Current Biology 22, 1-5, October 9, 2012 ©2012 Elsevier Ltd All rights reserved http://dx.doi.org/10.1016/j.cub.2012.07.043

### Report

# Decision Making during the Psychological Refractory Period

Ariel Zylberberg, 1,2,3,6 Brian Ouellette, 3,6 Mariano Sigman, 1 and Pieter R. Roelfsema 3,4,5,\*

<sup>1</sup>Laboratory of Integrative Neuroscience, Physics Department, FCEyN UBA and IFIBA, Conicet, Pabellón 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina <sup>2</sup>Instituto de Ingeniería Biomédica, Facultad de Ingeniería, Universidad de Buenos Aires, 1063 Buenos Aires, Argentina <sup>3</sup>Department of Vision and Cognition, Netherlands Institute for Neuroscience, an Institute of the Royal Netherlands Academy of Arts and Sciences, Meibergdreef 47, 1105 BA Amsterdam, The Netherlands

<sup>4</sup>Department of Integrative Neurophysiology, Center for Neurogenomics and Cognitive Research, VU University, 1081 HV Amsterdam, The Netherlands <sup>5</sup>Psychiatry Department, Academic Medical Center, P.O. Box 75867, 1070 AW Amsterdam, The Netherlands

### Summary

In spite of its massively parallel architecture [1], the human brain is fundamentally limited if required to perform two tasks at the same time [2, 3]. This limitation can be studied with the psychological refractory period (PRP) paradigm, where two stimuli that require speeded responses occur in close succession [4]. Interference generally takes the form of a delay in the time to respond to the second stimulus [5]. Previous studies suggested that sensory decisions require the accumulation of sensory evidence [6, 7] and that the PRP reflects the inability to form more than one decision at a time [4, 8]. In the present study, we used a psychophysical reverse-correlation technique [9, 10] to measure the timecourse of evidence accumulation during the PRP. We found that the accumulation of evidence could occur during the PRP albeit with a reduced efficiency, which implies that multiple decision processes can occur in parallel in the human brain. In addition to the reduced efficiency of evidence accumulation, our results uncover an additional delay in the routing of the decision to motor structures during the PRP, which implies that the process of sensory decision making is separable from the preparation of a motor response [11–13].

### Results

There are many parallel pathways that connect the sensory and motor areas of the human brain. Yet, human behavior is remarkably serial. If multiple sensory decisions have to be made in close succession, earlier decisions postpone the later ones, but the cause of this seriality is not well understood. A direct approach to study the interference between successive tasks is with the psychological refractory period (PRP) paradigm, in which two stimuli that require independent responses are presented in close succession [2–4, 14]. The ubiquitous finding is that the response time to the second stimulus

increases markedly if the time between stimuli is short ( $\sim$ 400 ms or less), as if the human brain is simply unable to make two decisions at the same time. Behavioral studies attributed the bottleneck to a limitation in response-selection [4], which is usually modeled as a process that accumulates information over time [15–19] and commits to a choice when the evidence (the "decision variable") reaches a threshold. Support for accumulation models has come from investigations in behaving monkeys, which identified neurons in association areas of the parietal and frontal cortex with activity that closely matches the dynamics of these decision variables [6, 7, 20–22].

Previous studies [8, 23] suggested that dual-task interference might reflect a reduced efficiency of evidence accumulation. One possibility is that sensory evidence can only accumulate for one task at a time [8]. Another possibility is that evidence accumulation proceeds in parallel for multiple decisions [23], but only one of them can be routed to a motor program at any one time [24]. This latter model implicitly distinguishes between two stages in response selection that have often been lumped together in previous work [6, 25]: the integration of sensory evidence and the subsequent routing of the decision to the motor response.

To measure how dual-task interference affects the evidence accumulation, we designed a PRP experiment in which a first task (T1) interfered with a second task (T2) that required the integration of noisy sensory evidence. We investigated how fluctuations in the sensory signal influenced the subject's decision with a "classification image" technique [9, 10, 26]. Specifically, the method allowed us to measure the influence of sensory evidence presented at different time points on the decision for T2. We examined three alternative models (sketched in Figure 1A): (1) integration of evidence for T2 is not possible during the PRP; (2) integration occurs normally during the PRP; and (3) integration occurs but is less efficient-reflecting some form of capacity sharing [27]. In addition, we determined whether the PRP bottleneck also causes a delay in the routing of the decision to the motor response, once evidence accumulation is complete.

