

Why does it take time to make a decision? The role of a global workspace in simple decision making

Mariano Sigman and Stanislas Dehaene

INTRODUCTION

The research that we report in this chapter was motivated by three surprisingly simple questions that turn out to have deep consequences for cognitive brain architecture:

- Why does it take so much time to take a decision?
- Why is decision time so variable?
- Why can we take only one decision at a time?

Why does it take so much time to take a decision?

Like any other form of computation, brain computing takes time. Thus, it may seem obvious that each of our decisions should take some minimal duration, ultimately linked to axonal, synaptic, and dendritic propagation delays. Yet the intriguing question is: Why does it take *so* much time to take a decision? We do not necessarily refer here to the arbitrary long durations of certain choices (although this too plays its part). It can take hours, days, or months to choose where and for how long to go on vacation, and with whom, and how, and so on and so on. Decision difficulty is illustrated by the metaphor of Buridan's ass (honoring the 14th-century French philosopher Jean Buridan), who supposedly died when placed in between a stack of hay and a pail of water, unable to decide whether to eat

or drink first. But far from the limits of starving or pathological decision making, simply deciding which of two numbers is larger, or which of two tones is of a higher pitch, takes between 500 and 1500 ms—a surprisingly long time, when considering synaptic delay and the timing of selective responses to sensory stimuli.

Responses to sensory stimulation in the primary visual cortex can be recorded as early as 50 ms, and the brainstem knows about the presence of a stimulus 20 ms after its presentation. Even after many relays of information and a few synapses, in about 180 ms information about a sensory stimulus is accessible to virtually all visual brain areas (Thorpe, Fize, & Marlot, 1996). The brain knows whether an object is a face, or a tree, at about 100 ms (monkey) or 170 ms (human), yet acting on this information to make a choice—a remarkably easy problem—takes 5, or 10, or 15 times more. Propagation of information and delay lines are thus simply not a good answer to explain why deciding takes so much time.

Why is decision time so variable?

Decision variability, in turn, may also seem to be a consequence of basic biology. Fluctuations in membrane voltage potential, in synaptic connectivity and efficiency—among other sources of biological noise—make the brain a highly stochastic machine. As a result, brain computations are unreliable, which explains why we sometimes make mistakes and why decisions are sometimes faster and other times slower.

Once again this simple intuition turns out to be wrong. While certainly biological components are noisy, they can combine into extremely precise machines assuring, among others, extremely reliable DNA replication or spiking devices (neurons) with remarkable precision. There are countless demonstrations of this fact. Bryant and Segundo (1976) first noticed that spike timing accuracy could be extremely precise and that reliability depends on the particulars of the input driving the neuron. More recently, this intriguing property received renewed attention and has been demonstrated in pyramidal neuron circuits (Mainen & Sejnowski, 1995), in vivo recordings of the H1 neuron in the visual system of the fly *Calliphora vicina* (de Ruyter van Steveninck, Lewen, Strong, Koberle, & Bialek, 1997), and other subcortical (Butts et al., 2007) and cortical structures in the mammalian brain (Cecchi, Sigman, Alonso, Martinez, & Chialvo, 2000; Gur, Beylin, & Snodderly, 1997; Kara, Reinagel, & Reid, 2000). Beyond the biophysical and anatomical details and their implications for the nature of the neural code, a remarkable corollary of these studies is that the brain can perform certain computations with exquisite precision. Thus, the question of why decisions (as well as other noisy neuronal

computations) are so variable requires a better explanation than simply biological noise.

Why can we take only one decision at a time?

A ubiquitous aspect of brain function is its quasi-modular and massively parallel organization, with a large number of processors (neurons, columns, or entire areas) operating simultaneously (Hubel & Wiesel, 1959, 1968). For example, in the visual modality, a yet undefined number of visual areas perform different feature analysis (Felleman & Van Essen, 1991; Van Essen, Anderson, & Felleman, 1992) and within each visual area, a parallel ensemble of cortical columns samples simultaneously the visual scene (Hubel & Wiesel, 1965). Multiple copies of retinotopic maps allow a fast coverage of the environment at a vast field of view. This vast parallel machine can perform, with seemingly no effort and extreme rapidity, tasks that, until recently, were judged virtually impossible for contemporary artificial machines, such as invariant object recognition.

The paradox, however, is that this extraordinary parallel machine is incapable of doing various mental calculations in parallel or, even to perform a single large arithmetic calculation that requires multiple steps. How come it is so easy to recognize moving objects, but so difficult to multiply 357 times 289? And why, if we can simultaneously coordinate walking, group contours, segment surfaces, talk, and listen to noisy speech, can we only make one decision at a time?

The proposal that we make is that an answer to all three puzzles—decision duration, decision variability, and slow seriality—arises from the brain architecture for decision making. In accordance with other investigators (Bundesen, 1990; Logan & Gordon, 2001; Sternberg, 1969, 2004), we contend that any simple decision task can be decomposed into separate processing stages (sensory, decision, and motor) and that the decision stage has specific properties of pooling information, attaining a threshold and broadcasting the result, which confer to mental decisions their peculiar properties.

CONTRIBUTIONS OF DIFFERENT STAGES TO RESPONSE-TIME MEAN AND VARIABILITY

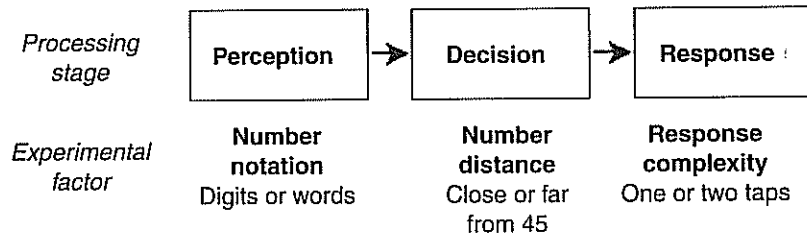
Factor analysis: Changing the mean without changing the variance

The results discussed in this chapter emerge from an exhaustive analysis of a very simple decision task. This task is number comparison, which

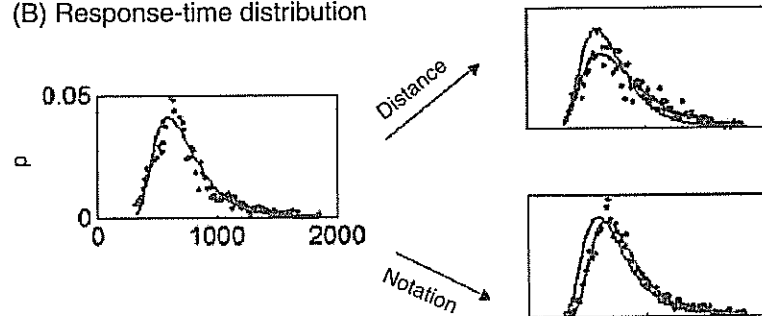
involves deciding whether a digit presented on the screen is larger or smaller than a fixed reference.

Different manipulations of the task can render it more difficult (Figure 1.1a) including: *notation* (target number presented in Arabic digits or in spelled words), *distance* (numerical distance between the target number

(A) Cognitive model



(B) Response-time distribution



(C) Response-time decomposition

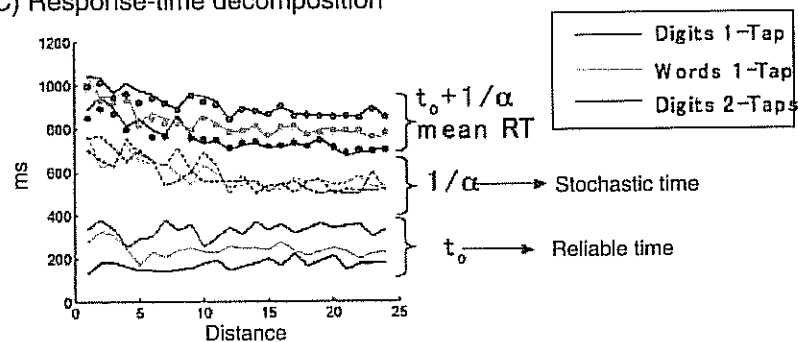


Figure 1.1 Features of the number comparison task. (a) Illustration of the number comparison task and the notation, distance, and response complexity factors. (b) A change in notation resulted in a rigid shift of the RT distribution; a change in numerical distance resulted in a wider distribution. (c) Decision (stochastic time) was only affected by numerical distance. Response and notation manipulations affected only the non-decision time. See color plate.

and the reference), and *response complexity* (whether subjects were asked to tap once or twice to indicate their choice). These manipulations change the difficulty of the task—the mean response time slows down when numerical distance decreases and when numbers are presented in spelled words (Dehaene, 1996; Moyer & Landauer, 1967; Sigman & Dehaene, 2005). These effects are additive—that is, the mean increase due to the distance factor is independent of the effect of notation (Dehaene, 1996; Pinel, Dehaene, Riviere, & LeBihan, 2001), thus establishing a first indication that they involve independent processing stages and can be factored from the entirety of the task (Sternberg, 1969, 2004).

A surprising observation resulted from an analysis of the effects of these factors on response-time variability. Although the precise dependence of the mean and variance may vary (Wagenmakers & Scott, 2007), a simultaneous increase in the dispersion and in the mean is expected from any stochastic (noisy) process. Indeed, the distance manipulation, as expected, resulted in a significant increase of the dispersion, which paralleled the increase in the mean. Strikingly, however, the notation and response complexity, while inflicting an important change in the mean, did not affect the dispersion of the RT distribution (Figure 1.1b).

Modeling decision making: The contribution of processing stages to RT mean and variance

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; Ratcliff, 1988; Schall, 2000; Schwarz, 2001; Shadlen & Newsome, 1996; Usher & McClelland, 2001; Vickers, 1970). Although many variants have been proposed, the basic idea is that perceptual evidence in favor of each of the available response alternatives is stochastically accumulated in time. A decision is taken whenever evidence in favor of one response exceeds a predetermined threshold. Decision thus results from a random walk of an internal abstract variable up to a fixed bound. In the simplest scheme, all the variance in response time is attributed to this integration process.

This analysis establishes a possible parsing of our task into components: A fixed component to transform the sensory information into an abstract internal variable (here broadly called the perceptual or P component), another fixed component to execute the response (motor or M component), and the accumulation of evidence itself (central or C component)—the latter being the only variable process.

This decomposition model was tested by analyzing how the shape of the response-time distributions varied with the experimental manipulations (Figure 1.1c). The histograms of response times were fitted to a

simple model based on a fixed onset delay (t_0 , the sum of the P and M component) plus a forced random walk with slope (α) and diffusion constant (σ) until it reaches a fixed threshold (T). T can be set to 1 without loss of generality and, for simplicity, we assumed that s was the same for all six experimental conditions, while α and t_0 could vary (none of the results depended qualitatively on the particular choice of σ). $1/\alpha$ characterizes the integration time (which explains all the variance), while t_0 captures fixed components that do not contribute to the variance. Our approach was to remain with the simplest possible model, whose sole purpose was to separate stochastic and invariant contributions to reaction times.

The results showed that the shapes of the distributions changed in two qualitatively different manners. A change in notation resulted in a rigid shift of the distribution; a change in numerical distance resulted in a widening of the RT distribution (Figure 1.1b). A parametric dependence with the distance (Figure 1.1c) revealed that t_0 changed with notation and response manipulations but was unaffected by distance. On the contrary, the decision time changed exclusively with distance, independently of the notation and response modality.

Furthermore, these results indicate that a cognitive task can be parsed in stages that provide a reliable contribution to response time, and other stages that provide a highly stochastic contribution to response time. Is this decomposition related in any manner with the serial and parallel nature of these computations? In the next section we show that indeed such a parallel decomposition can be demonstrated experimentally.

PARSING A COGNITIVE TASK INTO SERIAL AND PARALLEL PROCESSING STATES

When two tasks are presented simultaneously (or sequentially at a short interval), a delay in the execution of the second task is systematically observed (Kahneman, 1973; Pashler & Johnston, 1989; Smith, 1967; Telford, 1931). This interference effect is referred to as the Psychological Refractory Period (PRP). It has also been explained by a model that involves three stages of processing: (1) a perceptual component (P); (2) a central component (C); and (3) a motor component (M), where only the central component establishes a bottleneck (Pashler, 1994; Pashler & Johnston, 1989; Ruthruff, Pashler, & Klaassen, 2001; Schweickert, 1980; Schweickert & Townsend, 1989; Sternberg, 1969). PRP experiments have associated the central component to “response selection”, the mapping between sensory information and motor action (Pashler & Johnston, 1998). Other processes have also been implicated in the central bottleneck—not just response selection, such as mental rotation (Shepard & Metzler, 1971; Van Selst & Jolicoeur, 1994).

Interference analysis constitutes a very powerful experimental technique to understand the internal structure of a task. The logic of these experiments resembles the classic scattering methodology in physics, where the internal structure of an element (e.g., particle, molecule, etc.) is understood by colliding it with an experimental probe. In our experimental setup, the task under study (the number comparison task) was performed simultaneously or quasi-simultaneously with a probe task (the tone task). The tasks were chosen to involve entirely distinct sensory modalities (auditory versus visual) and responses (left versus right hand). The delay in the onset of the two task targets (number and tone), here called “stimulus onset asynchrony” or SOA, was controlled experimentally. To achieve a full separation of the three components, we presented the two tasks in both possible orders (Sigman & Dehaene, 2005): number followed by tone, or tone followed by number.

