Modification of response time variability in a decision-making task

Gijsbert Stoet[a,c], Hannes Ruge[b,d] and Lawrence H. Snyder[a]

Departments of [a]Anatomy and Neurobiology, [b]Psychology, Washington University School of Medicine, St Louis, Missouri, USA, [c]Institute of Psychological Sciences, University of Leeds, Leeds, UK and [d]Technical University of Dresden, Department of Psychology, Dresden, Germany

Correspondence to Gijsbert Stoet, Institute of Psychological Sciences, Faculty of Medicine and Health, University of Leeds, Leeds, UK
Tel: + 44 II3 343 8579; fax: + 44 II3 343 5749; e-mail: g.stoet@leeds.ac.uk

Experimental work was carried out at Washington University in St Louis and the University of Leeds, UK.

Previous single-unit recordings in monkeys showed that essential information regarding a decision is available earlier to posterior parietal cortex than expected based on simultaneously measured behavioral response times (RTs). We propose that this observation reflects a tendency of the brain to reduce RT variability in repetitive response sequences. To test this, we studied the effects of experimentally introduced variability in trial duration on RTs in humans. We found that humans adapt timing based on the timing of surrounding trials, essentially reducing RT variability in trial sequences. This implies that RTs do not always reflect the minimal time it takes to make a decision. Implications for the interpretation of behavioral data are discussed. *NeuroReport* 19:1321–1324 © 2008 Wolters Kluwer Health | Lippincott Williams & Wilkins.

Keywords: human, parietal lobe, primates, reaction time, rhesus monkey

Introduction

Previously, we have studied the relationship between response time (RT) and the latency of neural activity in posterior parietal cortex (PPC) [1,2]. Rhesus monkeys responded significantly faster in easy compared with difficult conditions in a task-switching paradigm. We found that in task-selective neurons of the PPC, the neural latency was not only shorter for easy than for difficult stimuli, but surprisingly, the neural latency difference between easy and difficult trials was substantially larger than the behavioral (RT) difference; monkeys performed more slowly in easy trials than one would expect, based on PPC activity. There is evidence for a similar dissociation of neuronal and behavioral timing in humans [3].

Here, we ask why subjects do not respond still faster to easy stimuli, given that their parietal cortices encode the correct response so quickly. This question is relevant for understanding how the PPC contributes to behavioral output, and generally for understanding the relationship between neural processes and behavior.

A simple explanation for the paradoxical effect in monkeys is that primates prefer to perform repetitive responses to stimuli with a certain rhythm, and accordingly, try to regularize their RT. In favor of this explanation is the well-established finding that primates perform many repetitive tasks in a rhythm [4,5].

Here, we tested whether humans would adapt RTs to reduce variability in the timing of decision-making tasks. In three experiments, volunteers performed tasks under conditions of varying difficulty and intertrial timing (i.e. sequences had irregular timing). We predicted that performance in short-lasting (or easy) trials would be slowed down when long-lasting (or difficult) trials were interleaved. Similarly, we predicted that performance in trials preceded by short intertrial intervals (ITI) would be slowed down when trials preceded by long ITIs were interleaved. Slowing down in easy and short trials would reduce variability as compared with a situation in which participants simply respond as quickly as possible in every trial.

Materials and methods

Participants

Fifty-two humans (median age: 22 years) volunteered in three experiments (*n*=8, 20, and 24, respectively).

Apparatus

Experiments were controlled by standard PCs running custom software. Stimuli were presented on a 14″ flat-screen monitor (experiments 1 and 2) or on a 17″ cathode-ray monitor (experiment 3).

Procedures

Participants sat behind the screen and PC keyboard, with approximately 60 cm between eyes and screen. They used their left and right index fingers for pressing the left or right shift keys of the keyboard. Participants performed sequences of trials, each consisting of stimulus presentation that required an immediate left or right button press. Stimuli were visible until a button was pressed.

Experiments 1 and 2 had an identical design, but entirely different sets of stimuli to ensure that effects found could not be attributed to particularities of the sensory features or the stimulus–response rules. In both experiments, there was a set of easy and difficult stimuli. In experiment 1, there were four easy stimuli (5×5 cm) whose colors were associated with either a left or right button press (i.e. red or orange=left, dark-green or bright-green=right). The four difficult stimuli were rectangles, each with a superimposed triangle; the combination of the color of the rectangle with the triangle orientation determined the response (i.e. left=blue rectangle + yellow triangle; left=yellow rectangle + blue inverted triangle; right=yellow rectangle + blue triangle; right-blue rectangle + yellow inverted triangle).

