) (Sigman & Dehaene, 2006) and is not predicted by the classic sequential model. We will later come back to this observation in more detail.

We then investigated which aspects of the PRP effect changed with extensive practice. There was a reduction in RT with learning, both in RT1 and RT2 and for all SOA conditions (Fig. 2B and Table 1). The effect was quite small, ranging between 10 and 70 ms depending on the condition, and marginally significant (see Table 1). Fig. 2 shows that the bulk of the learning effect seems to be on the first trials, as suggested by exponential models of learning (Heathcote, Brown, & Mewhort, 2000). To quantify this observation we fit an exponential model of RT with learning: $RT = A + B * \exp(-N/\tau)$, where *A* is the asymptotic value, *B* is the amplitude of the learning and τ is the characteristic learning scale and *N* the trial number since the beginning of learning. This analysis confirmed that learning occurred rapidly, in about 100 trials ($\tau_{RT1} = 106$ trials CI: [98, 117] trials; $\tau_{RT2} = 58$ trials CI: [53, 65] trials), indicating that the bulk of the learning effect on mean RTs is obtained in the first session (Fig. 2C and D). The amplitude of the effect was more pronounced for Task 2 trials ($B_{RT1} = 144$ ms CI: [136, 151] ms; B_{RT2} (short SOAs) = 275 ms CI: [259, 290] ms; B_{RT2} (long SOAs) = 231 ms CI: [119, 143] ms).

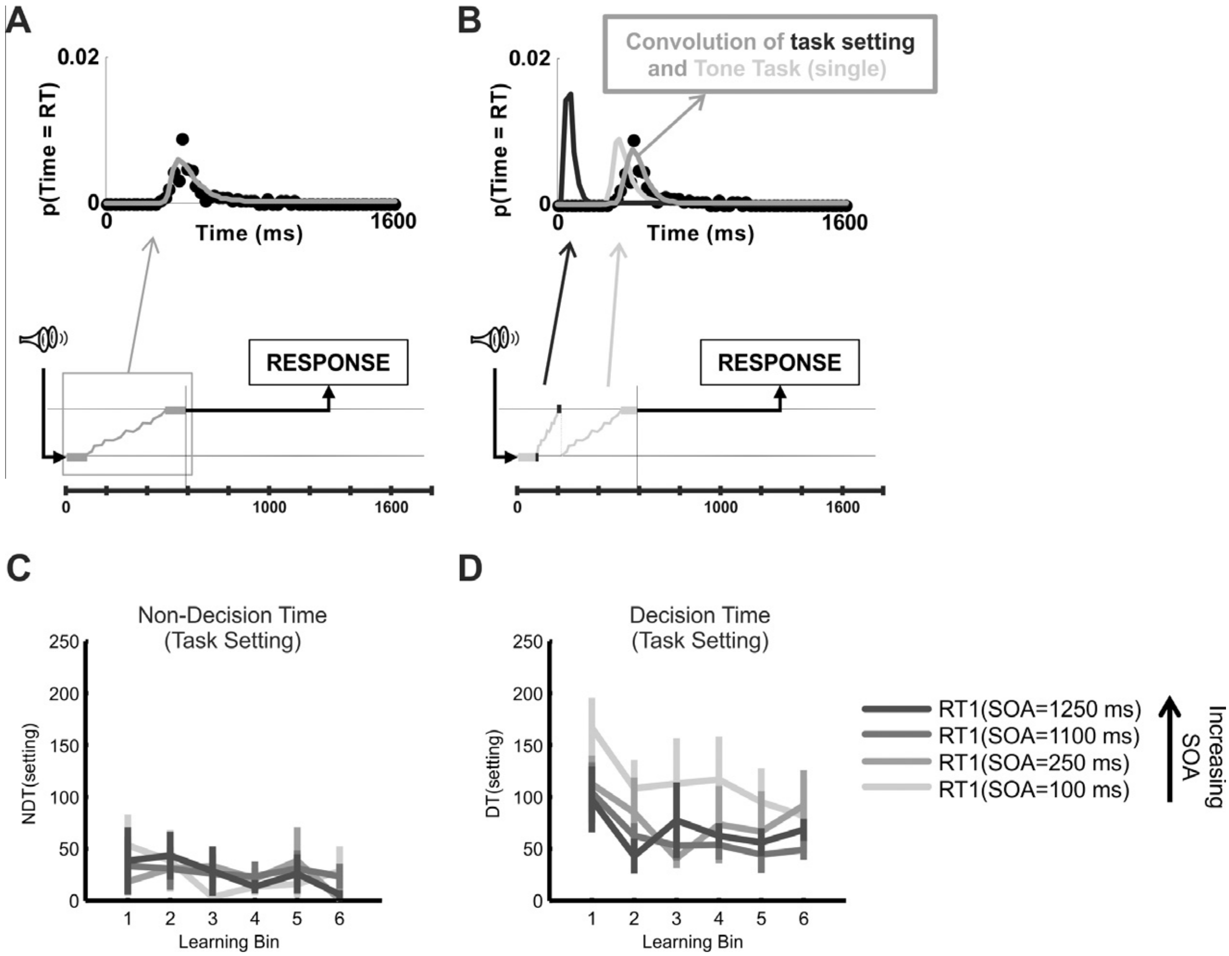

Fig. 5. Task Setting: A model of nested series of decisions. (A) Sketch of the single-barrier diffusion model used to measure DT and NDT from the response time distributions. (B) Sketch of the model used to model RT1 distributions as the convolutions of two shifted-Wald processes. The first decision models a task-setting procedure and the second decision the choice determined by the task. (C and D) DT and NDT contributions of the task-setting process. Most of the contribution results from DT which increases for shorter values of SOA. This effect decreases with learning. Trials were grouped in six bins to estimate the parameters of the model.