Under the assumptions of the PRP model, the P and M components can be carried out in parallel with another task, but the central stage is the only one that provides a bottleneck, in the sense that the central component of each task (C1 and C2) cannot be carried out simultaneously. These simple premises lead to a large number of concrete predictions on the family of curves for response time to the first and second tasks (here called RT1 and RT2) as a function of delay, and how these values change with a manipulation of the P, C, and M components of either task. The sets of predictions are summarized in Figure 1.2, and the logic leading to these predictions can be found in Sigman and Dehaene (2005). Here we summarize the main aspects.

When the number task (perturbed by different experimental manipulations) is performed first, all the manipulated variables should have a main effect on RT1, but only some of those effects (those that affect P and C components of the first task) should propagate to the response time of the second task. Furthermore, they should do so only at short inter-stimulus delays, within what we may call the “interference regime” (Figure 1.2a). In addition, SOA should not affect RT1 and should inflict no increase in RT2 for short SOA values and a linear increase with a slope of 1 for large SOA values. This assumes—as we do throughout this chapter—that RT2 is measured from the onset of the trial. In the convention in which RT2 is measured from the presentation of the second stimulus, the model predicts that RT2 decreases with a slope of -1 for short SOA durations and is constant for large SOA values.

When the number task came in second, the model predicts that there should be no effect of the manipulated variables on the first tone task. The response time to the second task should exhibit a constant increase (independent of delay) when the change affects the M and C components and should change only for large delays when the change affects the P

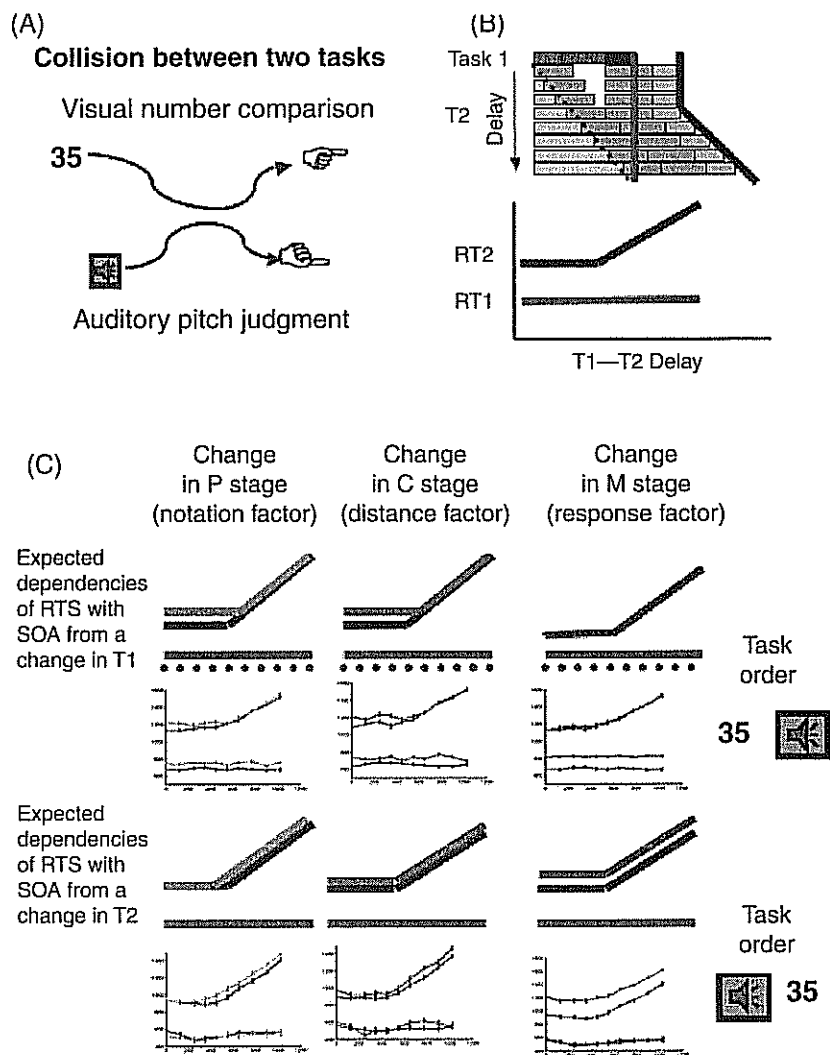


Figure 1.2 Sketch of basic task predictions. (a) Sketch of the dual-task. (b) Scheme of the main PRP effect. The vertical axis shows response time. The column on the left indicates the first task, and each color box within the column represents a different stage of processing: Perceptual component (P) (green), central component (C) (red) and motor component (M) (blue). The series of columns on the right indicate the processing time for task 2 at different delays (Δ), labeled in the x-axis. Response to the first task is independent of Δ . The response time to task 2 represented by the black line is unchanged for small Δ , while at sufficiently large Δ (non-interference regime) it increases linearly with a slope of 1, with Δ . (c) In the top panel the number task is performed first, in the bottom panel, second. In both cases, the number task is manipulated by the three factors of notation, distance, and response complexity. In all panels the code is identical: RT1 is colored grey, while RT2 is colored black. The “easy” condition is represented by a solid line and the “difficult” condition by a dotted line. All the data can be explained in terms of the PRP model: Notation (top row) affects the P component, distance (middle row) affects C, and response complexity (bottom row) affects M. See color plate.

component (Figure 1.2b). This is one of the critical predictions of the model since it implies that certain computations required for T2 perception can be performed during T1 queuing and thus argues for a coexistence of serial and parallel processing within the same task.

Every single prediction of the model was verified experimentally (Figure 1.2c). For example, in the response manipulation—when number was presented first and subjects were asked to tap the response button once or twice—we observed a large response cost (175 ms) on the first task, but none of that effect was propagated to the second task, typical of a post-bottleneck manipulation (a M1 component that operates in parallel with the C2 component). Similarly, the effect of notation was no longer observed for short SOA values when the number task was presented second, indicating a pre-bottleneck parallel stage (a P2 stage operating in parallel with P1 and C1 stages).

All the results taken together provide strong evidence that our three experimental factors (notation, distance, response complexity) mapped onto distinct stages of the PRP model. Notation manipulation affects a stage of processing prior to the bottleneck, the response manipulation to a parallel stage of processing occurring after the bottleneck, and the distance manipulation was the only factor to affect a serial stage of processing. It is striking, of course, that the task decomposition offered by the PRP and the response variability methods converged so tightly.

BRIDGING DYNAMICS AND ARCHITECTURE: A SYNTHETIC MODEL OF TASK ORGANIZATION

Taken together, the analyses of response variability and of interference suggest that the integration of evidence in time to reach a decision constitutes the only central process in a simple cognitive task that links perceptions to actions and thus that *there is a consistent parsing of a cognitive task based on variance or interference analysis*. This led us to develop a theoretical framework providing a synthesis of three basic aspects of cognitive architecture: (1) its chronometric organization; (2) its parallel and serial nature; and (3) its temporal reliability or stochasticity. We postulated that only the integration process establishes a serial bottleneck, while all other stages can proceed in parallel with stages of another task (Figure 1.3a).

While extremely simple, the model makes powerful mathematical predictions in experiments in which the order of the presentation of the two tasks and their relative offset in presentations are varied as described in the previous two sections. Moreover, it establishes a subtle, yet very critical prediction: The precise distribution of response times for the second task should be determined, according to the model, from the distributions of

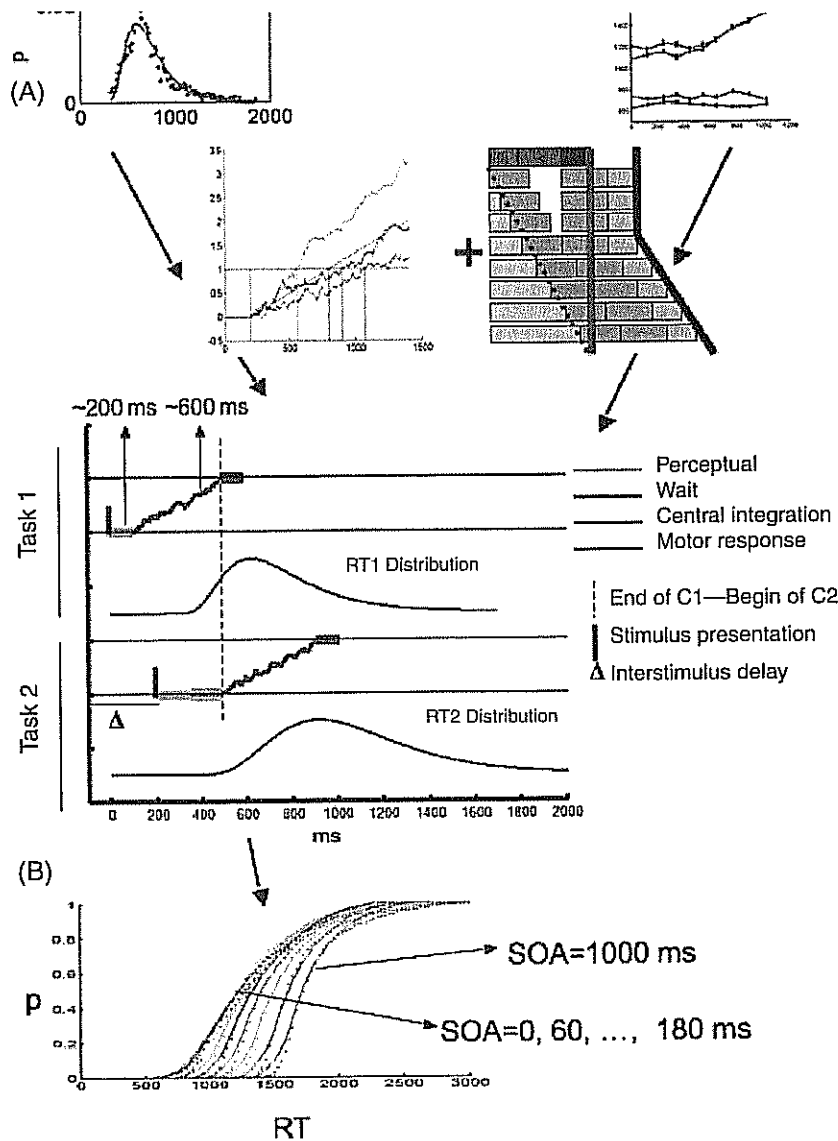


Figure 1.3 Features of parallel and serial processes. (a) A synthetic model that results from a consistent parsing based on stochastic and reliable components (left) and from interference experiments of the PRP. Perceptual and motor stages are reliable in time and can be carried out in parallel with stages of another task, while the central stage consists of a noisy integration (a random walk) until a decision threshold is reached. The central stage of task 2 cannot start until the central stage of task 1 is finished. Thus, this establishes a bottleneck and characterizes a serial process. (b) The distribution of reaction times to the second task is wider than that to the first task because it combines the intrinsic variance of task 2 (the time to reach threshold) and the variance in onset of the central stage of task 2, which is set by the end of the central stage of task 1. This distribution can be predicted correctly by the model. See color plate.

both tasks when performed in isolation or as a first task in a dual-task procedure.

If both tasks were performed in a completely sequential manner, the resulting response-time distribution would be the convolution of the two original distributions. If both tasks were performed in parallel, the response-time distribution of the compound task would be identical to the distribution when the task is performed in isolation. However, the PRP model states that only the central processing stage is sequential and thus that some operations can be done in parallel. The precise RT2 distribution can still be calculated (see Sigman & Dehaene [2005] for the details) and, crucially, this calculation is parametric, since it depends on the duration of the M1 and P2 components. Confronting the distributions of the first and second tasks thus provides access to the durations of parallel processing stages, which are not directly accessible to measurement.

We found a very good fit of the ensemble of RT distributions for varying SOAs. The fit required an additional parameter—a rigid shift of 125 ms in time for all distributions of RT1 (see below for the rationale of this parameter). The durations of the parallel processing stages obtained from the interference experiment were essentially identical to the ones based on the shape of the RT distributions to the first task, and which yielded estimates of $1/a$ (the time of integration), and t_0 (a fixed delay).

More recently, we explored a further prediction of this model. If the effect of interference is simply to reschedule processing stages within a task, varying the timing but not the characteristics (duration, precision, variability, etc.) of each processing stage, the quality of the decision is unaffected by a concurrent task. We thus studied the functional dependencies of response times and error rates in a dual-task experiment where each task involved a non-symbolic (analogic) decision. In this situation, as in most decision comparison tasks with a scalar measure of the evidence, the extent to which two stimuli can be discriminated is determined by their ratio, referred to as the Weber fraction. Consistent with the predictions, we observed that: (1) Response times replicated the main features observed in prior dual-task studies, showing a strong delay in the execution of the second task, which decreases as SOA increases, and no effect on the first responded task. (2) On the contrary, we did not observe any significant effect on the total number of errors or on the distribution of errors as a function of the numerical distance between the target and the reference, suggesting that the decision process itself is delayed but its workings are unaffected by task interference (Kamienkowski & Sigman, 2008).