In experiment 2, participants determined whether there was an odd or even number of blue circles (1 cm in diameter) in a yellow rectangle (5×4 cm), and had to indicate their decision using the left or right button. Easy stimuli contained two or three circles and difficult stimuli five or six circles (circles randomly positioned).

In experiment 3, participants determined the parity of the numbers 1 through 8 (about 3×3 cm, white on black background), and responded with one button to odd and another button to even numbers. Instead of varying difficulty levels to manipulate trial duration, we varied the ITIs, which were either 300, 600, or 900 ms (while ITI was held constant in experiments 1 and 2 at 200 ms).

An important experimental factor in all experiments was 'pure' versus 'mixed' block design. In 'pure' blocks, only one difficulty level (experiments 1 and 2) or one ITI (experiment 3) was used. In 'mixed' blocks, all difficulty levels (experiments 1 and 2) or all ITIs (experiment 3) were unpredictably interleaved.

Apart from the pure and mixed conditions, experiments 1 and 2 were performed under two different speed–accuracy conditions. We wanted accurate performance because this would match the accurate performance of monkeys, and it would lead to less interrupted response sequences. As a control for the high-accuracy condition, we also included a speed-prioritized condition.

In one half of each experiment, participants had to focus on accuracy; we asked the participants to perform at least 200 trials consecutively without making any errors. The block continued until the participant performed 200 consecutive successful trials. In this condition, we allowed a maximum RT of 6 s. In the other half of the experiment, participants were asked to work fast, which was encouraged by a maximum RT of 1.2 s in experiment 1, and 2.5 s in experiment 2. In both experiments 1 and 2, there were three blocks containing 100 trials each of just one (pure) condition, and an additional three blocks containing 200 trials each of mixed conditions. In experiment 3, pure and mixed blocks contained 200 and 600 trials, respectively. Block order was counterbalanced across participants in all experiments.

Results

In experiments 1 and 2, participants responded to simple and difficult trials in pure and mixed blocks. Average RTs and error rates were calculated for each participant and repeated measures analyses of variance (ANOVAs) were performed on these data; ANOVA factors are named 'pure/mix', 'easy/difficult', 'speed/accuracy', and 'ITI'. Only trials that followed at least two correct trials were analyzed.

Participants responded more slowly in difficult than in easy trials (as determined by two separate ANOVAs demonstrating the same pattern of a main effect of the factor easy/difficult [experiment 1: $F(1,7)=71.8$, experiment 2: $F(1,19)=180.8$, both $P<0.01$]). This effect of difficulty was different in pure and mixed blocks, which was expressed as a significant interaction between the factors pure/mix and easy/difficult [experiment 1: $F(1,7)=22.7$, $P<0.01$; experiment 2: $F(1,19)=4.7$, $P<0.05$]. In experiment 1, the RT difference between easy and difficult trials in pure blocks was 387 ± 39 ms (mean ± 1 SEM, $n=8$ participants); in mixed blocks this difference was reduced to only 225 ± 41 ms. Similarly, in experiment 2, the RT difference was reduced from 606 ± 41 ms (mean ± 1 SEM, $n=17$ participants) in pure blocks down to 531 ± 50 ms in mixed blocks (Fig. 1). Thus, the difference between easy and difficult trials was much larger when trials were performed separately than when mixed. We found no significant difference between difficult trials in the pure and the mixed blocks. Difficult trials had similar RTs in pure and mixed blocks (*t*-tests), and therefore the difference between blocks was due primarily of a lengthening of easy trial RTs in the mixed blocks (easy trials being faster in the pure than in the mixed blocks, [$t(7)=5.1$, $t(19)=6.3$, $P<0.01$ in both the experiments]).

To address the question of whether speed–accuracy tradeoffs may have affected performance, we reanalyzed the same data, but now with an additional factor for the speed-prioritized and accuracy-prioritized conditions. Participants responded more slowly in the accuracy-prioritized condition {on average 269 ± 38 ms and 243 ± 40 ms more slowly in experiments 1 [$F(1,7)=35.4$, $P<0.01$] and 2 [$F(1,19)=52.6$, $P<0.01$], respectively}. Error rates were lower in the accuracy-prioritized conditions (1.3 and 1.0% in experiments 1 and 2, respectively) than in the speed-prioritized conditions (8.4 and 4.4%, respectively) [experiment 1: $F(1,7)=30.0$, $P<0.01$, experiment 2: $F(1,19)=19.4$, $P<0.01$]. Importantly, the analyses of the accuracy-prioritized and speed-prioritized blocks separately show the same pattern.