The bottleneck model predicts that RT2 should decrease with SOA with a slope of -1 during the interference regime and then remain constant, thus creating an ''elbow'' in the RT2 curve, although trial to trial variability will tend to smooth this out. To quantify the shift in the elbow location we measured the slope of change in RT2 between consecutive SOAs values (Table 1). A slope of 0 (respectively -1) would imply that the considered SOAs lie outside (resp. inside) the interference regime, with intermediate values providing a continuous weighted estimate of the critical SOA range.

The slope of the PRP function (considering the change in RT between the shortest SOA conditions, 100 and 250 ms) showed a moderate effect of learning (Fig. 2B and E; linear regression of slope versus trials was significant, $b = 5 \times 10^{-4}$ $[4 \times 10^{-4}\ \ 6 \times 10^{-4}]$, $F = 0.1$, $p < 10^{-8}$; mean value of the slope for the first (last) five sessions = -1.01 ± 0.06 (-0.83 ± 0.11)). The difference in RT2 at SOA values of 100 and 250 ms remained highly significant even in the last sessions of training (124.26 ± 17.89 ms). These results indicate a progressive shift of the interference regime towards shorter SOA values, but suggest that, at least with this amount of learning, the serial queuing of central processing of both tasks remains present.

Further evidence for the persistence of a serial bottleneck comes from an analysis of RT1 and RT2 correlations. The serial model predicts that these correlations are high in the interference regime, where every extra millisecond taken by Task 1 slows down Task 2 by a corresponding amount. We measured RT1/RT2 correlations as a function of SOA, during the course of learning (Fig. 2F). For the shortest SOA value (100 ms), the correlations are very strong (mean correlation = (0.85 ± 0.04)) and remain essentially invariant during the course of learning (linear regression of correlations versus trials: $b = -7 \times 10^{-6}$ $[-0.3 \times 10^{-6},\ -15 \times 10^{-6}]$, $F = 3.56$, $p = 0.0596$)). For the SOA 250 ms conditions, correlations are also strong (mean correlation = (0.76 ± 0.07)) and decrease with practice (linear regression of correlations versus trials: $b = -6 \times 10^{-4}$ $[-4 \times 10^{-4},\ -7 \times 10^{-4}]$, $F = 72.09$, $p < 10^{-15}$). Within the

Table 1
Effects of learning: Statistical tests of the regression analysis the main variables (RT1, RT2 for different SOA values) with learning. Effects of SOA: *t*-test comparisons for different SOA values collapsing all experimental sessions. Interactions between learning and SOA: Statistical tests of the regression analysis of RT differences with learning. Effect Sizes in Learning and interaction of SOA with learning were calculated comparing RTs for the first five and the last five training sessions.

	X=	RT1		RT2	
		Effect size	Significance	Effect size	Significance
Effect of Learning	X(SOA = 100)	**57.35 ms**	**F=3.39; p < 0005**	**72.16 ms**	**F = 2.87; p < 0.01**
	X(SOA = 250)	**40.19 ms**	**F = 2.32; p < 0.05**	45.28 ms	F = 1.71, p > 0.05
	X(SOA = 1100)	**32.66 ms**	**F = 2.20; p < 0.05**	9.49 ms	F = 1.12; p > 0.05
	X(SOA = 1250)	20.55 ms	F = 1.56; p > 0.05	5.49 ms	F = 0.87; p > 0.05
Effect of SOA	{X(SOA = 100), X(SOA = 250)}	**−0.17 (−25.68 ms)**	**t = 7.48; p < 0.0001**	**−0.93 (−139.04 ms)**	**t = 30.24; p < 0.0001**
	{X(SOA = 250), X(SOA = 1100)}	**−0.03 (−28.04 ms)**	**t = 10.10; p < 0.0001**	**−0.24 (−201.85 ms)**	**t = 22.51; p < 0.0001**
	{X(SOA = 1100), X(SOA = 1250)}	0.01 (1.31 ms)	t = 0.37; p > 0.05	0.00 (0.20 ms)	t = 0.09; p > 0.05
Interaction between Learning and SOA	X(SOA = 100) – X(SOA = 250)	**−0.11 (−17.16 ms)**	**F = 2.55; p < 0.05**	**−0.18 (−26.88 ms)**	**F = 4.12; p < 0001**
	X(SOA = 250) – X(SOA = 1100)	−0.01 (−7.52 ms)	F = 1.59; p > 0.05	−0.04 (−35.80 ms)	F = 1.75; p > 0.05
	X(SOA = 1100) – X(SOA = 1250)	−0.08 (−12.11 ms)	F = 1.49; p > 0.05	−0.03 (−4.00 ms)	F = 0.59; p > 0.05

(*) Bold values indicates significant effects at $\alpha = 0.05$.

non-interference regime, correlations are weak throughout the course of learning (mean correlation for both large SOA values = (0.14 ± 0.04)).