In summary, our simple model was essentially capable of predicting the precise shape of a large family of distributions, corresponding to different SOA values, task orders, and experimental manipulations of the number task. We found, however, a consistent departure from this simple model: A

systematic and constant increase on RT1 that became evident from direct exploration and from the necessity of including a rigid shift in time on RT1 to account for a correct convolution of both tasks in a model of dual-task performance. In the last section, we return to this subtle but important departure from the simple sequential model, which will be the gate to a new series of experiments.

SERIAL AND PARALLEL PROCESSING IN THE HUMAN BRAIN: DYNAMICS, ARCHITECTURE, AND NETWORKS

Methodological developments to understand dynamics and architecture in the human brain

What is the neurophysiological basis of the postulated processing, and particularly of the central decision stage? The purely passive model of the PRP interference predicts that task interference results exclusively from a change in the dynamics of the processing stages within each task, without any additional engagement of other brain areas. This constitutes a challenge for neurophysiology since it implies observing shifts in time without any change in total amount of activity during dual-task interference. This challenge explains why most previous studies have relied on Event Related Potential (ERPs). These experiments have systematically showed delayed (and on occasions also reduced) components, such as the N2PC, P3, and lateralized readiness potentials (Arnell & Duncan, 2002; Arnell, Helion, Hurdelbrink, & Pasiaka, 2004; Brisson & Jolicoeur, 2007a, 2007b; Brisson, Robitaille, & Jolicoeur, 2007; Dell'acqua, Jolicoeur, Vespignani, & Toffanin, 2005; Luck, 1998; Osman & Moore, 1993; Sessa, Luria, Verleger, & Dell'Acqua, 2007). Using time-resolved fMRI (Formisano & Goebel, 2003; Kim, Richter, & Ugurbil, 1997; Menon, Luknowsky, & Gati, 1998), Dux and collaborators showed delayed activity in prefrontal cortex in a PRP paradigm (Dux, Ivanoff, Asplund, & Marois, 2006), suggesting that a frontal network was one of the fundamental nodes responsible for the central bottleneck of information processing. None of those studies, however, provided a complete analysis of the neurocognitive task architecture at the whole-brain level.

We sought to achieve such a full decomposition of each task into processing stages, in order to understand their parallel or serial nature as we did in the previous series of behavioral experiments. This ambitious goal required an important methodological advance: We needed to estimate timing information invariantly across different brain regions, to distinguish changes in onset latency and in duration (Bellgowan, Saad, & Bandettini, 2003), and then to cluster the timing information into distinct stages based on a precise model of task sequencing.

We first demonstrated that fMRI could be used to recover the precise timing of all the stages in a complex composite task, reconstructing the controlled stream of brain activations in a sequence of cognitive operations (sensory, motor, verbal) (Sigman, Jobert, Lebihan, & Dehaene, 2007). Our Fourier-based methodology—described thoroughly in Sigman et al. (2007)—showed that single-event fMRI can evaluate changes in activation timing with a rather good precision of 100–200 ms. We then used this methodology in a simplified version of the PRP experiment described in the previous section. Only four SOA (two short and two long) values were used, the number task was presented second, and numbers were presented only in Arabic digits and with a single response modality.

We performed independent identical experiments with time-resolved fMRI and with high-density ERP recordings (Sigman & Dehaene, 2008). The timing information from both imaging techniques was clustered into components, guided by the psychological model of the task sequence described in the previous section. This allowed us to parse the execution of the two tasks into a series of processing stages with different timing properties, to understand which nodes were involved in one or both tasks or in coordinating dual-task execution, and which stages proceeded in parallel with each other or imposed a serial bottleneck.

Factorizing a physiological stream into response components and investigating their temporal superposition

To understand the dynamics of different brain processes involved in the dual-task condition, we first decomposed the ERP data using scalp templates identified from the ERP recorded at the largest SOA in which the execution of both tasks does not overlap in time. We simply identified the main topographies at each local maximum of the total voltage power recorded over all electrodes (Figure 1.4a). These components could easily be identified as the N1 and P3 components corresponding to each task. While more components could potentially have been identified by a more sophisticated analysis, our aim was not to identify all the independent processes within a task, but rather to understand the dynamics and architecture of these basic response components within the interference regime.

To do so, we decomposed each ERP, at each SOA and each time point, into a linear combination of the four scalp templates—this is referred to simply as the “time course of the ERP components”. We observed that the time course of the components (Figure 1.4b) fitted with predictions derived from our sequential model, if one supposes that the N1 components map onto perceptual processes and the P3 components onto central processes (Del Cul, Baillet, & Dehaene, 2007; Sergent, Baillet, &

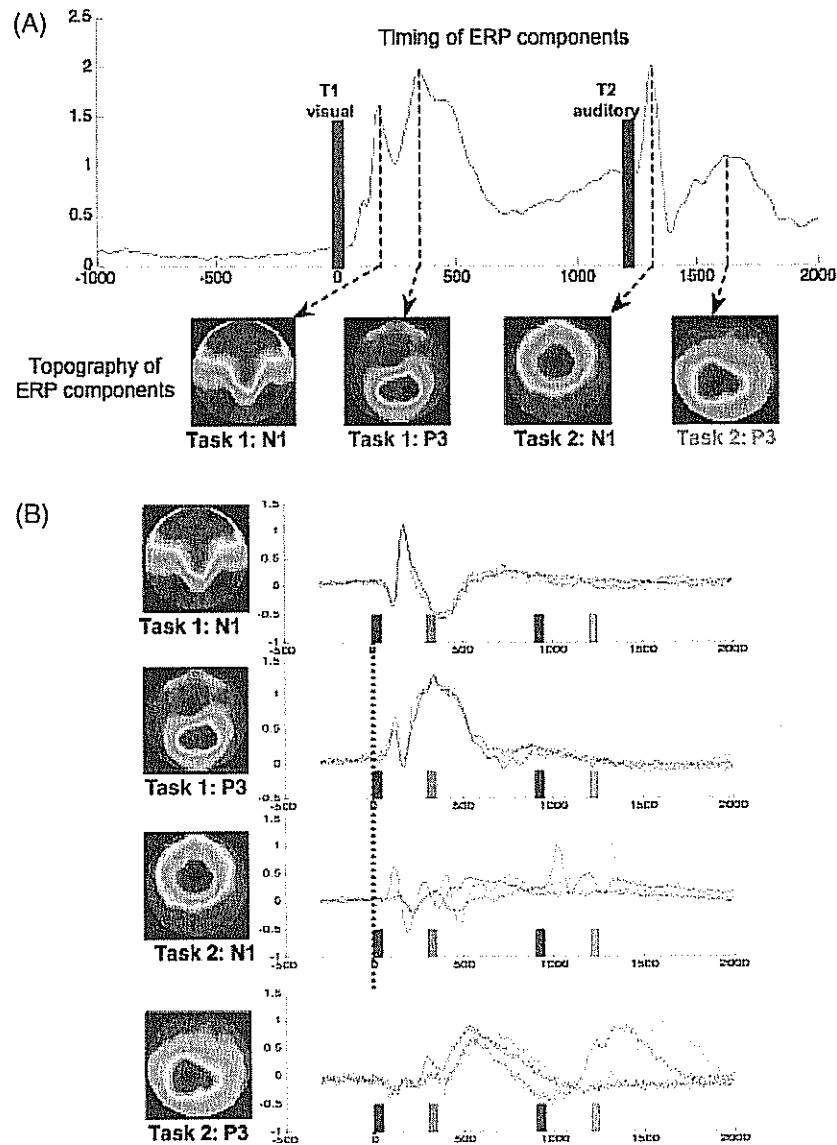


Figure 1.4 Dynamics of brain processes: N1 and P3. (a) Identifying the response components (N1 and P3), when the two tasks are performed without interference. (b) Dynamics of each component: Components of task 1 (first two panels) are unaffected by SOA. N1 component of T2 (third panel) follows stimulus presentation, P3 component reflects a bottleneck for short SOAs. See color plate.

Dehaene, 2005). The time course of the N1 and P3 components of the first task (denoted as T1_N1 and T1_P3) was unaffected by changes in SOA values (first and second rows in Figure 1.4b), indicating that, as predicted by the PRP model, the first task unfolded strictly identically within and outside the interference regime. This observation also testifies to the efficiency of the decomposition procedure, which was able to identify the visual components of the T1 task even when they were superimposed with simultaneously occurring auditory components.

The time course of the components of the second task showed a very distinct pattern. The T2_N1 component was strictly time locked to T2 onset, as expected for a perceptual component of task 2. It peaked at a fixed latency after T2 presentation, both within and outside the interference regime, and thus its latency increased linearly with SOA. The time course of the T2_P3 component of task 2, on the other hand, showed little effect of SOA within the interference regime and a shift proportional to the change in SOA in the non-interference regime, as expected for a central component of task 2. This is also expected from a motor component (since these predictions are valid for all post-bottleneck stages) and, given that the locus of the P3 for each task appears to be lateralized to the right for T1 and to the left for T2 (in accordance with the lateralization of the response of each hand), it is possible that this component may be indexing a motor component. Two observations make this interpretation unlikely: First, previous results have related the P3 to central processing (Sergent et al., 2005; Del Cul et al., 2007). Second, and most importantly, if the P3s indexed motor components, they could overlap in time (i.e., the P3 of the second task could be executed in parallel with the P3 of the first task). Our observation suggests the contrary, since T2_P3 appears to be locked to the ending of T1_P3.

Altogether, the data suggest that T2 presentation immediately engages a sensory processing stage, which unfolds as a series of damped oscillations over a period of about 300 ms after T2 onset (Figure 1.3, third row), followed by a central component that starts about 250 ms after T2 onset and peaks at 380 ms, but is systematically delayed by the simultaneous performance of another central process of a concurrent task.

As with the behavioral experiments, while the passive bottleneck model was capable of explaining the bulk of the observations, some details of the ERP analysis did suggest several departures from the simple model proposed: A modulation of the amplitude of the N1 component, reflecting sensory attenuation during concurrent task processing, the ramping of the N1 component, revealing task 2 sensory expectation after task 1 completion, and the emergence of the task 2 P3 component before stimulus presentation. These important departures will be discussed in the next section.

Brain networks of serial and parallel processing in the human brain

ERPs provide high temporal resolution, but they are notoriously imprecise for localization. Here, we took advantage of the fact that the PRP phenomenon induces large delays of several hundred milliseconds, which, as discussed previously, are measurable with fMRI (Sigman et al., 2007). For the context of this chapter, it is important to know that the analysis is based on the phase of the response that provides an estimate of the “temporal center of mass” and thus:

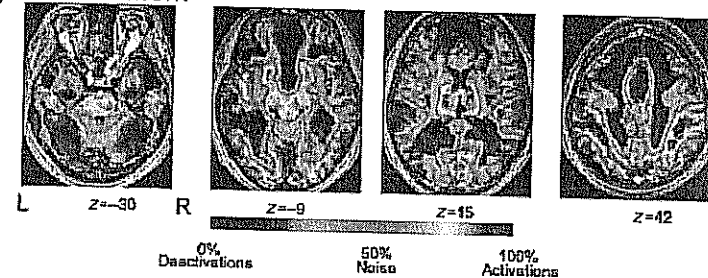
1. For a change in the onset of neural activation, only the phase of the hemodynamic response should vary, not the amplitude. The change in the phase should be identical to the change in delay.
2. For a change in duration of activation, both phase and amplitude should increase, with the slope of the phase change reflecting half of the actual change in the duration of neuronal activation (Sigman et al., 2007).

We therefore recorded whole-brain fMRI images at a sampling time (TR) of 1.5 s and computed the phase and amplitude of the hemodynamic response on each trial, for each subject and each voxel. A large network of brain areas (Figure 1.5a) exhibited phases consistently falling within the expected response latency for a task-induced activation (a liberal interval of 2–10 s). As expected for a complex dual-task experiment with visual and auditory stimuli, these regions included: (1) bilateral visual occipito-temporal cortices; (2) bilateral superior temporal auditory cortices; (3) motor, premotor, and cerebellar cortices; and (4) a large-scale bilateral parieto-frontal network. How is this large network organized in time during dual-task performance?

As with the ERPs, we relied on the model to determine our analytic strategy in order to parse the fMRI data (Figure 1.5b). Our method allowed us to fully parse the responsive brain areas into five different networks based on their temporal profile of response. The functional neuroanatomy of these networks (Figure 1.5b) was, for the most part, in tight accordance with the theoretical predictions.

For the brain network involved exclusively in task 1, timing of the activation, and hence the phase of the fMRI response, should not change with SOA. This first network (blue, no phase variation) comprised regions in extra-striate visual cortex, left motor-cortex, and the most medial part of the posterior-parietal cortex as well as an extended subcortical network. This network is expected for a visual number comparison task with a right-hand response.