The observed speed reduction in the easy trials in mixed blocks means that overall RT variability in these blocks was reduced compared with a situation in which people would have responded in easy trials as fast as they actually did in pure easy blocks. To quantify and statistically test the meaningfulness of this reduction in variability, we compared the actual variability in mixed blocks with the variability that we expected based on the data from the pure blocks. We pooled the data from pure easy and pure difficult trials together and then calculated the coefficient of variability (CV) as the standard deviation divided by the mean (this value is what we would 'expect' to observe in mixed blocks, assuming that the mixing of trials itself would not change anything about the way people respond). Figure 2 shows these expected values plotted against the CV values measured in the mixed blocks. Most of the data point fall below the unity line (six of eight in experiment 1, and 16 out of 21 in experiment 2), indicating that variability was reduced in most cases. On average, the CV in experiment 1 was 25 percentage points less in the mixed blocks than expected (based on the pure blocks [$t(7)=2.4$, $P<0.05$]). In experiment 2, this reduction was 10 percentage points [$t(19)=2.8$, $P<0.05$].

An alternative explanation for reduced variability in RT is that the inclusion of difficult trials might cause participants

to apply a conservative strategy of delaying their responses on all trials, including easy trials. This explanation is unlikely because reduced variability is seen even under accuracy-prioritized conditions, when participants are presumably already using a conservative strategy. To further rule out this explanation, we asked whether variability was systematically reduced when trial duration was manipulated by varying ITIs rather than by varying difficulty. We hypothesized that if the reduced variability in experiments 1 and 2 were truly because of a preference for rhythm, participants would show a similar effect when delays between trials were varied randomly. This is what was found. In experiment 3, there was a significant interaction between the factors pure/mix and ITI [$F_{(2,46)}=9.6$, $P<0.01$]: when ITIs were mixed, trials preceded by an ITI of 300 ms became 21 ± 7 ms more slowly [$t(23)=3.2$, $P<0.01$], whereas trials preceded by 600 and 900 ms ITIs were no different under mixed compared with pure conditions (Fig. 1).

Next, we consider switch costs in the mixed blocks of all three experiments. Switch costs are a performance cost associated with trials in which a participant switches from applying one rule to another [6]. Switch costs can be asymmetrical: the cost of switching from a hard to an easy trial can be greater than the cost of switching from an easy trial to a hard trial [7,8]. Asymmetrical switch costs could therefore contribute to the phenomenon we have observed: an asymmetrical slowing down of easy trials within blocks of mixed easy and hard trials. To exclude this possibility, we excluded all switch trials from the aforementioned analyses of experiments 1 and 2. This exclusion had no effect on our results. In fact, the results we report are the values with switch trials excluded. [In a separate analysis, we found significant switch costs in experiments 1 and 2 (298 ± 57 and 158 ± 21 ms, respectively) but not in experiment 3.]

We next consider changes in RT at a much finer scale, by looking at the distribution of differences in RTs on consecutive trials (CRTs). The rationale behind this analysis is that if people would strategically adapt their responses to reduce variability, changes in RTs between CRTs should be smaller than expected by chance. For example, for one
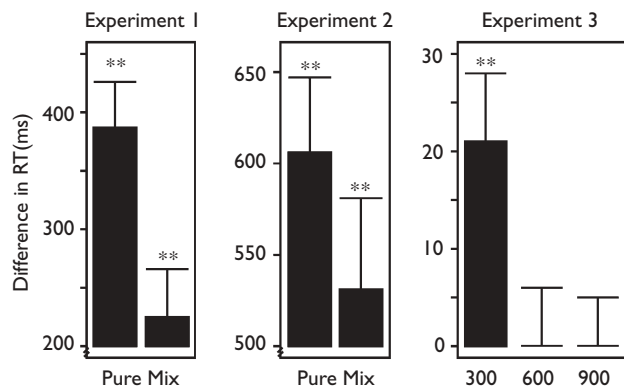

Fig. 2 Estimated and observed coefficients of variability. For each participant, the coefficients of variability [standard deviation (SD)/mean response time (RT)] are plotted for the mixed blocks in experiments 1 (top) and 2 (bottom). The expected values are based on the combined RTs of the pure blocks. The observed data are based on trials from the mixed blocks. The observed data are, averaged across participants, 25 and 10% points faster than the expected values ($P<0.05$), in experiments 1 and 2, respectively.

participant, we plotted the actual distribution of CRTs (black curve in Fig. 3a) and then compared the actual distribution of CRTs with the average CRT distribution of the 10 000 randomly shuffled sequences of the same RTs (gray curve in Fig. 3a). The data of this example show how observed CRTs are closer together (i.e. narrow distribution in black) than expected by chance (i.e. wider distribution shown in gray).