Three human participants performed two tasks as fast as possible. The first "tone task" (T1) was an auditory discrimination task. Participants decided whether the frequency of a single pure tone was high (880 Hz) or low (440 Hz) (Figures 1B and 1C). Task 2 was a visual task that required a luminance comparison: participants decided which of two flickering patches was, on average, brighter. We adjusted the difficulty of the luminance comparison so that subjects were ~75% correct (see Supplemental Experimental Procedures available online) and presented it at a variable stimulus-onset asynchrony (SOA) of 0, 120, 200, 520, or 600 ms. We instructed the participants to respond to the tone task before the visual task and collected 10,000 trials per subject so that we could examine how well the three models of Figure 1 fitted the data for every individual subject.

### **Evidence Integration during the PRP**

We first analyzed the effect of SOA on response times (RTs). As in previous studies [4], we observed a strong effect of SOA on

<sup>&</sup>lt;sup>6</sup>These authors contributed equally to this work

<sup>\*</sup>Correspondence: p.roelfsema@nin.knaw.nl

**Current Biology** Vol 22 No 19

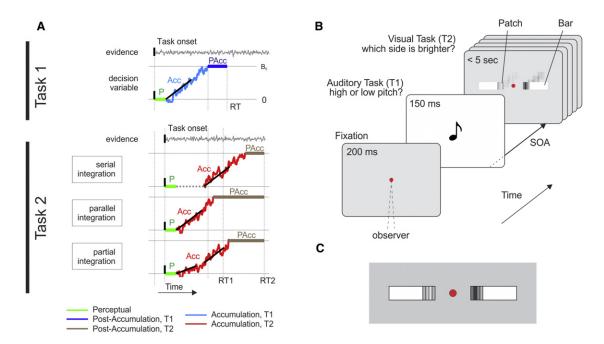


Figure 1. Models of Evidence Accumulation during the PRP and Design of the Experiment to Distinguish between Them

(A) A decision is made by the accumulation of sensory evidence. Noisy sensory evidence is shown in gray, and the accumulated evidence (Acc) for T1 and T2 is shown in light blue and red, respectively. For clarity, only a single barrier is shown, but two symmetrical bounds at  $\pm B_0$  are generally used to model two-alternative forced-choice decisions. The accumulation process ends when it reaches one of these barriers. Additional perceptual (P) and postaccumulation (PAcc) latencies add to the total response time. Three alternative models are shown for the second task of the PRP paradigm, which differ in how the accumulation of evidence for T2 is influenced by T1. If only one accumulation process can proceed at a time, then the integration for T2 is delayed at a short SOA ("serial integration") [8]. Alternatively, if the accumulation of evidence for the two tasks can proceed in parallel, no influence should be observed on the accumulation process ("parallel integration"). Dual-task interference could also be caused by a reduction in the efficacy of accumulation for T2 ("partial integration"). Note that all models can result in the same response-time to T2 (RT2) by adjusting PAcc.

(B) Experimental design. Subjects performed a tone discrimination task (T1) followed by a luminance discrimination task (T2), which involved deciding which of two patches was brightest. Each patch consisted of four bars and independent luminance noise was added to each bar and updated at a frequency of 25 Hz. The SOAs were randomly selected on each trial and could take values of [0, 120, 200, 520, 600] ms. Responses were made as fast as possible, using different hands for the two tasks.

(C) Magnification of one frame of the visual stimulus. The target patch with highest luminance is on the left.

RT in the second task (RT2,  $p < 10^{-8}$ , see Supplemental Experimental Procedures), which was on average 45% longer when the SOA was 0 ms than when it was 600 ms (Figure 2A). Accuracy in the visual task was only slightly influenced by SOA. It decreased from 77% at an SOA of 0 ms to 74% at an SOA of 600 ms ( $p < 10^{-8}$ , see Supplemental Experimental Procedures). Errors in the tone task were below 2% for every SOA. The RTs in the tone task (RT1) were slightly influenced by SOA, with a mean of 447 ms at an SOA of 0 ms and 426 ms at an SOA of 600 ms (Figure 2A). The analyses indicate a typical PRP effect: RT2 decreased with a slope close to -1 as a function of SOA inside the interference range and was unaffected by SOA beyond this range (Figure 2A).