(A) Active network



(B) Temporal parsing

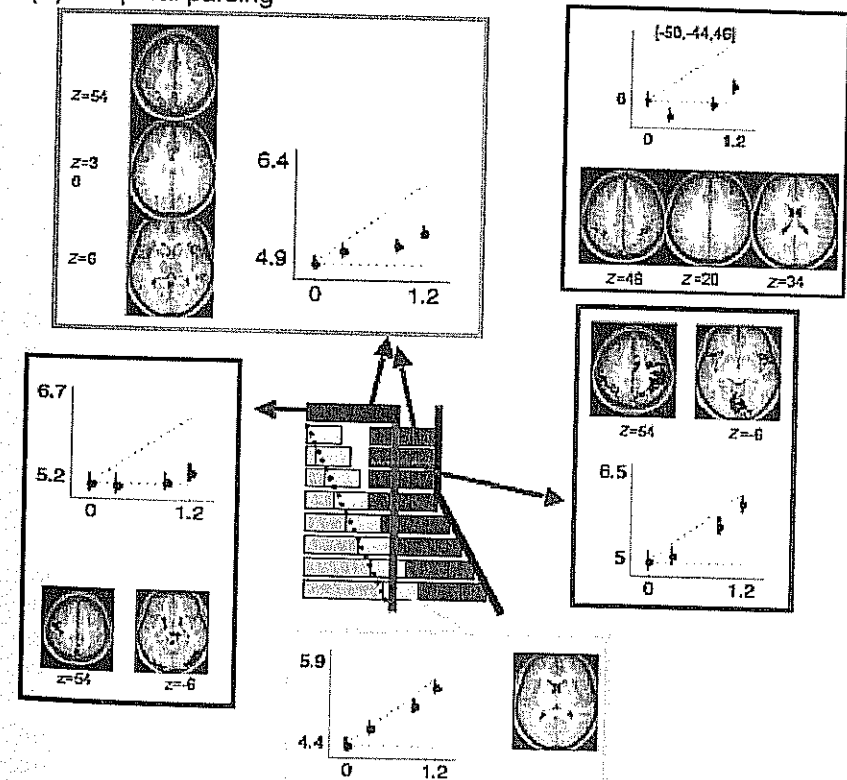


Figure 1.5 Dynamics of brain processes: fMRI. (a) Network activated (red) and inactivated (blue) during dual-task performance as identified by a phase coherence analysis. (b) Parsing the brain network in distinct dynamic processing stages, according to their phase profile. Blue: The brain network involved exclusively in task 1. Yellow: The brain network involved in the perceptual component of task 2. Red: The brain network involved in the central and motor components of task 2. Cyan: Regions involved in both tasks. Green: The brain network involved in executive coordination of both tasks. See color plate.

For the brain network involved in the perceptual component of task 2, the timing of activations, and hence the phase of the fMRI response, should increase in direct proportion with SOA. This second cluster (yellow, slope-1 linear phase response) was the smallest and involved exclusively bilateral auditory cortex, including Heschl's Gyrus and more lateral regions of temporal cortex—a plausible network for the sensory processing of the second task (auditory pitch judgment).

For the brain network involved in the central and motor components of task 2, activation should be unchanged for short SOA—due to queuing—and thus a delay should only be observed for long SOAs resulting in a non-linear phase dependence with SOA. fMRI voxels with this temporal profile were found in the right motor cortex, right Supplementary Motor Area (SMA) (remember that target T2 is responded to with the left hand), and bilateral intraparietal activation. Interestingly this cluster also included the most medial parts of the visual cortex. This unexpected finding might relate to the fact that subjects resumed attention to the fixation cross after conclusion of the two tasks.

The previous networks assume that the distinct dynamic processes of the two tasks always engage distinct brain regions, which of course is not a necessary implementation of the model—indeed it seems much more likely that some of the stages of tasks 1 and 2 engage identical brain regions. We therefore investigated one of the most interesting theoretical predictions—that the serial bottleneck may result from a broad shared network between both tasks, even when sensory and response modalities of both tasks are distinct. Regions involved in both tasks should show an increase in phase corresponding roughly to one half of these indicated by regions corresponding to task 2. We found that this profile corresponded to a massive cluster, involving an extended bilateral network that included the bilateral posterior parietal cortex, premotor cortex, SMA, anterior part of the insula, and the cerebellum. Actually, voxels involved in both tasks should increase with a slope of 0.5 in the noninterference regime, but show no increase in the interference regime, since the total duration of both tasks is unchanged (see Figure 1.2b). The resolution of our methodology was in the limit to resolve this issue and we could not distinguish reliably the slope of 0.5 from the slope of 0 in the interference regime.

For all these regions, if the effect of SOA is simply to alter the onset time of distinct processes, the amplitude of the fMRI activation should remain constant. This was indeed a very consistent experimental observation, in striking contrast with the observed broad repertoire of phase profiles. It indicates that all of the above changes corresponded to purely dynamical reorganizations of task components, without any change in activation strength.

In summary, we measured the phase of the fMRI response in the whole

brain at high temporal resolution in order to parse the responsive network into distinct processing stages. Sensory areas tracked the objective time of stimulus presentation, while a bilateral parieto-prefrontal network correlated with the dual-task sequential delay. An extended bilateral network, which included the bilateral posterior parietal cortex, premotor cortex, SMA, anterior part of the insula, and the cerebellum, was shared by both tasks during the extent of dual-task performance. Consistent with the EEG findings, this provides physiological evidence of the coexistence of serial and parallel processes within a cognitive task.

BEYOND THE SIMPLE SEQUENTIAL MODEL: A HIERARCHICAL SCHEME FOR DECISION MAKING AND EXECUTIVE CONTROL

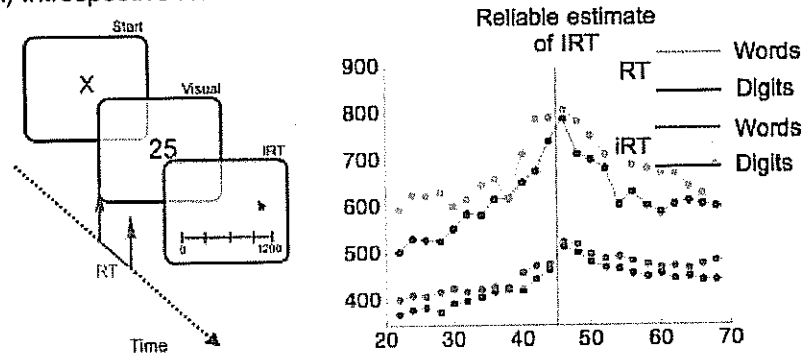
A consistent picture emerges from the previous behavioral and physiological experiments. A simple model of task architecture accounted for a large number of observations on mean and variability of response times, the precise shape of the distribution of response times of the first and second task in dual-task procedures, the effect of numerous experimental manipulations on these variables, and the precise dynamics of the different physiological components and networks identified with high resolution EEG and fMRI. However, in each instance, we observed minor but consistent and reproducible departures, which suggested that the model is necessarily incomplete.

The first marked departure came from our behavioral observations. Responses to the first task in the PRP paradigm, while independent of SOA, were found to be slower than when performing the task in isolation (Jiang, Saxe, & Kanwisher, 2004; Sigman & Dehaene, 2005). We reasoned that this may be related to an executive control stage engaged before performing the first task, in order to prepare for the instruction of performing the two tasks in a specific order. We hypothesized that in situations in which task order is unknown, this executive time should increase, reflecting a hierarchy of decision processes: First, in terms of which task to respond to and second, the specific decision involved in each task. We conducted a new series of experiments to understand the precise architecture of computations in this more ethological situation of unknown task order. Following a similar analysis of the effect of experimental factors on response-time distributions, we could conclude that in this situation of task uncertainty, executive components (engaging and disengaging in a task) had to be incorporated into our original model to account for critical behavioral observations (Sigman & Dehaene, 2006).

Evidence for the involvement of such executive components could also be derived from the human physiological data. In the ERP analysis we

observed that, while the timing of the peak of the N1 component of task 2 was in strict accordance with the predictions of the passive bottleneck model, several other observations deviated from this model. First, we observed a reduced amplitude of the sensory N1 within the interference regime compared with outside the interference regime. Second, the temporal course of the N1 component ramped prior to stimulus presentation, probably reflecting task expectation and preparation (Figure 1.6a). Finally, an auditory P3 component emerged at long SOAs, even before any

(A) Introspective RT is a reliable measure



(B) The dual-task delay is inaccessible to introspection

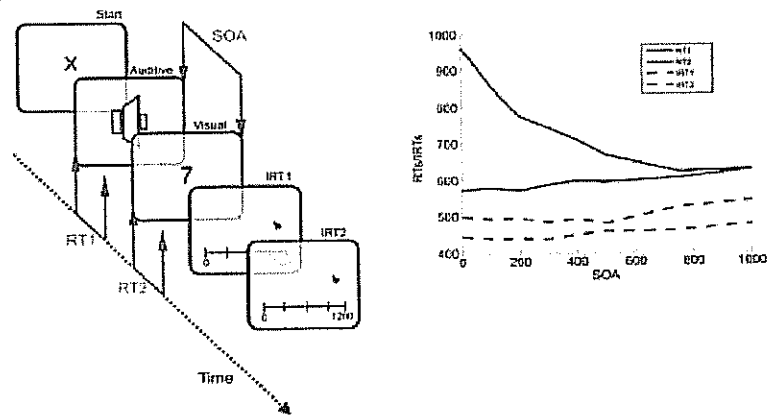


Figure 1.6 Temporal dynamics of processes. Introspection of response time and its failure during dual-task processing. (a) A number comparison task followed by an estimate of introspective response time (IRT) indicated that, although there is a consistent underestimation (reflecting a poor calibration of absolute IRT), subjects' estimate of IRT is very tightly correlated with RT. (b) On the contrary, the 300 ms delay of the PRP was unnoticed in introspection. Subjects always thought that they were taking a constant time for task 2, regardless of SOA. See color plate

auditory stimulus was presented. This anticipatory component peaked around 500 ms, thus coinciding nicely with the end of the visual P3 evoked by task 1. This ERP sequence is compatible with the hypothesis that, as soon as subjects completed task 1, they re-oriented their attention to prepare for task 2, reflecting an executive component of task engagement (Allport, Styles, & Hsieh, 1994; De Jong, 1993, 1995; Jentzsch, Leuthold, & Ulrich, 2007; Logan & Gordon, 2001; Meiran, Chorev, & Sapir, 2000; Ruthruff et al., 2001; Sigman & Dehaene, 2006). In addition, it suggests that the absence of top-down control may explain the amplitude attenuations observed during interference (Gilbert & Sigman, 2007).

We also found an indication of executive components involved in the active coordination of task processes in the fMRI experiment, by finding a cluster of voxels that was not expected by the purely sequential model. Those voxels showed a purely non-linear component with an increase in phase at the shortest SOA (Figure 1.5b, green cluster). This cluster, which was not predicted by the passive bottleneck model, involved exclusively a bilateral fronto-parietal network, previously found to be involved in processing bottlenecks in dual-task performance (Dux et al., 2006; Marois, Chun, & Gore, 2000; Marois and Ivanoff, 2005), in effortful but not in automatic tasks (Ashbridge, Walsh, & Cowey, 1997), and ubiquitously present in a large variety of goal-directed tasks (Duncan & Owen, 2000).

Cognitive theories have debated about the exact nature of the processes causing the PRP bottleneck. It has been suggested that it involves only a passive queuing of response-selection processes (Pashler, 1984) or, on the contrary, that it might involve a more extended set of processes that may vary with the exact nature of each experimental design, including executive components of task engaging and disengaging (Allport et al., 1994; Jentzsch et al., 2007; Logan & Gordon, 2001; Meiran et al., 2000; Ruthruff et al., 2001; Sigman & Dehaene, 2006, Pashler 1994) as well as delays in response initiation (De Jong, 1993, 1995; Meyer & Kieras 1997; Sigman & Dehaene, 2006). While a vast set of data can be explained simply in terms of response selection, certain aspects of the data argue clearly in favor of an "extended bottleneck" that may manifest in different manners related to executive function and consciousness, depending on context and specific paradigms. Indeed, we found it necessary to include executive components (engaging and disengaging in a task) in a more ethological situation of task uncertainty where the order of actions to achieve a complex goal is unknown. Furthermore, the extended array of areas affected by the PRP suggests that a broad array of processes causes the delay. Although this large set of areas might implement just a single cognitive stage of response selection, it seems more likely to correspond to the deployment of multiple hierarchically organized executive operations (Koechlin & Jubault, 2006; Koechlin, Ody, & Kouneiher, 2003).

We thus are led to envisage a model in which decisions occur serially at any of several possible levels of cognitive architecture. There are simple task decisions (Which key should I press for this stimulus?), but also higher-level executive decisions (Which of these two stimuli should I process first? Am I finished with the first stimulus? Can I orient attention to the second?). Importantly, our evidence so far suggests that all of these decisions involve overlapping parieto-prefrontal networks and are serially arranged in time, such that the need for a higher-level decision pushes all other lower-level decisions further back in time. This is reminiscent of Buridan's problem—sometimes our attempts to control our decisions actually delay our responses rather than improve them. However, the stochastic random-walk mechanism of decision making has one interesting property: Even when there is zero evidence favoring either alternative, the internal "Brownian motion" quickly leads to tipping over to one side or the other, in finite time only marginally slower than when actual evidence is present. Thus, we are never completely paralyzed by a decision—though we are often slowed by a serial chain of executive control operations.