We calculated the differences in consecutive RTs for a block of trials, and used a permutation test to determine whether CRTs were more similar to one another than would be expected by chance. For each participant and for each block of N trials, we first obtained the average absolute CRT of the ($N-1$) consecutive RT differences. Next, we shuffled the N trials 10 000 times, and after each shuffle we recomputed the average change in RT differences.

Comparisons of the difference scores of all three experiments indicated that the differences were significant at the population level [experiment 1: $t(7)=3.2$, $P<0.01$; experiment 2: $t(19)=8.1$, $P<0.01$; experiment 3: $t(23)=8.8$, $P<0.01$] (Fig. 3b). The results indicate that, in mixed

Fig. 1 Effects of difficulty in pure and mixed blocks. For experiments 1 and 2, the difference between easy and difficult trials in pure and mixed conditions is displayed (mean difference + 1 SEM, *$P<0.01$). In both pure and mixed conditions, participants were significantly faster in easy than in difficult trials, but the advantage of easy trials was significantly reduced in the mixed conditions. In experiment 3, differences between pure and mixed trials preceded by either a 300, 600, or 900 ms ITI are displayed. In the 300 ms condition, participants performed significantly faster in the pure than in the mixed condition.
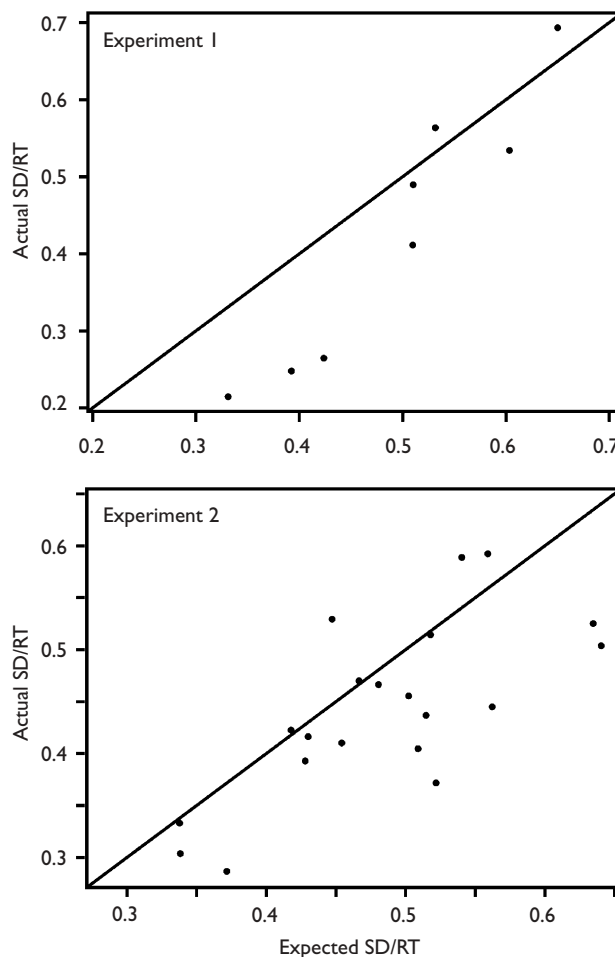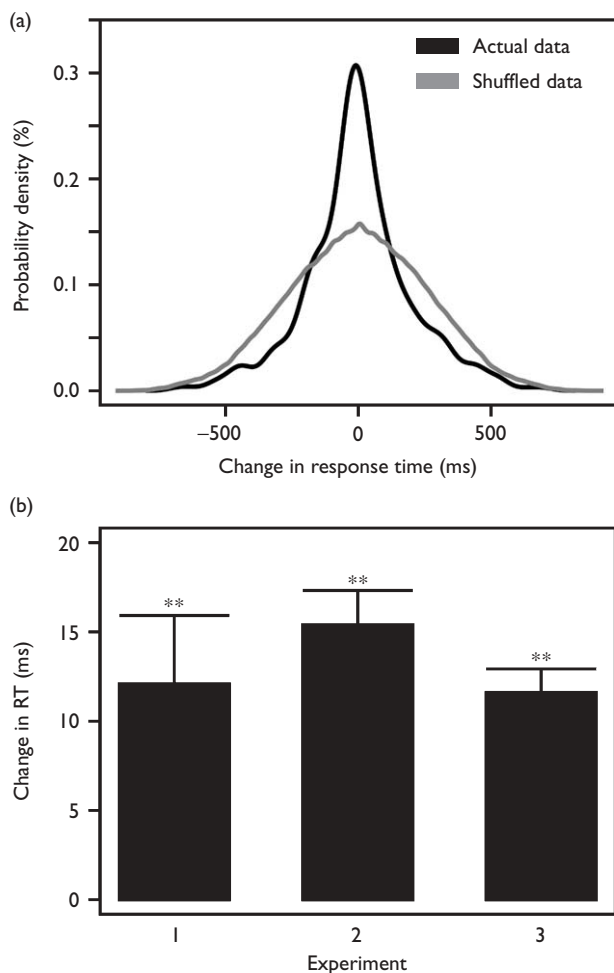
Fig. 3 Reduced variability in response time (RT) changes between individual trials. (a) Change in RT between trials is calculated by subtracting consecutive RTs. Shown is an example distribution (using a Gaussian smoothing kernel for display purposes only) of change in RTs of one participant (experiment I, pure block of 443 easy trials in the accuracy-prioritized condition) in black (SD=200 ms). The gray curve shows the distribution after those same trials have been randomly shuffled 10 times and the results after each shuffle averaged together. The narrower distribution (black) indicates the reduction of variability on a trial-by-trial basis. (b) Difference between observed and expected change in standard deviation of change in RT for each of the three experiments (mean ± 1 SEM, **$P < 0.01$).