The use of time-varying stimuli for T2 allowed us to go beyond the analysis of RTs and measure how the PRP influences the integration of sensory evidence. If the PRP interrupts the evidence integration, luminance samples during this epoch do not influence the subject's choice. Hence, the average of the luminance fluctuations during the PRP on correct and incorrect trials should be equal and the classification image should therefore be zero.

We measured classification images for each participant and illustrate the average classification image across participants for every SOA in Figure 2B (see Supplemental Experimental Procedures). Early sensory information had a strong influence on the decision (Figures 2B and 2C). The classification images

peaked around the third sample (around 100 ms) and gradually fell back to zero [28]. It is likely that the weak influence of the later samples is caused by the absence of evidence integration once the subject has made a decision about which luminance patch was brighter. Models that posit that the PRP is caused by the interruption of evidence accumulation predict that dual-task interference should be manifested by a delay in the onset of the integration, which should be particularly pronounced at the shorter SOAs [8]. This is not what we observed. The early samples always contributed to the decision, irrespective of SOA (Figures 2B and 2C). Yet, there was an influence of the PRP on the classification images because they were protracted in time for the shorter SOAs. Accordingly, the classification image at 0 ms SOA differed significantly from that at an SOA of 600 ms, outside the PRP (Monte-Carlo simulation,  $p < 10^{-8}$ , see Supplemental Experimental Procedures). These results, taken together, indicate that evidence integration occurs during the PRP, although it is less efficient.

When we aligned the classification images to RT2, we found that integration stopped at least 200 ms before the response [29]. The delay between the peak of the classification image and the response was larger for the shorter SOAs, which implies that the PRP also causes a delay once evidence accumulation is complete. Figure 2 shows the average classification image, but the effect of the PRP on evidence accumulation and on the delay between the decision and the motor response

Decision Making during the PRP

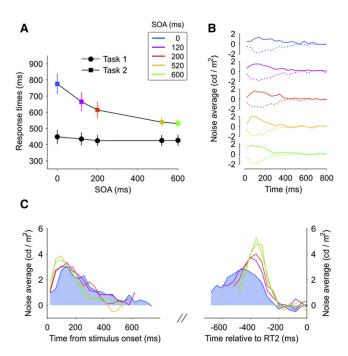


Figure 2. Response Times and Psychophysical Kernels

(A) Average response times for the tone (T1) and visual (T2) tasks. Different SOAs are indicated with different colors (see legend). Error bars indicate SEM.

(B) Classification images, Each panel corresponds to a different SOA, Two classification images are shown in each panel: the solid line shows the difference in the luminance fluctuations between correct and incorrect trials of the "target" patch. On correct trials, the luminance of the target patch tended to be higher. The dashed line shows the same subtraction for the "distractor" patch, which tended to have a lower luminance on correct trials. (C) Time course of the classification images ("target"-"distractor"), after aligning the luminance fluctuations to the onset of the visual stimulus (left) or to the time of the response (right). Noise averages in the left portion of the graph are drawn until one of the three participants had responded in more than 70% of trials at that SOA. Samples after the subject's response were excluded from the trial averages. On the right, the fluctuations were aligned to the time of the response before computing the classification images. Noise averages in this portion of the graph are drawn from the times at which every participant has started integration in at least 30% of trials. The classification images of individual participants are shown in Figure S1A. We obtained similar psychophysical kernels with a linear regression analysis (Figure S1B).

of task 2 was significant in each of the three participants (Figure S1A).

We also investigated the influence of luminance samples on RT with a regression analysis (Figure S1B). The results were highly consistent with the classification image analysis, because luminance fluctuations during the PRP did influence RT but their effect was weaker than outside the PRP.

## Disentangling the Contributions to the PRP during the Decision Process

We found that dual-task interference reduces the efficiency of evidence accumulation but that it is not entirely halted. To estimate the relative contribution of delays in evidence accumulation and in the mapping of the sensory decision onto a behavioral response to the PRP, we fitted a model (Supplemental Experimental Procedures) that accumulates evidence until a decision variable reaches a bound for one or the other decision [15–19].

The model simulated the integration of luminance in a decision variable and includes the three models of Figure 1B for different values of the model parameters (see Supplemental Experimental Procedures; parameter values are shown in Table S1). The model accounted for the complex influence of dual-task interference on the accumulation of evidence and provided an excellent fit to the distribution of response times across SOAs with an average R2 of 0.98 (Figure 3A). The classification images had not been used for model fitting. We could therefore test whether the model based its response on the same luminance samples as the subjects, by calculating the classification images of the model and comparing them to the experimental ones. The agreement between the classification images of the model and participants was excellent (the average R<sup>2</sup> across SOA was 0.89) (Figure 3B; Figure S2), which indicates that the model used the same information as the subjects did.