A MOMENT OF THOUGHT: EVIDENCE FROM INTROSPECTION AND METACOGNITION

Our approach to understanding cognitive architecture involved a variety of methodologies. A common aspect to most of them was combining chronometry (through additive factor analysis) with other experimental tools such as interference, EEG, and time-resolved fMRI. Recently, we also investigated whether the chronometric technique could be used to address, in a precise quantitative manner, the introspective knowledge of one's own mental content. A fascinating question is whether there are limits to our own introspection (Hart, 1965; Jack & Roepstorff, 2003, 2004; Lutz, Lachaux, Martinerie, & Varela, 2002; Nelson, 1996; Overgaard, 2006; Wundt, 1897/1999), and whether introspective access to specific processing stages presents a precise, well-determined relation to our neurocognitive architecture. Indeed, it has often been speculated that a key property of the parallel stages of a task is that they operate outside of our awareness and are therefore opaque to introspection (e.g., Posner, Snyder, & Davidson, 1980). Might the decision stage also coincide with the conscious stage in task processing?

To explore which processing stages were accessible to introspection during the PRP, we again asked participants to perform the number comparison task, either in isolation or in a PRP design with a concurrent tone-discrimination. The novel aspect was that, following each of their responses, participants also indicated their introspective estimation of

response time (IRT) by clicking with the mouse on a continuous graded scale spanning 0 to 1200 ms and labeled every 300 ms.

The results we obtained were extremely clear cut (Figure 1.6). IRT turned out to be an extremely reliable and sensitive measure, tightly correlated with objective RT in a single-task context and sensitive to factors (i.e., notation) affecting response time by less than 50 ms. In a psychological refractory period task, however, the objective processing delay resulting from interference with a second concurrent task was totally absent from introspective estimates. That is, during the interference regime, the participants were totally unaware that their responses to the second task had been slowed by as much as 300 ms (Figure 1.6). Thus, awareness and accessibility to introspection seemed to be tightly linked to the availability of the central decision system—when this system was free, all stages of a cognitive task contributed to introspective response time, but when it was occupied by another task, participants were no longer aware of the duration of the perceptual and queuing stages of the PRP task.

DISCUSSION: COGNITIVE ARCHITECTURE OF THE NEVER-HALTING MACHINE

Several cognitive theories shared the hypothesis that while most mental and neural operations are modular, certain specific controlled processes require a distinct functional architecture that can establish flexible links among existing processors. This process has been called the central executive (Baddeley, 1986), the supervisory attentional system (Shallice, 1988), the anterior attention system (Posner, 1994), the capacity limited stage (Chun & Potter, 1995), the global workspace (Baars, 1989; Dehaene, Kerszberg, & Changeux, 1998), or the dynamic core (Tononi & Edelman, 1998). Yet in all of these different versions, it shares three critical aspects: (1) binding of information from many different modules; (2) uniqueness of its contents; and (3) sustained activity for a few hundred milliseconds. These aspects are also characteristic of consciousness and thus it has been proposed that engagement of the central system may be a requirement for conscious processing (Baars, 1989; Dehaene & Naccache, 2001; Tononi & Edelman, 1998). It is proposed, in a two-stage model, that an early stage permits the rapid, initial evaluation of the visual world. Only during a second, capacity-limited stage, does information become conscious when the neural population that represents it is mobilized into a brain-scale state of coherent activity that involves many neurons distributed throughout an extended network of the brain (Baars, 1989; Dehaene & Naccache, 2001; Tononi & Edelman, 1998).

Combining models from psychophysics and human physiology, we could disentangle the unfolding of processing stages during a compound

cognitive task and relate this architecture of stages to their dynamic organization and stochasticity. The data converged to a model that proposes:

- an initial perceptual processing of incoming stimulus, which is performed in a modular (parallel) fashion and is very reliable in time, as indexed by the N1 latency of the event-related potential
- followed by a central process involved in the flexible coordination of information according to specific task requirements, which is intrinsically serial and involves a stochastic integration, as traditionally used to model decision making in single tasks, and as indexed by the P3 component of the ERP.

Our findings of the dynamics of T2 processing in the PRP fit well with the ERPs of other interference experiments such as the attentional-blink and masking experiments (Del Cul et al., 2007; Sergent et al., 2005). In these experiments it was found that the first ~270 ms, indexed by P1 and N1 components, were independent of conscious access and this was followed sequentially by a central distributed workspace, indexed by a P3 component, involving prefrontal and parietal network, which was engaged only in trials in which the stimulus accessed conscious processing.

These results have led to the proposal that the P3 component of the EEG may be related to access to a global coherent workspace associated with flexible coordination of information, which in turn may mediate conscious reportability (Del Cul et al., 2007; Sergent et al., 2005). According to this theory, a distributed set of neurons with long axons provides a global “broadcasting” system enabling communication between arbitrary and otherwise not directly connected brain processors (Baars, 1989; Dehaene et al., 1998; Dehaene & Naccache, 2001). Global neuronal workspace theory can explain why response selection generally imposes a dual-task bottleneck. In most psychological tasks, the relation between stimuli and responses is entirely arbitrary and thus requires the temporary mapping between otherwise independent processors. Establishing such a new arbitrary interconnection should involve central workspace mediation.

Supporting this interpretation, interference is drastically reduced for highly practiced or non-arbitrary tasks (Greenwald, 2003; Lien, McCann, Ruthruff, & Proctor, 2005; Lien, Proctor, & Allen, 2002). More evidence for the existence of a common global workspace involved in conscious processing and in the executive and flexible coordination of processing modules results from behavioral experiments that have combined the basic features of different manifestations of central processing such as the PRP (two rapid responses) or the attentional blink (extinction of a second rapidly presented stimulus). These experiments suggest that processing limitations may arise in part from a common bottleneck (Jolicoeur, 1999;

Marois & Ivanoff, 2005; Wong, 2002). Finally, our observation that the objective processing delay resulting from interference with a second concurrent task is totally absent from introspective estimates further suggests that subjective introspection of time spent on a task tightly correlates with the period of availability of global workspace activity (Corallo, Sackur, Dehaene, & Sigman, 2008).

Simply to illustrate these ideas, we provide a basic sketch of the postulated brain dynamics during the blink and the PRP (Figure 1.7). In both experimental designs, certain aspects of T2 processing—indexed by the N1 and mediated by a transient activation of sensory cortex—can proceed while T1 is engaged in a broad and coherent state of processing occupying the central workspace. While the nature of this sensory memory of T2 is not fully understood and requires theoretical and experimental investigation (Gegenfurtner & Sperling, 1993; Graziano & Sigman, 2008; Sperling, 1960), it appears to constitute a labile form of memory that can be overridden by the presence of a mask, consistent with current theories and experimental evidence of masking by object substitution (Del Cul et al., 2007; Di Lollo, Enns, & Rensink, 2000). The presence of a mask after

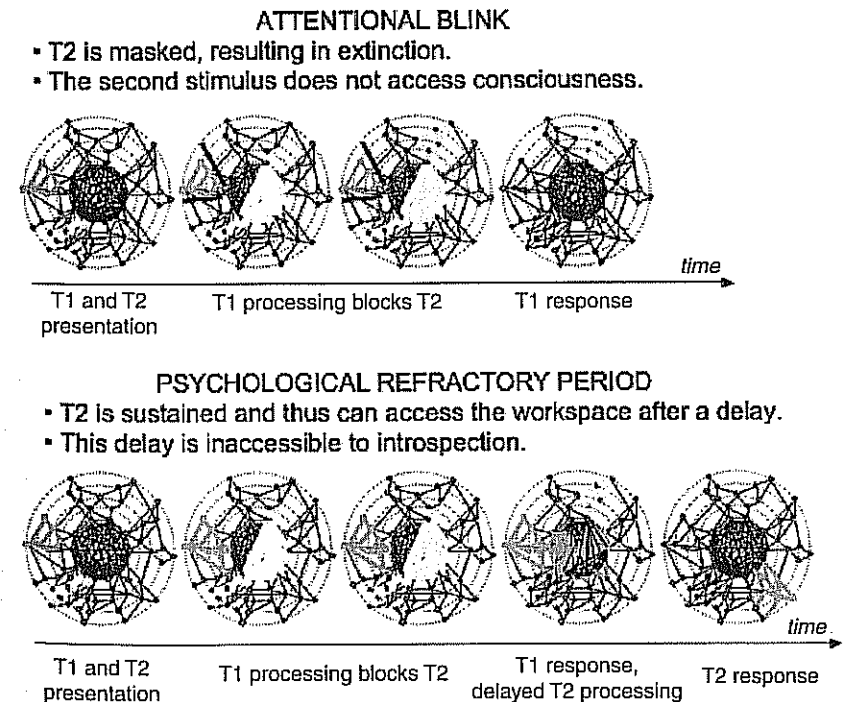


Figure 1.7 A simple sketch of the dynamics of modular and workspace processing during the PRP and the attentional blink. See color plate.

T2 can thus lead to the attentional blink (AB) phenomenon, whereby participants fail to detect the presence of T2 although it demonstrably caused a strong initial perceptual activation. In the absence of masking, T2 can access the workspace and T2 is therefore processed with the delay observed as the PRP phenomenon. AB and PRP are therefore envisaged as two very similar phenomena, arising from the same constraints of parallel-serial architecture.

While there are no direct correlates of these findings in single-cell awake monkey physiology (there is currently no demonstration of the PRP in non-human primates), some chronometric aspects can be related to the dynamics of neuronal activation in single task. The neurophysiological bases of simple perceptual decision making have been widely studied in tactile (Hernandez, Zainos, & Romo, 2002; Romo, Hernandez, & Zainos, 2004; Romo & Salinas, 1999) and visual discrimination tasks (Bichot, Schall, & Thompson, 1996; Britten, Newsome, Shadlen, Celebrini, & Movshon, 1996; Britten, Shadlen, Newsome, & Movshon, 1992; Hanes & Schall, 1996; Schall & Bichot, 1998; Shadlen & Newsome, 2001; Thompson, Bichot, & Schall, 1997; Thompson, Hanes, Bichot, & Schall, 1996). These studies have revealed direct physiological correlates of the accumulation process postulated in formal response-time models. Some neurons appear to code for the current perceptual state. For instance, neurons in the middle temporal visual area (MT) appear to encode the amount of evidence for motion in a certain direction (Britten et al., 1992; Shadlen & Newsome, 1996). Other neurons, distributed in multiple areas, including posterior parietal cortex, dorsolateral prefrontal cortex, and frontal eye fields, appear to integrate this sensory information and thus show stochastically increasing firing rates in the course of decision making (Kim & Shadlen, 1999; Shadlen & Newsome, 1996, 2001). In agreement with the accumulation model of decision making, the rate of increase varies with the quality of sensory evidence (Mazurek, Roitman, Ditterich, & Shadlen, 2003; Shadlen & Newsome, 1996, 2001), and the response is emitted when the firing exceeds a threshold (Hanes & Schall, 1996). Furthermore, accumulation of information about the upcoming response appears in the firing train after a latency of about 200 ms (Gold & Shadlen, 2000; Roitman & Shadlen, 2002), which is relatively fixed for a given task and might thus index the duration of the initial perceptual stage.

Further evidence of such a dynamic arrangement comes from recordings from the primary visual cortex of awake monkeys, which have shown that a visual stimulus evokes a first transient response, followed by a sustained wave of activity (Lamme & Roelfsema, 2000; Lamme, Super, Landman, Roelfsema, & Spekreijse, 2000; Lee, Yang, Romero, & Mumford, 2002; Li, Piech, & Gilbert, 2006, 2008; Luck, Chelazzi, Hillyard, & Desimone, 1997; Reynolds, Chelazzi, & Desimone, 1999;

Roelfsema, Lamme, & Spekreijse, 1998; Roelfsema, Lamme, Spekreijse, & Bosch, 2002) at a latency of about 200 ms, nicely coincident with the latency of integration in parietal and prefrontal areas and of the engagement of the P3 process. These two responses have very different functional dependencies: The first transient response is largely determined by stimulus properties and can be explained by classical bottom-up receptive field properties. On the contrary, the second response can be modulated by different cognitive and contextual factors. For instance, it is amplified if the local stimulus is salient and attended (as in figure-ground experiments) and it can virtually disappear if the stimulus is masked (which presumably precludes its access to consciousness) or by anesthetics. This specificity suggests an engagement of the central workspace system in the second wave of activity and that the same neuron may be involved in distinct processing stages within the same task. Further experiments are required to determine whether this second wave of activity shows a dual-task delay characteristic of the serial processing bottleneck.

Altogether, neurophysiological and brain-imaging studies suggest that, beyond an initial and relatively reliable perceptual delay of about 200 ms, a decision stage begins that involves a process of stochastic accumulation of evidence and the joint activation of a distributed network of areas, with partially changing topography as a function of the nature of the task, but with frequent if not systematic co-activation of parietal and premotor regions. Our results suggest that this accumulation system is slow (which might be required to flexibly combine information in different sensory, mnemonic, and motor modalities), is variable (probably due to an active, dynamic reverberating and integrating state), and is responsible for establishing the PRP bottleneck. Thus, while certain neural computations (probably mediated by hard-wired circuits confined to small portions of sensory space) can be very fast, precise, and parallel, linking such information together in a coherent workspace results in a slow, variable, and intrinsically sequential computation.