Note that both the change in the slope of the PRP effect and of the correlations between RT1 and RT2 for short SOA values progressed with learning even beyond a few hundred trials, when the mean RTs had reached a plateau. This finding indicates that learning may result in changes in task architecture which are not simply visible by mean RT analysis, and motivates our subsequent decomposition of RTs using random-walk models.

Finally we explored whether an important departure from the passive bottleneck model, the observed effect of SOA on RT1, changed with practice. As seen in Fig. 2C, during the first sessions, RT1 is on average almost 100 ms larger for short SOA values (difference between blue and green, and red and yellow curves). In the last session, however, this difference virtually vanishes. The reduction of the effect of SOA on RT1 is accompanied by an overall decrease of RT1 (Table 1).

Summarizing, we observe the following effects with practice in dual-task performance: (1) A reduction of mean response times in both tasks. (2) A shift of the extent of interference regime ("the elbow location") to shorter SOA values, (3) the reduction or virtually disappearance of an effect of SOA on RT1 and (4) a constancy of the PRP bottleneck effect, even after extensive practice, for SOA values shorter than 200 ms, as evidenced by the persistence of high correlations between RT1 and RT2 and by the consistent decrease of RT2 for short SOA values.

3.2. Changes in single-task architecture and contributions to RT during the course of learning

As described in the previous section, we observed a small reduction of RT with learning, mainly in the first sessions of training. The question we investigate here is whether the differential contributions of different stages to RT varied significantly throughout the course of learning, beyond this initial regime. RT distributions were different prior and after extensive learning (see Fig. 3A). To provide a quantitative measure of the change of different contributions to RT, we analyzed the evolution of the different parameters of the models throughout the course of learning. This analysis was performed independently for the Arabic digits and spelled words numerical task.[1]

The single-barrier model (see methods and Fig. 1) assumes a fixed onset delay referred to as Non-Decision-Time (*NDT*), followed by a forced random walk to a threshold b with slope (drift rate) v and diffusion constant σ. Only *NDT* and v are free parameters. The fixed delay (NDT) indicates a non-stochastic contribution to response time. The Decision Time (DT), the main stochastic contribution to RT is determined by b/v. Since the threshold is fixed, an increase in v implies a reduction in decision time. Analysis revealed a progressive increase of drift rate v with learning (Fig. 3B top-left, $p < 0.002$ for Words and $p < 0.015$ for Digits, see Table 2 for all details of the regression). NDT showed a very moderate, non significant increase with learning (Fig. 3B left column, second-row, $p > 0.18$ for Words and $p > 0.46$). The fraction of RT devoted to decision time (*DT*/RT) decreased during the course of learning (Regression on mean curves for both Notations: Digits: Effect Size (mean ± std error) = (−0.056 ± 0.005); $F = 13.01$; $p < 0.01$; and Words: Effect Size (mean ± std error) = (−0.056 ± 0.005); $F = 22.49$; $p < 0.0005$).

The EZ model assumes two symmetrical barriers at the cost of an additional free parameter, the distance between both barriers, referred as a. We observed the same dependence for the shared parameters (Fig. 3B, 2nd column) v and *NDT*. We observed a progressive increase in drift rate v throughout the course of learning which was significant for Words ($p < 0.0036$, see Table 2) and marginally significant for Digits ($p < 0.066$, see Table 2) and NDT did not change with learning (see Table 2). The separation between both barriers a did not change with learning (see Table 2).

[1] Interestingly the tone task – in the single-task condition – showed virtually no effects of practice (see Fig. 4A). This task was responded very rapidly and thus it is likely that response times were close to saturation. Another difference is that the visual task involved to distinct perceptual modalities (words and digits) while the auditory task always was presented with the same two tones.

Table 2
Statistical tests of the regression analysis the estimated parameters of the Numerical Task as Single-Task (and both Notations) for the three models with learning.

