Part III

Psychophysics

Chapter 7

Measuring the Cost of Deploying Top-down Visual Attention

7.1 Introduction

Imagine walking into a crowded restaurant, looking for your friend whom you are supposed to meet. You will be looking around, scanning the faces of the patrons for your friend's without paying much attention to the interior design or the furniture. Entering the same restaurant with the intention of finding a suitable table, on the other hand, will have you looking at pretty much the same scene, and yet your perception will be biased for the arrangement of the furniture, mostly ignoring the other guests.

Task or agenda affect our visual perception by a set of processes that we commonly term top-down attention, to distinguish them from stimulus-driven bottom-up attention (Treisman and Gelade 1980; Itti and Koch 2001a). It enables us to preferentially perceive what is important for the task at hand. Without top-down attention to the relevant parts of a scene, we may even miss large changes (Rensink et al. 1997; Simons and Levin 1998; Simons and Rensink 2005) until we are explicitly cued, for instance by pointing (exogenous cue) or by describing the changed part of the image (endogenous cue). If cueing changes our visual perception so effectively, what is the cost of deploying attention to a new task?

Humans can detect object categories in natural scenes in as little as 150 ms (Thorpe et al. 1996; Potter and Levy 1969), and Li et al. (2002) demonstrated that this can be achieved even when spatial attention is tied to a demanding task elsewhere in the visual field. At this fast processing speed, there is not enough time for feedback within the visual hierarchy, suggesting purely feed-forward, bottom-up processing. Because of their block design, these experiments allow subjects to prepare for the given task well in advance, giving them ample time to bias their visual system accordingly. Here we are interested in the reaction time cost for adjusting the visual system for a new task from trial to trial. Thus, we ask how efficient it is to bias the visual system from the top down to allow

for subsequent efficient processing of stimuli in a purely bottom-up fashion.

Wolfe et al. (2004) approached a similar question for visual search by cueing odd-one-out search tasks in sets of 6–18 items, finding a reaction time cost of up to 200 ms for picture cues and 700 ms for word cues for the mixed versus blocked condition. These endogenous cues take about 200 ms to become fully effective. However, with their design Wolfe et al. (2004) were not able to separate the cost for deploying top-down attention from the cost for other processes such as perceiving and interpreting the cue.

We address this question by adapting the task switching paradigm, recently reviewed by Monsell (2003), to fast natural scene categorization tasks (Walther et al. 2006). Task switching was introduced by Jersild (1927), who had students work through lists of simple computation tasks (adding and subtracting 3 from numbers). He found that blocks, in which the two tasks alternate, require considerably more time than blocks with single tasks. This result was later verified by Spector and Biederman (1976).

In general, task switching experiments require subjects to perform two or more tasks that typically relate to different attributes of the stimulus, e.g., reporting whether a number is odd or even vs. whether it is larger or smaller than 5. Subjects are tested in blocks with single tasks and in mixed task blocks. In mixed blocks the tasks can either alternate in a prespecified sequence, e.g. "AABBAABB" (Allport et al. 1994; Rogers and Monsell 1995; De Jong 2000), or the task order can be unpredictable, and a task cue is presented before stimulus onset (Sudevan and Taylor 1987; Meiran 1996). See Koch (2005) for a comparison of the two paradigms.

In either case, there will be trials with task repeats and trials with task switches. Reaction times (RTs) tend to be longer for switch trials than for repeat trials. The difference is termed "switch cost." Even though no actual switch needs to happen in repeat trials, RTs will still be longer than in single task blocks, giving rise to a "mixing cost." Both switch and mixing cost depend on the preparation time from the presentation of the cue or, in the absence of a cue, from the end of the previous trial to the stimulus onset of the current trial. It is frequently observed that even with long preparation times of up to 5 seconds, there is still a considerable residual cost (Sohn et al. 2000; Kimberg et al. 2000).

Switch cost is generally assumed to be due to task-set reconfiguration, including a shift of attention between stimulus attributes, selection of the correct response action, and, depending on the task, reconfiguration of other task-specific cognitive processes. Mixing cost captures the extra effort involved in *potentially* (but not actually) having to switch to another task compared to a single task condition, such as time for cue perception and interpretation.