The challenge for the next few years will be to understand which precise biophysical mechanisms are involved in this coherent workspace and why they result in such dynamic and architectonically distinctive properties. At present, our results simply suggest that this bottleneck might occur because the cerebral accumulation system is broadly distributed and largely shared across tasks, and thus must be entirely "mobilized", at any given moment, by whichever task is currently performed. Considerable research has examined the neurophysiology of a single cognitive operation, but much less is known on how we chain these basic operations into complex tasks. The results described here constitute a first step toward a progressive understanding of the chaining of simple computations into complex compound tasks.

REFERENCES

- Allport, D., Styles, E., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421–452). Cambridge, MA: MIT Press.
- Arnell, K. M., & Duncan, J. (2002). Separate and shared sources of dual-task cost in stimulus identification and response selection. *Cognitive Psychology, 44*, 105–147.
- Arnell, K. M., Helion, A. M., Hurdelbrink, J. A., & Pasiëka, B. (2004). Dissociating sources of dual-task interference using human electrophysiology. *Psychonomic Bulletin & Review, 11*, 77–83.
- Ashbridge, E., Walsh, V., & Cowey, A. (1997). Temporal aspects of visual search studied by transcranial magnetic stimulation. *Neuropsychologia, 35*, 1121–1131.
- Baars, B. J. (1989). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.
- Bellgowan, P. S., Saad, Z. S., & Bandettini, P. A. (2003). Understanding neural system dynamics through task modulation and measurement of functional MRI amplitude, latency, and width. *Proceedings of the National Academy of Sciences USA, 100*, 1415–1419.
- Bichot, N. P., Schall, J. D., & Thompson, K. G. (1996). Visual feature selectivity in frontal eye fields induced by experience in mature macaques. *Nature, 381*, 697–699.
- Brisson, B., & Jolicoeur, P. (2007a). Cross-modal multitasking processing deficits prior to the central bottleneck revealed by event-related potentials. *Neuropsychologia, 45*, 3038–3053.
- Brisson, B., & Jolicoeur, P. (2007b). A psychological refractory period in access to visual short-term memory and the deployment of visual-spatial attention: Multitasking processing deficits revealed by event-related potentials. *Psychophysiology, 44*, 323–333.
- Brisson, B., Robitaille, N., & Jolicoeur, P. (2007). Stimulus intensity affects the latency but not the amplitude of the N2pc. *Neuroreport, 18*, 1627–1630.
- Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience, 13*, 87–100.
- Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience, 12*, 4745–4765.
- Bryant, H. L., & Segundo, J. P. (1976). Spike initiation by transmembrane current: A white-noise analysis. *Journal of Physiology, 260*, 279–314.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523–547.
- Butts, D. A., Weng, C., Jin, J., Yeh, C. I., Lesica, N. A., Alonso, J. M. et al. (2007). Temporal precision in the neural code and the timescales of natural vision. *Nature, 449*, 92–95.
- Cecchi, G. A., Sigman, M., Alonso, J. M., Martinez, L., & Chialvo, D. (2000). Noise in neurons is message dependent. *Proceedings of the National Academy of Sciences USA, 97*, 5557–5561.
- 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*, 109–127.
- Corallo, G., Sackur, J., Dehaene, S., & Sigman, M. (2008). Limits on introspection: Distorted subjective time during the dual-task bottleneck. *Psychological Science, 19*, 1110–1117.
- De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 965–980.
- De Jong, R. (1995). The role of preparation in overlapping-task performance. *Quarterly Journal of Experimental Psychology: A, 48*, 2–25.
- de Ruyter van Steveninck, R. R., Lewen, G. D., Strong, S. P., Koberle, R., & Bialek, W. (1997). Reproducibility and variability in neural spike train. *Science, 275*, 1805–1808.
- Dehaene, S. (1996). The organization of brain activation in number comparison: Event-related potential and the additive-factor methods. *Journal of Cognitive Neuroscience, 8*, 47–68.
- 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 USA, 95*, 14529–14534.
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition, 79*, 1–37.
- Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology, 5*, e260.
- Dell'acqua, R., Jolicoeur, P., Vespignani, F., & Toffanin, P. (2005). Central processing overlap modulates P3 latency. *Experimental Brain Research, 165*, 54–68.
- Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129*, 481–507.
- Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23*, 475–483.
- Dux, P. E., Ivanoff, J., Asplund, C. L., & Marois, R. (2006). Isolation of a central bottleneck of information processing with time-resolved fMRI. *Neuron, 52*, 1109–1120.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1*, 1–47.
- Formisano, E., & Goebel, R. (2003). Tracking cognitive processes with functional MRI mental chronometry. *Current Opinion in Neurobiology, 13*, 174–181.
- Gegenfurtner, K. R., & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 845–866.
- Gilbert, C. D., & Sigman, M. (2007). Brain states: Top-down influences in sensory processing. *Neuron, 54*, 677–696.
- Gold, J. I., & Shadlen, M. N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature, 404*, 390–394.
- Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences, 5*, 10–16.

- Graziano, M., & Sigman, M. (2008). The dynamics of sensory buffers: Geometric, spatial and experience dependent shaping of iconic memory. *Journal of Vision*, 8, 1–13.
- 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, 859–868.
- Gur, M., Beylin, A., & Snodderly, D. M. (1997). Response variability of neurons in primary visual cortex (V1) of alert monkeys. *Journal of Neuroscience*, 17, 2914–2920.
- Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science*, 274, 427–430.
- Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, 56, 208–216.
- Hernandez, A., Zainos, A., & Romo, R. (2002). Temporal evolution of a decision-making process in medial premotor cortex. *Neuron*, 33, 959–972.
- Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology*, 148, 574–591.
- Hubel, D. H., & Wiesel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 28, 229–289.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 148, 574–591.
- Jack, A. I., & Roepstorff, A. (2003). Trusting the subject I. *Journal of Consciousness Studies*, 9, 9–10.
- Jack, A. I., & Roepstorff, A. (2004). Trusting the subject II. *Journal of Consciousness Studies*, 11, 7–8.
- Jentsch, I., Leuthold, H., & Ulrich, R. (2007). Decomposing sources of response slowing in the PRP paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 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, 390–396.
- Jolicoeur, P. (1999). Concurrent response-selection demands modulate the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1097–1113.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Kamienkowski, J. E., & Sigman, M. (2008). Delays without mistakes: Response time and error distributions in dual-task. *PLoS ONE*, 3, e-3196.
- Kara, P., Reinagel, P., & Reid, R. C. (2000). Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron*, 27, 635–646.
- Kim, J. N., & Shadlen, M. N. (1999). Neural correlates of a decision in the dorso-lateral prefrontal cortex of the macaque. *Nature Neuroscience*, 2, 176–185.
- Kim, S. G., Richter, W., & Ugurbil, K. (1997). Limitations of temporal resolution in functional MRI. *Magnetic Resonance in Medicine*, 37, 631–636.
- Koechlin, E., & Jubault, T. (2006). Broca's area and the hierarchical organization of human behavior. *Neuron*, 50, 963–974.
- Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, 302, 1181–1185.
- Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, 23, 571–579.
- Lamme, V. A., Super, H., Landman, R., Roelfsema, P. R., & Spekreijse, H. (2000). The role of primary visual cortex (V1) in visual awareness. *Vision Research*, 40, 1507–1521.
- Lee, T. S., Yang, C. F., Romero, R. D., & Mumford, D. (2002). Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. *Nature Neuroscience*, 5, 589–597.
- Li, W., Piech, V., & Gilbert, C. D. (2006). Contour saliency in primary visual cortex. *Neuron*, 50, 951–962.
- Li, W., Piech, V., & Gilbert, C. D. (2008). Learning to link visual contours. *Neuron*, 57, 442–451.
- 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, 122–144.
- 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, 396–409.
- Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, 40, 77–111.
- Logan, G. D., & Gordon, R. (2001). Executive control of visual attention in dual tasks. *Psychological Review*, 108, 393–434.
- Luce, R. D. (1986). *Response times*. New York: Oxford University Press.
- Luck, S. J. (1998). Sources of dual-task interference: Evidence from human electrophysiology. *Psychological Science*, 9, 223–227.
- Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2 and V4 of macaque visual cortex. *Journal of Neurophysiology*, 77, 24–42.
- Lutz, A., Lachaux, J. P., Martinerie, J., & Varela, F. J. (2002). Guiding the study of brain dynamics by using first-person data: Synchrony patterns correlate with ongoing conscious states during a simple visual task. *Proceedings of the National Academy of Sciences USA*, 99, 1586–1591.
- Mainen, Z. F., & Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science*, 268, 1503–1506.
- Marois, R., Chun, M. M., & Gore, J. C. (2000). Neural correlates of the attentional blink. *Neuron*, 28, 299–308.
- Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences*, 9, 296–305.
- Mazurek, M. E., Roitman, J. D., Ditterich, J., & Shadlen, M. N. (2003). A role for neural integrators in perceptual decision making. *Cerebral Cortex*, 13, 1257–1269.
- Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology*, 41, 211–253.
- Menon, R. S., Luknowsky, D. C., & Gati, J. S. (1998). Mental chronometry using

We thus are led to envisage a model in which decisions occur serially at any of several possible levels of cognitive architecture. There are simple task decisions (Which key should I press for this stimulus?), but also higher-level executive decisions (Which of these two stimuli should I process first? Am I finished with the first stimulus? Can I orient attention to the second?). Importantly, our evidence so far suggests that all of these decisions involve overlapping parieto-prefrontal networks and are serially arranged in time, such that the need for a higher-level decision pushes all other lower-level decisions further back in time. This is reminiscent of Buridan's problem—sometimes our attempts to control our decisions actually delay our responses rather than improve them. However, the stochastic random-walk mechanism of decision making has one interesting property: Even when there is zero evidence favoring either alternative, the internal "Brownian motion" quickly leads to tipping over to one side or the other, in finite time only marginally slower than when actual evidence is present. Thus, we are never completely paralyzed by a decision—though we are often slowed by a serial chain of executive control operations.

A MOMENT OF THOUGHT: EVIDENCE FROM INTROSPECTION AND METACOGNITION

Our approach to understanding cognitive architecture involved a variety of methodologies. A common aspect to most of them was combining chronometry (through additive factor analysis) with other experimental tools such as interference, EEG, and time-resolved fMRI. Recently, we also investigated whether the chronometric technique could be used to address, in a precise quantitative manner, the introspective knowledge of one's own mental content. A fascinating question is whether there are limits to our own introspection (Hart, 1965; Jack & Roepstorff, 2003, 2004; Lutz, Lachaux, Martinerie, & Varela, 2002; Nelson, 1996; Overgaard, 2006; Wundt, 1897/1999), and whether introspective access to specific processing stages presents a precise, well-determined relation to our neurocognitive architecture. Indeed, it has often been speculated that a key property of the parallel stages of a task is that they operate outside of our awareness and are therefore opaque to introspection (e.g., Posner, Snyder, & Davidson, 1980). Might the decision stage also coincide with the conscious stage in task processing?

To explore which processing stages were accessible to introspection during the PRP, we again asked participants to perform the number comparison task, either in isolation or in a PRP design with a concurrent tone-discrimination. The novel aspect was that, following each of their responses, participants also indicated their introspective estimation of

response time (IRT) by clicking with the mouse on a continuous graded scale spanning 0 to 1200 ms and labeled every 300 ms.

The results we obtained were extremely clear cut (Figure 1.6). IRT turned out to be an extremely reliable and sensitive measure, tightly correlated with objective RT in a single-task context and sensitive to factors (i.e., notation) affecting response time by less than 50 ms. In a psychological refractory period task, however, the objective processing delay resulting from interference with a second concurrent task was totally absent from introspective estimates. That is, during the interference regime, the participants were totally unaware that their responses to the second task had been slowed by as much as 300 ms (Figure 1.6). Thus, awareness and accessibility to introspection seemed to be tightly linked to the availability of the central decision system—when this system was free, all stages of a cognitive task contributed to introspective response time, but when it was occupied by another task, participants were no longer aware of the duration of the perceptual and queuing stages of the PRP task.

DISCUSSION: COGNITIVE ARCHITECTURE OF THE NEVER-HALTING MACHINE

Several cognitive theories shared the hypothesis that while most mental and neural operations are modular, certain specific controlled processes require a distinct functional architecture that can establish flexible links among existing processors. This process has been called the central executive (Baddeley, 1986), the supervisory attentional system (Shallice, 1988), the anterior attention system (Posner, 1994), the capacity limited stage (Chun & Potter, 1995), the global workspace (Baars, 1989; Dehaene, Kerszberg, & Changeux, 1998), or the dynamic core (Tononi & Edelman, 1998). Yet in all of these different versions, it shares three critical aspects: (1) binding of information from many different modules; (2) uniqueness of its contents; and (3) sustained activity for a few hundred milliseconds. These aspects are also characteristic of consciousness and thus it has been proposed that engagement of the central system may be a requirement for conscious processing (Baars, 1989; Dehaene & Naccache, 2001; Tononi & Edelman, 1998). It is proposed, in a two-stage model, that an early stage permits the rapid, initial evaluation of the visual world. Only during a second, capacity-limited stage, does information become conscious when the neural population that represents it is mobilized into a brain-scale state of coherent activity that involves many neurons distributed throughout an extended network of the brain (Baars, 1989; Dehaene & Naccache, 2001; Tononi & Edelman, 1998).