The excellent fit implies that we can use the model to separate the PRP delay into a contribution caused by the decreased efficiency of evidence integration and another contribution caused by a delay in routing the decision to the motor response (Figure 3C). The PRP prolonged the accumulation of evidence by an average of 86 ms (51 ± 7 ms, 148 ± 12 ms, and 60 ± 10 ms in the three participants), accounting for an average of  $\sim$ 35% of the increase in RT at zero SOA. The gain of the evidence accumulation was reduced to 0.75, 0.41, and 0.74 for the three participants, which indicates that substantial evidence accumulation occurred during the PRP. In addition, the PRP increased the subsequent routing time by an average of 159 ms (139  $\pm$  8 ms, 177  $\pm$  14 ms, and 161 ± 16 ms in the three participants), accounting for an average of ~65% of the increase in RT2 at zero SOA. These results, taken together, indicate that the PRP-bottleneck caused a reduction in the efficiency of evidence accumulation and a longer delay in routing this decision to a motor response.

In spite of the reduced efficacy, evidence integration was not halted by the PRP. We considered the possibility that the residual integration might have occurred at low levels of the visual system, before the PRP-bottleneck. At first sight, such an explanation may appear unlikely, for two reasons. First, we presented luminance samples at the same spatial location, and later samples therefore masked the previous ones. Second, the task demanded the comparison of the luminances at two locations, which presumably goes beyond low levels of the visual system. Nevertheless, we investigated the contribution of prebottleneck levels to our luminance integration task in a second experiment (Supplemental Experimental Procedures). Six subjects mentally traced a curve through a series of three bifurcations. They had to choose the branch with highest luminance at each bifurcation while keeping gaze at a central fixation (Figure 4A), making three motor responses to indicate the decisions made at each bifurcation. Evidence accumulation was highly serial (Figure 4B); the kernel for the second decision was delayed relative to the first one, and in turn the kernel for the third decision was delayed relative to the second one (p  $< 10^{-8}$  for both comparisons, permutation test). These results rule out that low levels of the visual system integrated the luminance information in the main PRP task.

### Discussion

It is generally assumed that dual-task interference during the PRP reflects the serial nature of the decisional processes [2, 4] and the inability to select two responses at the same **Current Biology** Vol 22 No 19

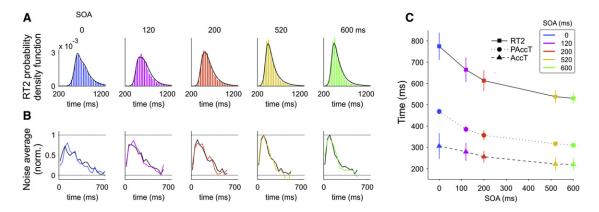


Figure 3. Fit of a Diffusion Model of Dual-Task Interference

- (A) Response time distributions as a function of SOA, for the participants (colored bars) and the model (black lines). Response times are shown relative to the onset of the visual stimulus of task 2. The fraction of variance explained by the model fit (R<sup>2</sup>) is, from left to right, 0.99, 0.98, 0.99, 0.98, and 0.98. Bin width is 40 ms.
- (B) Classification images for the data (colored lines) and model (black lines), as a function of SOA, normalized to the peak value across SOA. The fraction of variance explained by the model fit (R<sup>2</sup>) is, from left to right, 0.78, 0.92, 0.91, 0.93, and 0.9. In (A and B), the model was fitted independently for each participant, and the resulting fits were averaged across participants.
- (C) The model separated RT2 (solid lines) into a time of accumulation (dashed lines, AccT; time at which the boundary is reached) and a postaccumulation time (dotted lines, PAccT). Error bars indicate SEM across participants. Response time distributions and classification images for individual participants are shown in Figure S2, and model parameters are shown in Table S1.

time [4]. Here, we determined how the PRP influences the integration of sensory evidence and the mapping of the decision onto a motor response. Our results allow us to draw two main conclusions. First, the integration of sensory evidence does not constitute an absolute bottleneck because evidence accumulation for a second task can occur during processing of the first task, albeit with a reduced efficiency. The control experiment demonstrated that early, prebottleneck levels are not responsible for the residual evidence accumulation during the PRP. Second, the reduction in the efficiency of evidence accumulation explains only a fraction of the PRP. We observed that the PRP also caused a long delay after evidence integration was complete, which was directly visible as a large increase in the delay between the classification image and RT2 (Figure 2C, right).