		Wald		EZ		RDM	
Drift Rate (v)	Words	**0.0002**	**$F = 18.81$; $p < 0005$**	**0.0144**	**$F = 16.59$; $p < 0005$**	**0.0141**	**$F = 15.14$; $p < 0005$**
	Digits	**0.0002**	**$F = 9.50$; $p < 0.05$**	0.0070	$F = 4.52$; $p = 0.0663$	**0.0072**	**$F = 5.62$; $p < 0.05$**
Non-Decision Time (*NDT*)	Words	0.0026	$F = 2.17$; $p = 0.179$	0.0017	$F = 1.49$; $p = 0.2576$	0.0023	$F = 1.69$; $p = 0.2293$
	Digits	0.0010	$F = 0.58$; $p = 0.467$	−0.0006	$F = 0.15$; $p = 0.7084$	−0.0005	$F = 0.12$; $p = 0.7391$
Boundary Separation (a)	Words			0.0002	$F = 0.02$; $p = 0.8933$	0.0009	$F = 0.37$; $p = 0.5588$
	Digits			0.0004	$F = 0.26$; $p = 0.6214$	0.0022	$F = 4.36$; $p = 0.0703$
Initial Bias (z)	Words					0.0009	$F = 1.09$; $p = 0.3279$
	Digits					**0.0027**	**$F = 6.12$; $p < 0.05$**
Variability of NDT (st)	Words					0.0002	$F = 0.01$; $p = 0.9124$
	Digits					−0.0009	$F = 0.41$; $p = 0.5418$

(*) Bold values indicates significant effects at $\alpha = 0.05$.

The RDM further incorporates two additional parameters (Fig. 3B, 3rd column), the variability of non-decision time st and an initial bias towards one of the barriers (z) and hence has a total of five free parameters: {a, NDT, v, z, st}. Analysis of the evolution of the parameter models through of learning revealed, as with the simpler models, an increase in v (see Table 2). All other parameters, with the exception of z only for Digits, did not show any effect of learning (Table 2). The mean value of z for both notations was not-significantly different from (a/2) (a(mean ± std across participants) = (0.10 ± 0.03) and z(mean ± std across participants) = (0.06 ± 0.02)) indicating that, as expected, none of the responses had an a priori bias. The variance of non-decision time (st) was on average (mean ± std across participants = 0.26 ± 0.04) which is within normal ranges of what has been found in prior studies on other tasks (Ratcliff & Smith, 2004; Ratcliff, Thapar, & McKoon, 2001; Ratcliff, Thapar, & McKoon, 2004; Ratcliff, Thapar, Gomez, & McKoon, 2004).

In summary, a distribution analysis of RT showed that even beyond the extent where we could observe changes in mean RT, there is a reorganization of the contribution of the different stages of a task to RT variability, revealing a shift from decision to non-decision time. We observed a consistent pattern of all three models revealing an increase of v with learning.

3.3. Relating single and dual-task performance: departures from the passive bottleneck model

The sequential bottleneck model predicts a highly specific relationship between dual and single-task performance. In particular, it predicts no difference between RT1 and RT of the first task when performed in isolation (RT_{single}). In this experiment the tone discrimination task was performed as the first task in the dual-task experiment. Contrary to the prediction of the passive bottleneck model, we observed that the mean RT1 was significantly larger than RT_{single} (Fig. 4A, paired t-test comparison comparing RT1 with RT_{single} $p < 0001$ for all SOA values). Previous experiments have observed a comparable slowing of the first task of about 100 ms (Sigman & Dehaene, 2005).

The difference between RT1 and RT_{single} decreased with training ($p < 0.01$ for all SOA values), but remained positive even after extensive training ($p < 0.01$ for all SOA values, before training 150.94 ms (averaged across all SOA values), after training 87.66 ms, see Fig. 4A and B). We then investigated which stages of task execution accounted for this increase in RT and whether this changed with learning (Fig. 4B).

As previously, we first parsed RTs into decision time and non-decision time using the Wald (single-barrier model) and subsequently extended these results to the two-barrier models (EZ and RDM).

For each SOA value of the double task, for each independent session we computed the RT distributions which were submitted to the Wald model to obtain the parameter values. For each condition (learning session and SOA), DT and NDT of the RT1 of the dual-task experiment were normalized to the values of DT and NDT of RT_{single}.

Decision Time and Non-Decision Time were greater in RT1 as compared with RT_{single}. Indeed, mean (RT1) was greater than mean (RT_{single}) for every SOA value and learning bin, and hence the ratio is systematically greater than 1 (Fig. 4C and D).

The dependencies of these two contributions to response time with SOA and learning showed qualitatively distinct patterns. First, the increase in *Non-Decision Time* was independent of SOA (Fig. 4C, see Table 3) and reached a non-zero asymptote, i.e. remained significant even after extensive practice (Fig. 4D). The effect in *Decision Time* was strongly dependent of SOA and vanished after learning was completed (Fig. 4C, see Table 3).

Both contributions of response time decreased with learning but the effect was much more pronounced for the Decision Time contribution at short SOA values (Fig. 4D and Table 4).

The results were highly consistent and reproducible for the three models (Supplementary Fig. 2, Tables 3 and 4). In all cases we observed that a strong dependence of v with SOA (which simply accounts for the observation of the DT dependency with SOA). On the contrary, NDT of all three models were insensitive to SOA. The additional parameters of the models {a, z, st} did not change with SOA (Supplementary Fig. 2). In all three models NDT and its variability showed very similar patterns. Both were insensitive to SOA and both increased for RT1 as compared to RT_{single}.

In summary, we observed a systematic departure from the passive bottleneck model: an increase of RT1 compared to RT_{single}. This effect decreases with learning, and reflects both stochastic and non-stochastic components of RT, with

Table 3
Results of an ANOVA with DT or NDT as independent variable and {SOA, Learning} as dependent variables, and Subject as Random Variable.