Combining models from psychophysics and human physiology, we could disentangle the unfolding of processing stages during a compound

cognitive task and relate this architecture of stages to their dynamic organization and stochasticity. The data converged to a model that proposes:

- an initial perceptual processing of incoming stimulus, which is performed in a modular (parallel) fashion and is very reliable in time, as indexed by the N1 latency of the event-related potential
- followed by a central process involved in the flexible coordination of information according to specific task requirements, which is intrinsically serial and involves a stochastic integration, as traditionally used to model decision making in single tasks, and as indexed by the P3 component of the ERP.

Our findings of the dynamics of T2 processing in the PRP fit well with the ERPs of other interference experiments such as the attentional-blink and masking experiments (Del Cul et al., 2007; Sergent et al., 2005). In these experiments it was found that the first ~270 ms, indexed by P1 and N1 components, were independent of conscious access and this was followed sequentially by a central distributed workspace, indexed by a P3 component, involving prefrontal and parietal network, which was engaged only in trials in which the stimulus accessed conscious processing.

These results have led to the proposal that the P3 component of the EEG may be related to access to a global coherent workspace associated with flexible coordination of information, which in turn may mediate conscious reportability (Del Cul et al., 2007; Sergent et al., 2005). According to this theory, a distributed set of neurons with long axons provides a global “broadcasting” system enabling communication between arbitrary and otherwise not directly connected brain processors (Baars, 1989; Dehaene et al., 1998; Dehaene & Naccache, 2001). Global neuronal workspace theory can explain why response selection generally imposes a dual-task bottleneck. In most psychological tasks, the relation between stimuli and responses is entirely arbitrary and thus requires the temporary mapping between otherwise independent processors. Establishing such a new arbitrary interconnection should involve central workspace mediation.

Supporting this interpretation, interference is drastically reduced for highly practiced or non-arbitrary tasks (Greenwald, 2003; Lien, McCann, Ruthruff, & Proctor, 2005; Lien, Proctor, & Allen, 2002). More evidence for the existence of a common global workspace involved in conscious processing and in the executive and flexible coordination of processing modules results from behavioral experiments that have combined the basic features of different manifestations of central processing such as the PRP (two rapid responses) or the attentional blink (extinction of a second rapidly presented stimulus). These experiments suggest that processing limitations may arise in part from a common bottleneck (Jolicoeur, 1999;

Marois & Ivanoff, 2005; Wong, 2002). Finally, our observation that the objective processing delay resulting from interference with a second concurrent task is totally absent from introspective estimates further suggests that subjective introspection of time spent on a task tightly correlates with the period of availability of global workspace activity (Corallo, Sackur, Dehaene, & Sigman, 2008).

Simply to illustrate these ideas, we provide a basic sketch of the postulated brain dynamics during the blink and the PRP (Figure 1.7). In both experimental designs, certain aspects of T2 processing—indexed by the N1 and mediated by a transient activation of sensory cortex—can proceed while T1 is engaged in a broad and coherent state of processing occupying the central workspace. While the nature of this sensory memory of T2 is not fully understood and requires theoretical and experimental investigation (Gegenfurtner & Sperling, 1993; Graziano & Sigman, 2008; Sperling, 1960), it appears to constitute a labile form of memory that can be overridden by the presence of a mask, consistent with current theories and experimental evidence of masking by object substitution (Del Cul et al., 2007; Di Lollo, Enns, & Rensink, 2000). The presence of a mask after

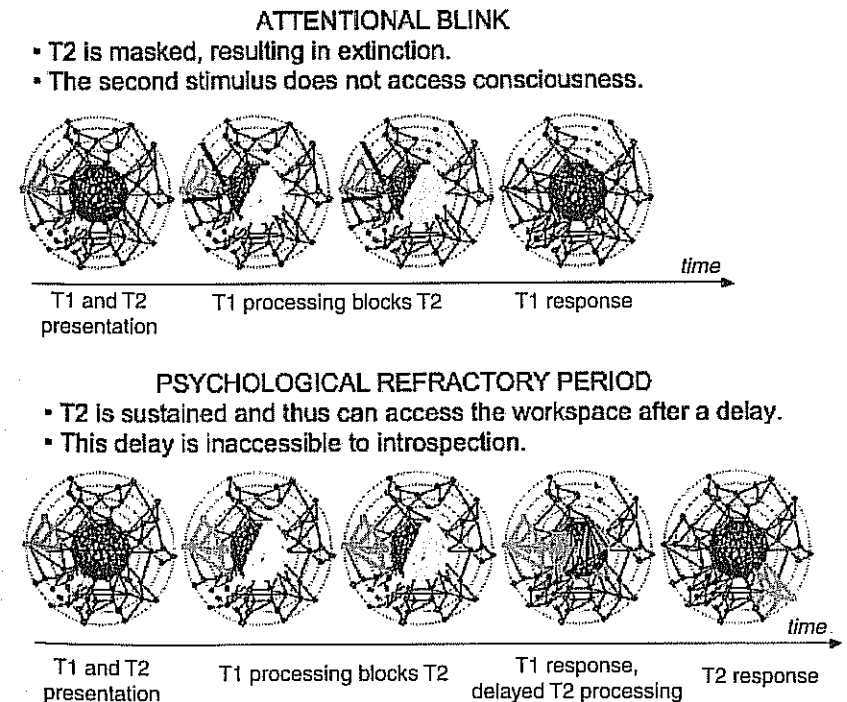


Figure 1.7 A simple sketch of the dynamics of modular and workspace processing during the PRP and the attentional blink. See color plate.

T2 can thus lead to the attentional blink (AB) phenomenon, whereby participants fail to detect the presence of T2 although it demonstrably caused a strong initial perceptual activation. In the absence of masking, T2 can access the workspace and T2 is therefore processed with the delay observed as the PRP phenomenon. AB and PRP are therefore envisaged as two very similar phenomena, arising from the same constraints of parallel-serial architecture.

While there are no direct correlates of these findings in single-cell awake monkey physiology (there is currently no demonstration of the PRP in non-human primates), some chronometric aspects can be related to the dynamics of neuronal activation in single task. The neurophysiological bases of simple perceptual decision making have been widely studied in tactile (Hernandez, Zainos, & Romo, 2002; Romo, Hernandez, & Zainos, 2004; Romo & Salinas, 1999) and visual discrimination tasks (Bichot, Schall, & Thompson, 1996; Britten, Newsome, Shadlen, Celebrini, & Movshon, 1996; Britten, Shadlen, Newsome, & Movshon, 1992; Hanes & Schall, 1996; Schall & Bichot, 1998; Shadlen & Newsome, 2001; Thompson, Bichot, & Schall, 1997; Thompson, Hanes, Bichot, & Schall, 1996). These studies have revealed direct physiological correlates of the accumulation process postulated in formal response-time models. Some neurons appear to code for the current perceptual state. For instance, neurons in the middle temporal visual area (MT) appear to encode the amount of evidence for motion in a certain direction (Britten et al., 1992; Shadlen & Newsome, 1996). Other neurons, distributed in multiple areas, including posterior parietal cortex, dorsolateral prefrontal cortex, and frontal eye fields, appear to integrate this sensory information and thus show stochastically increasing firing rates in the course of decision making (Kim & Shadlen, 1999; Shadlen & Newsome, 1996, 2001). In agreement with the accumulation model of decision making, the rate of increase varies with the quality of sensory evidence (Mazurek, Roitman, Ditterich, & Shadlen, 2003; Shadlen & Newsome, 1996, 2001), and the response is emitted when the firing exceeds a threshold (Hanes & Schall, 1996). Furthermore, accumulation of information about the upcoming response appears in the firing train after a latency of about 200 ms (Gold & Shadlen, 2000; Roitman & Shadlen, 2002), which is relatively fixed for a given task and might thus index the duration of the initial perceptual stage.

Further evidence of such a dynamic arrangement comes from recordings from the primary visual cortex of awake monkeys, which have shown that a visual stimulus evokes a first transient response, followed by a sustained wave of activity (Lamme & Roelfsema, 2000; Lamme, Super, Landman, Roelfsema, & Spekreijse, 2000; Lee, Yang, Romero, & Mumford, 2002; Li, Piech, & Gilbert, 2006, 2008; Luck, Chelazzi, Hillyard, & Desimone, 1997; Reynolds, Chelazzi, & Desimone, 1999;

Roelfsema, Lamme, & Spekreijse, 1998; Roelfsema, Lamme, Spekreijse, & Bosch, 2002) at a latency of about 200 ms, nicely coincident with the latency of integration in parietal and prefrontal areas and of the engagement of the P3 process. These two responses have very different functional dependencies: The first transient response is largely determined by stimulus properties and can be explained by classical bottom-up receptive field properties. On the contrary, the second response can be modulated by different cognitive and contextual factors. For instance, it is amplified if the local stimulus is salient and attended (as in figure-ground experiments) and it can virtually disappear if the stimulus is masked (which presumably precludes its access to consciousness) or by anesthetics. This specificity suggests an engagement of the central workspace system in the second wave of activity and that the same neuron may be involved in distinct processing stages within the same task. Further experiments are required to determine whether this second wave of activity shows a dual-task delay characteristic of the serial processing bottleneck.

Altogether, neurophysiological and brain-imaging studies suggest that, beyond an initial and relatively reliable perceptual delay of about 200 ms, a decision stage begins that involves a process of stochastic accumulation of evidence and the joint activation of a distributed network of areas, with partially changing topography as a function of the nature of the task, but with frequent if not systematic co-activation of parietal and premotor regions. Our results suggest that this accumulation system is slow (which might be required to flexibly combine information in different sensory, mnemonic, and motor modalities), is variable (probably due to an active, dynamic reverberating and integrating state), and is responsible for establishing the PRP bottleneck. Thus, while certain neural computations (probably mediated by hard-wired circuits confined to small portions of sensory space) can be very fast, precise, and parallel, linking such information together in a coherent workspace results in a slow, variable, and intrinsically sequential computation.

The challenge for the next few years will be to understand which precise biophysical mechanisms are involved in this coherent workspace and why they result in such dynamic and architectonically distinctive properties. At present, our results simply suggest that this bottleneck might occur because the cerebral accumulation system is broadly distributed and largely shared across tasks, and thus must be entirely "mobilized", at any given moment, by whichever task is currently performed. Considerable research has examined the neurophysiology of a single cognitive operation, but much less is known on how we chain these basic operations into complex tasks. The results described here constitute a first step toward a progressive understanding of the chaining of simple computations into complex compound tasks.

REFERENCES

- Allport, D., Styles, E., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421–452). Cambridge, MA: MIT Press.
- Arnell, K. M., & Duncan, J. (2002). Separate and shared sources of dual-task cost in stimulus identification and response selection. *Cognitive Psychology, 44*, 105–147.
- Arnell, K. M., Helion, A. M., Hurdelbrink, J. A., & Pasiaka, B. (2004). Dissociating sources of dual-task interference using human electrophysiology. *Psychonomic Bulletin & Review, 11*, 77–83.
- Ashbridge, E., Walsh, V., & Cowey, A. (1997). Temporal aspects of visual search studied by transcranial magnetic stimulation. *Neuropsychologia, 35*, 1121–1131.
- Baars, B. J. (1989). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.
- Bellgowan, P. S., Saad, Z. S., & Bandettini, P. A. (2003). Understanding neural system dynamics through task modulation and measurement of functional MRI amplitude, latency, and width. *Proceedings of the National Academy of Sciences USA, 100*, 1415–1419.
- Bichot, N. P., Schall, J. D., & Thompson, K. G. (1996). Visual feature selectivity in frontal eye fields induced by experience in mature macaques. *Nature, 381*, 697–699.
- Brisson, B., & Jolicoeur, P. (2007a). Cross-modal multitasking processing deficits prior to the central bottleneck revealed by event-related potentials. *Neuropsychologia, 45*, 3038–3053.
- Brisson, B., & Jolicoeur, P. (2007b). A psychological refractory period in access to visual short-term memory and the deployment of visual-spatial attention: Multitasking processing deficits revealed by event-related potentials. *Psychophysiology, 44*, 323–333.
- Brisson, B., Robitaille, N., & Jolicoeur, P. (2007). Stimulus intensity affects the latency but not the amplitude of the N2pc. *Neuroreport, 18*, 1627–1630.
- Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience, 13*, 87–100.
- Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience, 12*, 4745–4765.
- Bryant, H. L., & Segundo, J. P. (1976). Spike initiation by transmembrane current: A white-noise analysis. *Journal of Physiology, 260*, 279–314.
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523–547.
- Butts, D. A., Weng, C., Jin, J., Yeh, C. I., Lesica, N. A., Alonso, J. M. et al. (2007). Temporal precision in the neural code and the timescales of natural vision. *Nature, 449*, 92–95.
- Cecchi, G. A., Sigman, M., Alonso, J. M., Martinez, L., & Chialvo, D. (2000). Noise in neurons is message dependent. *Proceedings of the National Academy of Sciences USA, 97*, 5557–5561.
- 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*, 109–127.
- Corallo, G., Sackur, J., Dehaene, S., & Sigman, M. (2008). Limits on introspection: Distorted subjective time during the dual-task bottleneck. *Psychological Science, 19*, 1110–1117.
- De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 965–980.
- De Jong, R. (1995). The role of preparation in overlapping-task performance. *Quarterly Journal of Experimental Psychology: A, 48*, 2–25.
- de Ruyter van Steveninck, R. R., Lewen, G. D., Strong, S. P., Koberle, R., & Bialek, W. (1997). Reproducibility and variability in neural spike train. *Science, 275*, 1805–1808.
- Dehaene, S. (1996). The organization of brain activation in number comparison: Event-related potential and the additive-factor methods. *Journal of Cognitive Neuroscience, 8*, 47–68.
- 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 USA, 95*, 14529–14534.
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition, 79*, 1–37.
- Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology, 5*, e260.
- Dell'acqua, R., Jolicoeur, P., Vespignani, F., & Toffanin, P. (2005). Central processing overlap modulates P3 latency. *Experimental Brain Research, 165*, 54–68.
- Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129*, 481–507.
- Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23*, 475–483.
- Dux, P. E., Ivanoff, J., Asplund, C. L., & Marois, R. (2006). Isolation of a central bottleneck of information processing with time-resolved fMRI. *Neuron, 52*, 1109–1120.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1*, 1–47.
- Formisano, E., & Goebel, R. (2003). Tracking cognitive processes with functional MRI mental chronometry. *Current Opinion in Neurobiology, 13*, 174–181.
- Gegenfurtner, K. R., & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 845–866.
- Gilbert, C. D., & Sigman, M. (2007). Brain states: Top-down influences in sensory processing. *Neuron, 54*, 677–696.
- Gold, J. I., & Shadlen, M. N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature, 404*, 390–394.
- Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences, 5*, 10–16.