compared to pure blocks, participants not only reacted more slowly on easy trials (experiments 1 and 2) and on short trials (experiment 3), but also showed reduced changes in RT on a trial-by-trial basis.

Discussion

We found that participants reduced their RT variability by slowing down on easy and fast trials. We found this effect not only in average RTs of easy (or fast) and difficult (or slow) trials, but also on a finer trial scale. The data of this study corroborate our hypothesis that humans adapt their responses to reduce variability, and that this effect is mostly because of lengthening their RTs on short and easy trials.

Our finding of an adaptation of response timing to reduce variability in RTs offers an explanation for the paradoxical observation in the monkey [2] described in the introduction. It is not entirely surprising that the brain prefers to reduce variability. After all, behavior with low variability (e.g. rhythmic behavior) is ubiquitous in the animal kingdom (e.g. walking, scratching, breathing, swimming, chewing, and drumming [4]). Traditionally, rhythmic and discrete goal-directed behaviors have been studied as separate phenomena, but more recently the relationship between the two has been addressed [9–11].

Our current data and previous data [2] show that RTs do not always reflect the minimum processing time of stimulus information. And yet, this is a common assumption underlying models of behavior in the cognitive sciences. Our research raises the awareness of this problem; we do not solve it, but we hope our findings stimulate research into solutions for this important problem.

Conclusion

Monkeys do not always respond as fast as they could, given that their PPC encodes the correct response earlier than behavior indicates. We hypothesized that the brain prefers to reduce variability at the cost of maximum speed. We confirmed this in three behavioral experiments in humans. RTs cannot always be taken as an indication of the time to make a decision in the most efficient way, but may reflect adaptive strategies as well, such as a preference for regular timing over speed.

References

1. Stoet G, Snyder LH. Single neurons in posterior parietal cortex (PPC) of monkeys encode cognitive set. *Neuron* 2004; **42**:1003–1012.
2. Stoet G, Snyder LH. Correlates of stimulus-response congruence in posterior parietal cortex (PPC). *J Cogn Neurosci* 2007; **19**:194–203.
3. Ruge H, Stoet G, Naumann E. Attentional set mixing: effects on target selection and selective response activation. *Psychophysiology* 2006; **43**:413–421.
4. Delcomyn F. Neural basis of rhythmic behavior in animals. *Science* 1980; **210**:492–498.
5. Repp B. Sensorimotor synchronization: a review of the tapping literature. *Psychon Bull Rev* 2005; **12**:969–992.
6. Monsell S. Task switching. *Trends Cogn Sci* 2003; **7**:134–140.
7. Allport DA, Styles EA, Hsieh S. Shifting intentional set: exploring the dynamic control of tasks. In: Umilta C, Moscovitch M, editors. *Attention and Performance XV*. Cambridge: MIT Press; 1994. pp. 421–452.
8. Monsell S, Yeung N, Azuma R. Reconfiguration of task-set: is it easier to switch to the weaker task? *Psychol Res* 2000; **63**:250–264.
9. Grosjean M, Rosenbaum DA, Elsinger C. Timing and reaction time. *J Exp Psychol: General* 2001; **130**:256–272.
10. MacDonald CJ, Meck WH. System-level integration of interval timing and reaction time. *Neurosci Biobehav Rev* 2004; **28**:747–769.
11. Schaal S, Sternad D, Osu R, Kawato M. Rhythmic arm movement is not discrete. *Nat Neurosci* 2004; **7**:1137–1144.