	Model	SOA	Learning	SOA × Learning
Non-Decision Time	Wald	$F = 0.67$; $p = 0.4742$	$F = 0.18$; $p = 0.9660$	$F = 0.71$; $p = 0.6230$
	EZ	$F = 0.52$; $p = 0.5258$	$F = 0.54$; $p = 0.7460$	$F = 2.69$; $p = 0.0627$
	RDM	$F = 0.35$; $p = 0.5960$	$F = 0.20$; $p = 0.9559$	**$F = 3.09$; $p < 0.05$**
Decision Time	Wald	**$F = 18.06$; $p < 0.05$**	$F = 1.86$; $p = 0.1618$	**$F = 3.18$; $p < 0.05$**
	EZ	**$F = 25.62$; $p < 0.05$**	$F = 2.67$; $p = 0.0639$	**$F = 9.62$; $p < 0.0005$**
	RDM	**$F = 30.37$; $p < 0.05$**	**$F = 4.39$; $p < 0.05$**	**$F = 7.03$; $p < 0005$**

(*) Bold values indicates significant effects at $\alpha = 0.05$.

Table 4
Statistical tests of the regression analysis the DT and NDT for the shortest and largest SOA, and the three models with learning.

	Model	SOA = 100 ms		SOA = 1250 ms	
		Slope	Significance	Slope	Significance
Non-Decision Time	Wald	−0.002	$F = 0.03$; $p = 0.8725$	−0.019	$F = 3.15$; $p = 0.1506$
	EZ	0.009	$F = 0.27$; $p = 0.6280$	−0.022	$F = 2.31$; $p = 0.2035$
	RDM	**0.028**	**$F = 10.89$; $p < 0.05$**	−0.008	$F = 0.78$; $p = 0.4265$
Decisión Time	Wald	**−0.093**	**$F = 84.08$; $p < 0001$**	**−0.039**	**$F = 37.75$; $p < 0005$**
	EZ	**−0.136**	**$F = 28.59$; $p < 0.01$**	−0.013	$F = 0.35$; $p = 0.5875$
	RDM	**−0.262**	**$F = 51.59$; $p < 0005$**	−0.054	$F = 1.88$; $p = 0.2425$

(*) Bold values indicates significant effects at $\alpha = 0.05$.

only the non-stochastic contributions (unaffected by SOA) remaining significant after extensive practice.

3.4. A model of nested series of decisions

In our previous work, we had suggested that dual-task execution could involve a nested series of decisions: first choosing a task-schedule (which task to respond first) and then executing the two successive task decisions (Sigman & Dehaene, 2006). We showed in particular that the former, executive decision was influenced by SOA (it is harder to focus on the auditory task when the distracting visual information is presented simultaneously or at a short time interval). We implemented a model of two chained decisions, by modeling each decision with a shifted Wald-distribution and then convolving the two resulting distributions (Fig. 5A and B). The parameters (NDT and DT) of the RT_{single} were obtained from the single-task experiment. We then convolved this with a shifted-Wald distribution with unknown DT and NDT values, corresponding to the executive decision. We obtained these parameters comparing the resulting convolution with the RT1 distribution. A caveat which must be taken into consideration is that single-task (both auditory and visual) were responded with the right-hand, while in the PRP, the first task was responded with the right-hand and the second task with the left-hand.

We observed that the task-engaging stage had a Non-Decision-time which was unaffected by SOA and decreased moderately with learning, and a Decision-Time which increased significantly for short SOA values and showed a very strong decrease with learning (Fig. 5C and D; Effect of SOA: NDT: $p = 0.1496$, $t = 1.70$, $df = 5$, and DT: $p = 0.0010$, $t = 6.81$, $df = 5$; paired t-test between values at short and large SOAs; and Effect of Learning: NDT: $p = 0.0273$ and DT: $p = 0.0482$).

4. Discussion

A broad literature has examined the process of task automatization (Compton & Logan, 1991; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977), and within the PRP paradigm, a series of studies have explored the changes of the PRP effect after extensive practice (Greenwald, 2003; Greenwald & Shulman, 1973; Lien et al., 2002; Maquestiaux, Lague-Beauvais, Ruthruff, & Bherer, 2008; Pashler, Johnston, & Ruthruff, 2001; Ruthruff, Hazeltine, & Remington, 2006; Ruthruff, Johnston, Van Selst, Whitsell, & Remington, 2003; Ruthruff, Van Selst, Johnston, & Remington, 2006; Ruthruff et al., 2001; Van Selst, Ruthruff, & Johnston, 1999). These studies have consistently found a marked reduction of interference during the PRP. However, the nature of this speedup has been debated. Does it reflect a complete parallelization of the two tasks or, alternatively, a persisting central seriality but with a speedup in the central stages of each task?

The main novelty of the present work is that we combined interference analysis (delays in dual-task experiment) and random-walk models of decision making to measure the decision and non-decision contributions to RT. This allowed us to explore the reorganization of task architecture throughout the course of learning, investigating the contributions of different stages of task execution to response time, response time variability and serial processing bottlenecks.