- Graziano, M., & Sigman, M. (2008). The dynamics of sensory buffers: Geometric, spatial and experience dependent shaping of iconic memory. *Journal of Vision*, 8, 1–13.
- 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, 859–868.
- Gur, M., Beylin, A., & Snodderly, D. M. (1997). Response variability of neurons in primary visual cortex (V1) of alert monkeys. *Journal of Neuroscience*, 17, 2914–2920.
- Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science*, 274, 427–430.
- Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, 56, 208–216.
- Hernandez, A., Zainos, A., & Romo, R. (2002). Temporal evolution of a decision-making process in medial premotor cortex. *Neuron*, 33, 959–972.
- Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology*, 148, 574–591.
- Hubel, D. H., & Wiesel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 28, 229–289.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 148, 574–591.
- Jack, A. I., & Roepstorff, A. (2003). Trusting the subject I. *Journal of Consciousness Studies*, 9, 9–10.
- Jack, A. I., & Roepstorff, A. (2004). Trusting the subject II. *Journal of Consciousness Studies*, 11, 7–8.
- Jentsch, I., Leuthold, H., & Ulrich, R. (2007). Decomposing sources of response slowing in the PRP paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 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, 390–396.
- Jolicoeur, P. (1999). Concurrent response-selection demands modulate the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1097–1113.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Kamienkowski, J. E., & Sigman, M. (2008). Delays without mistakes: Response time and error distributions in dual-task. *PLoS ONE*, 3, e-3196.
- Kara, P., Reinagel, P., & Reid, R. C. (2000). Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron*, 27, 635–646.
- Kim, J. N., & Shadlen, M. N. (1999). Neural correlates of a decision in the dorso-lateral prefrontal cortex of the macaque. *Nature Neuroscience*, 2, 176–185.
- Kim, S. G., Richter, W., & Ugurbil, K. (1997). Limitations of temporal resolution in functional MRI. *Magnetic Resonance in Medicine*, 37, 631–636.
- Koechlin, E., & Jubault, T. (2006). Broca's area and the hierarchical organization of human behavior. *Neuron*, 50, 963–974.
- Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, 302, 1181–1185.
- Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, 23, 571–579.
- Lamme, V. A., Super, H., Landman, R., Roelfsema, P. R., & Spekreijse, H. (2000). The role of primary visual cortex (V1) in visual awareness. *Vision Research*, 40, 1507–1521.
- Lee, T. S., Yang, C. F., Romero, R. D., & Mumford, D. (2002). Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. *Nature Neuroscience*, 5, 589–597.
- Li, W., Piech, V., & Gilbert, C. D. (2006). Contour saliency in primary visual cortex. *Neuron*, 50, 951–962.
- Li, W., Piech, V., & Gilbert, C. D. (2008). Learning to link visual contours. *Neuron*, 57, 442–451.
- 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, 122–144.
- 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, 396–409.
- Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, 40, 77–111.
- Logan, G. D., & Gordon, R. (2001). Executive control of visual attention in dual tasks. *Psychological Review*, 108, 393–434.
- Luce, R. D. (1986). *Response times*. New York: Oxford University Press.
- Luck, S. J. (1998). Sources of dual-task interference: Evidence from human electrophysiology. *Psychological Science*, 9, 223–227.
- Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2 and V4 of macaque visual cortex. *Journal of Neurophysiology*, 77, 24–42.
- Lutz, A., Lachaux, J. P., Martinerie, J., & Varela, F. J. (2002). Guiding the study of brain dynamics by using first-person data: Synchrony patterns correlate with ongoing conscious states during a simple visual task. *Proceedings of the National Academy of Sciences USA*, 99, 1586–1591.
- Mainen, Z. F., & Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science*, 268, 1503–1506.
- Marois, R., Chun, M. M., & Gore, J. C. (2000). Neural correlates of the attentional blink. *Neuron*, 28, 299–308.
- Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences*, 9, 296–305.
- Mazurek, M. E., Roitman, J. D., Ditterich, J., & Shadlen, M. N. (2003). A role for neural integrators in perceptual decision making. *Cerebral Cortex*, 13, 1257–1269.
- Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology*, 41, 211–253.
- Menon, R. S., Luknowsky, D. C., & Gati, J. S. (1998). Mental chronometry using

- latency-resolved functional MRI. *Proceedings of the National Academy of Sciences USA*, 95, 10902–10907.
- Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part I. Basic mechanisms. *Psychological Review*, 104, 3–65.
- Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequalities. *Nature*, 215, 1519–1520.
- Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist*, 51, 102–116.
- Osman, A., & Moore, C. M. (1993). The locus of dual-task interference: Psychological refractory effects on movement-related brain potentials. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 1292–1312.
- Overgaard, M. (2006). Introspection in science. *Consciousness and Cognition*, 15, 629–633.
- Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 358–377.
- Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, 116, 220–244.
- Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology*, 41A, 19–45.
- Pashler, H., & Johnston, J. C. (1998). Attentional limitations in dual-task performance. In H. Pashler (Ed.), *Attention* (pp. 155–189). Hove: Psychology Press/Erlbaum.
- Pinel, P., Dehaene, S., Riviere, D., & LeBihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *Neuroimage*, 14, 1013–1026.
- Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Sciences USA*, 91, 7398–7403.
- Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology*, 109, 160–174.
- Ratcliff, R. (1988). Continuous versus discrete information processing modeling accumulation of partial information. *Psychological Review*, 95, 238–255.
- Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience*, 19, 1736–1753.
- Roelfsema, P. R., Lamme, V. A. F., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, 395, 376–381.
- Roelfsema, P. R., Lamme, V. A., Spekreijse, H., & Bosch, H. (2002). Figure-ground segregation in a recurrent network architecture. *Journal of Cognitive Neuroscience*, 14, 525–537.
- Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience*, 22, 9475–9489.
- Romo, R., Hernandez, A., & Zainos, A. (2004). Neuronal correlates of a perceptual decision in ventral premotor cortex. *Neuron*, 41, 165–173.
- Romo, R., & Salinas, E. (1999). Sensing and deciding in the somatosensory system. *Current Opinion in Neurobiology*, 9, 487–493.
- Ruthruff, E., Pashler, H. E., & Klaassen, A. (2001). Processing bottlenecks in dual-task performance: Structural limitation or strategic postponement? *Psychonomic Bulletin and Review*, 8, 73–80.
- Schall, J. D. (2000). From sensory evidence to a motor command. *Current Biology*, 10, R404–406.
- Schall, J. D., & Bichot, N. P. (1998). Neural correlates of visual and motor decision processes. *Current Opinion in Neurobiology*, 8, 211–217.
- Schwarz, W. (2001). The ex-Wald distribution as a descriptive model of response times. *Behavior Research Methods, Instruments & Computers*, 33, 457–469.
- Schweickert, R. (1980). Critical-path scheduling of mental processes in a dual task. *Science*, 209, 704–706.
- Schweickert, R., & Townsend, J. T. (1989). Trichotomy: Interactions of factors prolonging sequential and concurrent mental processes in stochastic discrete mental (PERT) networks. *Journal of Mathematical Psychology*, 33, 328–347.
- Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, 8, 1391–1400.
- Sessa, P., Luria, R., Verleger, R., & Dell'Acqua, R. (2007). P3 latency shifts in the attentional blink: Further evidence for second target processing postponement. *Brain Research*, 1137, 131–139.
- Shadlen, M. N., & Newsome, W. T. (1996). Motion perception: Seeing and deciding. *Proceedings of the National Academy of Sciences USA*, 93, 628–633.
- Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*, 86, 1916–1936.
- Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701–703.
- Sigman, M., & Dehaene, S. (2005). Parsing a cognitive task: A characterization of the mind's bottleneck. *PLoS Biology*, 3, e37.
- Sigman, M., & Dehaene, S. (2006). Dynamics of the central bottleneck: Dual-task and task uncertainty. *PLoS Biology*, 4, e220.
- Sigman, M., & Dehaene, S. (2008). Brain mechanisms of serial and parallel processing during dual-task performance. *Journal of Neuroscience*, 28, 7585–7598.
- Sigman, M., Jobert, A., LeBihan, D., & Dehaene, S. (2007). Parsing a sequence of brain activations at psychological times using fMRI. *Neuroimage*, 35, 655–668.
- Smith, M. C. (1967). Theories of the psychological refractory period. *Psychological Bulletin*, 67, 202–213.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74, 1–29.
- Sternberg, S. (1969). The discovery of processing stages: Extension of Donders' method. *Acta Psychologica*, 30, 276–315.
- Sternberg, S. (2004). Separate modifiability and the search for processing modules.

- In N. Kanwisher & J. Duncan (Eds.), *Attention and performance XX* (pp. 125–142). Oxford: Oxford University Press.
- Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology*, *14*, 1–36.
- Thompson, K. G., Bichot, N. P., & Schall, J. D. (1997). Dissociation of visual discrimination from saccade programming in macaque frontal eye field. *Journal of Neurophysiology*, *77*, 1046–1050.
- Thompson, K. G., Hanes, D. P., Bichot, N. P., & Schall, J. D. (1996). Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. *Journal of Neurophysiology*, *76*, 4040–4055.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the visual system. *Nature*, *381*, 520–522.
- Tononi, G., & Edelman, G. M. (1998). Consciousness and complexity. *Science*, *282*, 1846–1851.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*, 550–592.
- Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information processing in the primate visual system: An integrated systems perspective. *Science*, *255*, 419–423.
- Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *4*, 631–650.
- Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, *13*, 37–58.
- Wagenmakers, E., & Scott, B. (2007). On the linear relation between the mean and the standard deviation of a response time distribution. *Psychological Review*, *114*, 830–841.
- Wong, K. F. E. (2002). The relationship between attentional blink and psychological refractory period. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 54–71.
- Wundt, W. M. (1897/1999). *Outlines of psychology*. Bristol: Thoemmes Press.

CHAPTER TWO

Frames, brains, and content domains: Neural and behavioral effects of descriptive context on preferential choice

David R. Mandel and Oshin Vartanian

INTRODUCTION

At 7:51 am on Friday, January 12, 2007, a man wearing jeans, a long-sleeve T-shirt, and a Washington Nationals baseball cap began to play his violin at the L'Enfant Plaza metro station in Washington DC. Over the next 43 minutes, 1097 people passed him by. Among them, only seven stopped to listen for at least a minute. Twenty-seven gave him money, most without breaking their pace, for a grand total of \$32 and change. Only one person, who gave the man \$20—more than half of what he earned—realized that the “fiddler” was Joshua Bell, one of the world’s most celebrated musicians, who had just played six timeless pieces of music on a violin handcrafted by Stradivari in 1713 and worth an estimated US\$3.5 million. Two days earlier, Bell had performed at a theater in Boston where merely “pretty good” seats sold for \$100. This study, organized by *The Washington Post* (Weingarten, 2007), poignantly illustrates the importance of context on subjective valuation. As Weingarten put it, “He [Bell] was, in short, art without a frame.” The Bell demonstration, of course, was not designed to carefully disentangle the possible causal determinants of people’s ostensible indifference toward beauty in a mundane environment, but rather to conjure in our minds the idea that, in two disparate contexts, the *same* man playing the *same* music on the *same* exquisite instrument could be valued and treated so differently.

The argument that context matters hence presupposes that there are