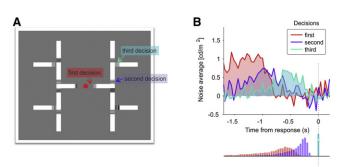


Figure 4. Serial Integration of Evidence for Multiple Decisions

- (A) Subjects had to trace a curve by choosing the brightest luminance directions at three locations. As in our main experiment, the luminance at each bifurcation varied rapidly over time (25 Hz) so that we could measure evidence accumulation for the three decisions.
- (B) The classification images, aligned to the time of the third response. Evidence for the three decisions accumulated serially, implying that low levels of the visual system do not automatically integrate luminance information for this task. The bottom panel shows the distribution of response times for the three decisions, aligned to the last response.

There is an ongoing debate about the characteristics of the bottleneck, with some theories suggesting that it operates serially and can only be made available for one task at a time [4, 8] and other theories positing that it is parallel because it reflects a limited resource that can be shared by tasks [27, 30]. Our results support capacity sharing theories by showing that the PRP does not delay the onset of evidence accumulation but that it does reduce its efficacy. By using a classification image technique, the present study identified the accumulation of evidence and the routing of the decision to motor areas as two of the processes that depend on this "central" resource.

The present study determined how task 1 interferes with task 2 and it thereby complements studies that investigated "backwards" influences of task 2 on task 1 during the PRP [31–33]. For instance, Hommel [31] demonstrated that the first response in a dual-task design is faster when it is compatible with the second, supporting parallel processing for both tasks. The present results support capacity sharing models in a complementary manner by showing that evidence accumulation for task 2 can take place during the PRP, albeit with reduced efficiency.

The present results also have implications for theories on decision making within a single task. Previous studies suggested that decision making and action selection are inseparable [25, 34]. According to this idea, decisions are taken by accumulating evidence in those motor structures that determine the motor response. If a decision is communicated through an eye movement, then an area like LIP, involved in eye movement planning, would accumulate the evidence, whereas another area—like the parietal reach region —would accumulate evidence if the task requires an arm movement. However, a recent neurophysiological study demonstrated that the integration of sensory evidence in area LIP also occurs if the evidence is not directly linked to a movement plan [13] and that neuronal signals for evidence integration and eye movement planning follow different time courses. Our results support such a separation between sensory decisions and

Please cite this article in press as: Zylberberg et al., Decision Making during the Psychological Refractory Period, Current Biology (2012), http://dx.doi.org/10.1016/j.cub.2012.07.043

**Decision Making during the PRP** 

5

the resulting action. A large fraction of the PRP was caused by a delay in the processes that occurred after the accumulation of evidence but before the initiation of the response.

We have previously suggested a framework for sequential decision making, where complex tasks are decomposed into a sequence of simpler operations that require the integration of information from the senses or from a previous processing stage [11]. To explain the present results, such a model would first accumulate sensory evidence to determine the brightest patch, and the sensory decision would be followed by a second integration process that "routes" it to the appropriate motor response [24]. The use of the classification image method in a PRP paradigm has provided new insights into the cause of the processing bottleneck, showing that the reduced efficacy of accumulation of sensory evidence accounts for a fraction of the PRP effect, whereas a substantial delay occurs in the routing of the decision to motor structures.

### **Supplemental Information**

Supplemental Information includes three figures, one table, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2012.07.043.

#### Acknowledgments

This research was supported by the Human Frontiers Science Program. A.Z. was supported by a grant from the Peruilh Foundation, Faculty of Engineering, Buenos Aires University. B.O. was supported by Human Frontier Science Program Grant RGP0007/2007-C and P.R.R. by a NWO-VICI grant.