The observed effect of learning on the PRP is consistent with previous findings. We found that the regime of interference was reduced, but that nevertheless, for sufficient short SOA values, the PRP effect remained intact. It is of course possible that with even more practice sessions the interference regime may be further reduced and that even for the SOA values we have explored we could have found a reduction of the PRP effect. However, our results suggests that for such simple, yet arbitrary tasks (as compared for instance to highly compatible ideomotor tasks with very regular and overtrained stimulus–response mappings), even after extensive practice, a strong residual of the PRP effect is found. The quantitative parameters of the PRP vary (particularly the location of the "elbow" in the curve relating RT2 to SOA), but the qualitative pattern of central seriality is unaffected.

This persistence of bottleneck-type interference seems potentially consistent with the "global workspace theory". The global workspace constitutes a broadcasting system enabling communication to take place between arbitrary and otherwise not directly connected processes (Baars, 1989; Dehaene & Naccache, 2001; Dehaene et al., 1998). According to this theory, when the relation between stimuli and responses is entirely arbitrary, a temporary mapping between otherwise independent processors must be established. This is achieved through the mediation of a central workspace. While this interference can be reduced for highly non-arbitrary tasks (Greenwald, 2003; Lien & Proctor, 2000; Lien, Schweickert, & Proctor, 2003; Lien et al., 2002), thousands of training trials are, according to our results, insufficient to route information from arbitrary stimuli to arbitrary responses, bypassing the workspace system.

We observed that the reduction of the basic PRP effect (the range in which the two tasks interfere) was accompanied by a progressive increase in the fraction of processing time devoted to non-decision time. This consistent progression towards more reliable and more task overlap may explain the general observation of task automatization (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977) and is also at least broadly consistent with the idea that learning may result in earlier stages in cortical processing coming to represent complex task-relevant features over the course of training (Gilbert, Sigman, & Crist, 2001; Sigman & Gilbert, 2000; Sigman et al., 2005). Our results are also in line with the hypothesis that memory retrieval and slow algorithmic decision coexist and compete in a race (Compton & Logan, 1991; Klapp et al., 1991a, 1991b). As in these studies, we propose a model of a fixed cognitive architecture where only the parameters change through the course of learning.

The final objective of this paper was to explore the impact of extensive practice on executive components involved in dual-task coordination. The involvements of such components is more evident in situations in which the effectors of each task share a common representation (De Jong, 1993, 1995) and in experiments in which the order of actions to achieve a complex goal is not known in advance (Sigman & Dehaene, 2006). The evidence that response times in PRP include some executive component is inconsistent with a simple passive queuing model of the PRP. It favors models of cognitive architecture which postulate more than one bottleneck. Previous studies found that, besides the response selection bottleneck, there is another one corresponding to response initiation, which is more pronounced when the same effector is used (i.e. bimanual tasks) than when combining effectors that do not share an abstract motor representation as in voice-manual dual-task experiment (De Jong, 1993; Logan & Burkell, 1986). Here we found that task-setting executive components are also present and sensitive to the critical interference parameter, the SOA. After extensive practice, we could detect a remaining very short (100 ms, see Fig. 5) but very reliable contribution of task engagement which, in turn was completely unaffected by SOA. Thus, our results suggest that the executive decisions involved in the scheduling of two consecutive tasks also benefit from an automatization process and, after extensive practice, are accomplished within a very short and temporally reliable processing stage.

Encouragingly, our analysis of the different contributions to RT during single-task and during dual-task interference yielded consistent results across the three models examined. This suggests that, for the specific distributions analyzed in this experiment, the task analyses provided by those models are roughly equivalent. This conclusion need not hold in other situations, of course. The reason why these models generally agreed in the present experiment depends upon several features of the dataset. First, there is apparently no response bias: RTs were identical for the HIGH and LOW pitch responses (a *t*-test comparison between these distributions was not-significantly different from zero for all subjects and learning bin). This is very different for instance in the lexical decision task, which has been widely studied using two-barrier models (Ratcliff & McKoon, 2008; Ratcliff & Rouder, 1998; Wagenmakers, Van der Mass, & Grasman, 2007), in which the two-choice boundaries "Is a word" and "Is not a word" are not symmetric. Second, the number of errors is extremely low – and independent of experimental conditions. The independence of the variance of NDT with SOA and learning was not intuitive a priori and required explicit investigation.

In summary, extensive practice significantly improves the speed and reliability of our decisions, but for the arbitrary tasks that we studied here, it does not alleviate or circumvent the serial bottleneck. The inability to process two tasks at once appears, once more, as a robust and—in some senses at least—structural feature of the cognitive architecture.

Acknowledgements

The authors thank David Cun for programming assistance and Noriko Coburn for overseeing the data collection. MS and SD are funded by the Human Frontiers Science Program. MS is also supported by the SECYT, PICT 38366. JEK has a fellowship of the National Research Council of Argentina (CONICET).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.cognition.2010.12.010.

References

Allport, D., Styles, E., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421–452). Cambridge, MA: MIT Press.

Baars, B. J. (1989). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 21*(1), 109–127.

Compton, B. J., & Logan, G. D. (1991). The transition from algorithm to retrieval in memory-based theories of automaticity. *Memory & Cognition, 19*(2), 151–158.

De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance, 19*(5), 965–980.

De Jong, R. (1995). The role of preparation in overlapping-task performance. *Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology, 48*(1), 2–25.

Dehaene, S., Kerszberg, M., & Changeux, J. P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences of the United States of America, 95*(24), 14529–14534.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition, 79*(1–2), 1–37.

Diederich, A., & Busemeyer, J. R. (2003). Simple matrix methods for analyzing diffusion models of choice probability, choice response time, and simple response time. *Journal of Mathematical Psychology, 47*(3), 304–322.

Donchin, E., & Coles, M. G. H. (1998). Context updating and the P300. *Behavioural Brain Research, 21*, 152–154.

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology, 41*, 260–274.

Gilbert, C. D., & Sigman, M. (2007). Brain states: Top-down influences in sensory processing. *Neuron, 54*(5), 677–696.

Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron, 31*(5), 681–697.

Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences, 5*(1), 10–16.

Gold, J. I., & Shadlen, M. N. (2002). Banburismus and the brain: Decoding the relationship between sensory stimuli, decisions, and reward. *Neuron, 36*(2), 299–308.

Grasman, R. P., Wagenmakers, E. J., & van der Maas, H. L. (2009). On the mean and variance of response times under the diffusion model with an application to parameter estimation. *Journal of Mathematical Psychology, 53*(2), 55–68.

Greenwald, A. G. (2003). On doing two things at once. III. Confirmation of perfect timesharing when simultaneous tasks are ideomotor compatible. *Journal of Experimental Psychology – Human Perception and Performance, 29*(5), 859–868.

Greenwald, A. G., & Shulman, H. G. (1973). On doing two things at once. II. Elimination of the psychological refractory period effect. *Journal of Experimental Psychology, 101*(1), 70–76.

Heathcote, A. (2004). Fitting Wald and ex-Wald distributions to response time data: An example using functions for the S-PLUS package. *Behavior Research Methods Instruments & Computers, 36*(4), 678–694.

Heathcote, A., Brown, S., & Mewhort, D. J. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7*(2), 185–207.

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

Jiang, Y., Saxe, R., & Kanwisher, N. (2004). Functional magnetic resonance imaging provides new constraints on theories of the psychological refractory period. *Psychological Science, 15*(6), 390–396.

Kamienkowski, J. E., & Sigman, M. (2008). Delays without mistakes: Response time and error distributions in dual-task. *PLoS ONE, 3*(9), e3196. doi:10.1371/journal.pone.0003196.

Klapp, S. T., Boches, C. A., Trabert, M. L., & Logan, G. D. (1991a). Automatizing alphabet arithmetic I: Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*(2), 179–195.

Klapp, S. T., Boches, C. A., Trabert, M. L., & Logan, G. D. (1991b). Automatizing alphabet arithmetic II: Are there practice effects after automaticity is achieved? *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*(2), 196–209.

Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the Nelder–Mead Simplex Method in low dimensions. *SIAM Journal of Optimization, 9*(1), 112–147.

Laming, D. (1968). *Information theory of choice-reaction times*. New York: Academic Press.

Lien, M. C., McCann, R. S., Ruthruff, E., & Proctor, R. W. (2005). Dual-task performance with ideomotor-compatible tasks: Is the central processing bottleneck intact, bypassed, or shifted in locus? *Journal of Experimental Psychology: Human Perception and Performance, 31*(1), 122–144.

Lien, M. C., & Proctor, R. W. (2000). Multiple spatial correspondence effects on dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance, 26*(4), 1260–1280.

Lien, M. C., Proctor, R. W., & Allen, P. A. (2002). Ideomotor compatibility in the psychological refractory period effect: 29 years of oversimplification. *Journal of Experimental Psychology: Human Perception and Performance, 28*(2), 396–409.

Lien, M. C., Schweickert, R., & Proctor, R. W. (2003). Task switching and response correspondence in the psychological refractory period paradigm. *Journal of Experimental Psychology: Human Perception and Performance, 29*(3), 692–712.

Link, S. W. (1992). *The wave theory of difference and similarity*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika, 40*, 77–111.

Logan, G. D., & Burkell, J. (1986). Dependence and independencein responding to double stimulation: A comparison of stop, change, and dual-task paradigms. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 549–563.

Logan, G. D., & Gordon, R. (2001). Executive control of visual attention in dual tasks. *Psychological Review, 108*(2), 393–434.

Luce, R. D. (1986). *Response times*. New York: Oxford University Press.

Machens, C. K., Romo, R., & Brody, C. D. (2005). Flexible control of mutual inhibition: A neural model of two-interval discrimination. *Science, 307*(5712), 1121–1124.

Maquestiaux, F., Lague-Beauvais, M., Ruthruff, E., & Bherer, L. (2008). Bypassing the central bottleneck after single-task practice in the psychological refractory period paradigm: Evidence for task automatization and greedy resource recruitment. *Memory & Cognition, 36*(7), 1262–1282.