Received: March 28, 2012 Revised: June 11, 2012 Accepted: July 19, 2012

Published online: August 23, 2012

### References

- Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1, 1–47.
- Welford, A. (1952). The 'psychological refractory period' and the timing of high speed performance - a review and a theory. Br. J. Psychol. 43, 2–19.
- Telford, C.W. (1931). The refractory phase of voluntary and associative responses. J. Exp. Psychol. 14, 1–36.
- Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. Psychol. Bull. 116, 220–244.
- Pashler, H. (1984). Processing stages in overlapping tasks: evidence for a central bottleneck. J. Exp. Psychol. Hum. Percept. Perform. 10, 358-377
- Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. 30, 535–574.
- Roitman, J.D., and Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J. Neurosci. 22, 9475–9489.
- Sigman, M., and Dehaene, S. (2005). Parsing a cognitive task: a characterization of the mind's bottleneck. PLoS Biol. 3, e37.
- Eckstein, M.P., and Ahumada, A.J., Jr. (2002). Classification images: a tool to analyze visual strategies. J. Vis. 2, 1x.
- Neri, P., and Heeger, D.J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. Nat. Neurosci. 5, 812–816.
- Zylberberg, A., Dehaene, S., Roelfsema, P.R., and Sigman, M. (2011).
  The human Turing machine: a neural framework for mental programs.
  Trends Cogn. Sci. 15, 293–300.
- Freedman, D.J., and Assad, J.A. (2011). A proposed common neural mechanism for categorization and perceptual decisions. Nat. Neurosci. 14, 143–146.

- Bennur, S., and Gold, J.I. (2011). Distinct representations of a perceptual decision and the associated oculomotor plan in the monkey lateral intraparietal area. J. Neurosci. 31, 913–921.
- Smith, M.C. (1967). Theories of the psychological refractory period. Psychol. Bull. 67, 202–213.
- Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. Ergonomics 13, 37–58.
- McClelland, J.L. (1979). On the time relations of mental processes: an examination of systems of processes in cascade. Psychol. Rev. 86, 287.
- 17. Luce, R.D. (1986). Response Times (New York: Oxford University Press).
- Gold, J.I., and Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. Trends Cogn. Sci. 5, 10–16.
- Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. 20, 873–922.
- Shadlen, M.N., and Newsome, W.T. (1996). Motion perception: seeing and deciding. Proc. Natl. Acad. Sci. USA 93, 628–633.
- Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J. Neurophysiol. 86, 1916–1936.
- Schall, J.D. (2001). Neural basis of deciding, choosing and acting. Nat. Rev. Neurosci. 2, 33–42.
- Logan, G.D., and Gordon, R.D. (2001). Executive control of visual attention in dual-task situations. Psychol. Rev. 108, 393–434.
- Zylberberg, A., Fernández Slezak, D., Roelfsema, P.R., Dehaene, S., and Sigman, M. (2010). The brain's router: a cortical network model of serial processing in the primate brain. PLoS Comput. Biol. 6, e1000765.
- Shadlen, M.N., Kiani, R., Hanks, T.D., and Churchland, A.K. (2008).
  Neurobiology of decision making: an intentional framework. In Better Than Conscious?: Decision Making, the Human Mind, and Implications for Institutions, C. Engel and W. Singer, eds. (Cambridge, MA: The MIT Press), pp. 71–101.
- Ahumada, A.J., Jr. (1996). Perceptual classification images from Vernier acuity masked by noise. Perception 25, 1831–1840.
- Tombu, M., and Jolicoeur, P. (2003). A central capacity sharing model of dual-task performance. J. Exp. Psychol. Hum. Percept. Perform. 29. 3–18.
- Ludwig, C.J.H., Gilchrist, I.D., McSorley, E., and Baddeley, R.J. (2005).
  The temporal impulse response underlying saccadic decisions.
  J. Neurosci. 25, 9907–9912.
- Resulaj, A., Kiani, R., Wolpert, D.M., and Shadlen, M.N. (2009). Changes of mind in decision-making. Nature 461, 263–266.
- Kahneman, D. (1973). Attention and Effort (Englewood Cliffs, NJ: Prentice-Hall, Inc).
- Hommel, B. (1998). Automatic stimulus-response translation in dualtask performance. J. Exp. Psychol. Hum. Percept. Perform. 24, 1368– 1384.
- Miller, J., and Alderton, M. (2006). Backward response-level crosstalk in the psychological refractory period paradigm. J. Exp. Psychol. Hum. Percept. Perform. 32, 149–165.
- Watter, S., and Logan, G.D. (2006). Parallel response selection in dualtask situations. Percept. Psychophys. 68, 254–277.
- Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. Philos. Trans. R. Soc. Lond. B Biol. Sci. 362, 1585–1599.