McLeod, P. (1977). A dual task response modality effect: Support for multiprocessor models of attention. *Quarterly Journal of Experimental Psychology, 29*, 651–667.

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology, 41*(3), 211–253.

Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance, 10*(3), 358–377.

Pashler, H., Carrier, M., & Hoffman, J. (1993). Saccadic eye movements and dual-task interference. *Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology, 46*(1), 51–82.

Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *QJEP, 41A*, 19–45.

Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology, 52*, 629–651.

Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Sciences of the United States of America, 91*(16), 7398–7403.

Posner, M. I., & McLeod, P. (1982). Information processing models – In search of elementary operations. *Annual Review of Psychology, 33*, 477–514.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59–108.

Ratcliff, R. (1988). Continuous versus discrete information processing modeling accumulation of partial information. *Psychological Review, 95*(2), 238–255.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*(4), 873–922.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9*(5), 347–356.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111*(2), 333–367.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review, 9*(3), 438–481.

Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging, 16*, 323–341.

Ratcliff, R., Thapar, A., & McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language, 50*, 408–424.

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging, 19*, 278–289.
Reddi, B. A., & Carpenter, R. H. (2000). The influence of urgency on decision time. *Nature Neuroscience, 3*(8), 827–830.
Ruthruff, E., Hazeltine, E., & Remington, R. W. (2006). What causes residual dual-task interference after practice? *Psychological Research, 70*(6), 494–503.
Ruthruff, E., Johnston, J. C., Van Selst, M., Whitsell, S., & Remington, R. (2003). Vanishing dual-task interference after practice: Has the bottleneck been eliminated or is it merely latent? *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 280–289.
Ruthruff, E., Pashler, H. E., & Klaassen, A. (2001). Processing bottlenecks in dual-task performance: Structural limitation or strategic postponement? *Psychonomic Bulletin & Review, 8*(1), 73–80.
Ruthruff, E., Van Selst, M., Johnston, J. C., & Remington, R. (2006). How does practice reduce dual-task interference: Integration, automatization, or just stage-shortening? *Psychological Research, 70*(2), 125–142.
Schall, J. D. (2000). From sensory evidence to a motor command. *Current Biology, 10*(11), R404–406.
Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review, 84*, 1–66.
Schumacher, E. H., Seymour, T. L., Glass, J. M., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing n dual-task performance: Uncorking the central attentional bottleneck. *Psychological Science, 12*, 101–108.
Schwarz, W. (2001). The ex-Wald distribution as a descriptive model of response times. *Behavior Research Methods, Instruments & Computers, 33*(4), 457–469.
Shadlen, M. N., & Newsome, W. T. (1996). Motion perception: Seeing and deciding. *Proceedings of the National Academy of Sciences of the United States of America, 93*(2), 628–633.
Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.
Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*, 127–191.
Sigman, M., & Dehaene, S. (2005). Parsing a cognitive task: A characterization of the Mind's bottleneck. *PLoS Biology, 3*(2), e37.
Sigman, M., & Dehaene, S. (2006). Dynamics of the central bottleneck: Dual-task and task uncertainty. *PLoS Biology, 4*(7), e220.
Sigman, M., & Dehaene, S. (2008). Brain mechanisms of serial and parallel processing during dual-task performance. *Journal of Neuroscience, 28*(30), 7585–7598.
Sigman, M., & Gilbert, C. D. (2000). Learning to find a shape. *Nature Neuroscience, 3*(3), 264–269.
Sigman, M., Pan, H., Yang, Y., Stern, E., Silbersweig, D., & Gilbert, C. D. (2005). Top-down reorganization of activity in the visual pathway after learning a shape identification task. *Neuron, 46*(5), 823–835.
Smith, M. C. (1967). Theories of the psychological refractory period. *Psychological Bulletin, 67*, 202–213.
Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology, 14*, 1–36.
Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108*(3), 550–592.
Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin and Review, 14*, 1011–1026.
Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with Matlab: A DMAT Primer. *Behavior Research Methods, 40*, 61–72.
Van Selst, M., Ruthruff, E., & Johnston, J. C. (1999). Can practice eliminate the psychological refractory period effect? *Journal of Experimental Psychology: Human Perception and Performance, 25*(5), 1268–1283.
Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics, 13*(1), 37–58.
Wagenmakers, E. J., van der Maas, H. L., Dolan, C. V., & Grasman, R. P. (2008). EZ does it! Extensions of the EZ-diffusion model. *Psychonomic Bulletin & Review, 15*(6), 1229–1235.
Wagenmakers, E. J., van der Maas, H. L., & Grasman, R. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review, 14*(1), 3–22.
Wong, K. F., & Wang, X. J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience, 26*(4), 1314–1328.
Zylberberg, A., Fernandez Slezak, D., Roelfsema, P. R., Dehaene, S., & Sigman, M. (2010). The brain's router: A cortical network model of serial processing in the primate brain. *PLoS Computational Biology, 6*(4), e1000765. doi:10.1371/journal.pcbi.1000765.