When attempting to determine the cost of shifting attention, switch cost is the more interesting effect. However, cost for attention shifts is confounded with other costs such as response selection in its contribution to switch cost. To disentangle these effects, we propose a paradigm with four

tasks divided into two task groups of two tasks each. Tasks within groups relate to the same stimulus attribute and hence do not require an attention shift, while switching between tasks from different groups requires shifting attention to a different stimulus attribute. Since the only difference in within-group versus between-group switches is the necessity to shift attention, the difference in switch cost between the two conditions will give us a measure for its cost.

This project is a collaboration with Dr. Fei-Fei Li. She and I initiated the project and supervised SURF student Lisa Fukui for early pilot studies. I conducted the experiments and analyzed the data reported in this chapter.

7.2 Methods

7.2.1 Subjects

Six right-handed subjects (one female, five male) with normal or corrected to normal vision participated in the experiments, including the author. Subjects (ages 20 to 29, average 23) were recruited from the Caltech academic community and paid for their participation. All subjects passed the Ishihara screening test for color vision without error and gave written informed consent.

One subject's data were excluded from the analysis because his RT was consistently longer than two standard deviations above the RTs of all other subjects.

7.2.2 Apparatus

Stimuli were presented on a 20" Dell Trinitron CRT monitor (1024×768 pixels, 3×8 bit RGB) at a refresh rate of 120 Hz. The display was synchronized with the vertical retrace of the monitor. Stimulus presentation and recording of the subjects' response was controlled with a Pentium 4 PC running Matlab R14 with the psychophysics toolbox (Brainard 1997). Subjects were positioned approximately 100 cm from the computer screen.

7.2.3 Stimuli

Images of the natural or man-made scenes were taken from a large commercially available CD-ROM library and from the world wide web (Li et al. 2002; Thorpe et al. 1996), allowing access to several thousand stimuli. The images were converted to 256 gray levels and rescaled to subtend an area of $4.4^\circ \times 6.6^\circ$ of visual angle. Each image belonged to one of three classes containing a clearly visible animal (e.g., birds, fish, mammals, insects), clearly visible means of transport (e.g., trains, cars, airplanes, bicycles), or neither (distracter images). See figure 7.1 for examples. We used more than 1000 natural scenes of each of these classes, and each image was presented no more than twice during the test sessions.

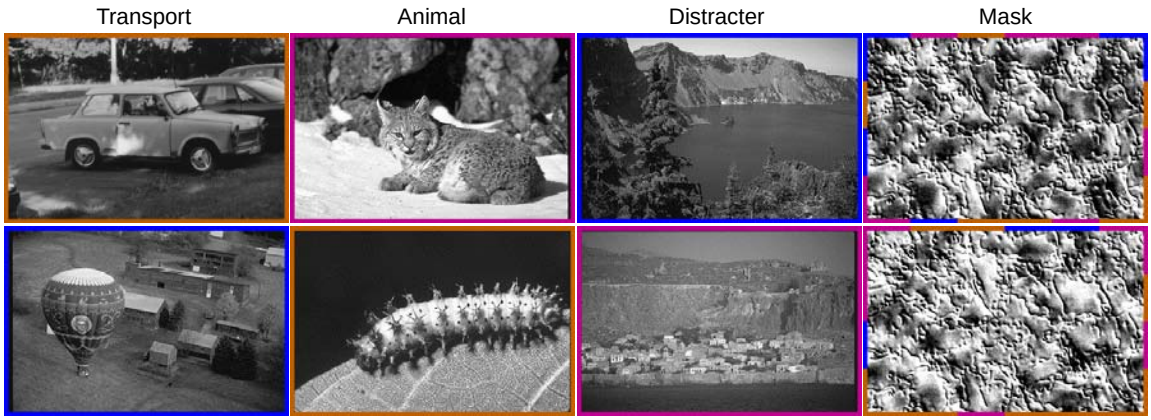

Figure 7.1: Example stimuli for means of transport, animals, and distracters, as well as example masks. Masks are created by superimposing a naturalistic texture on a mixture of white noise at different spatial frequencies (Li et al. 2002), surrounded by a frame with broken-up segments of orange, blue, and purple. Note that the thickness of the color frames is exaggerated threefold in this figure for illustration.

A 2.7' (2 pixels) thick orange, blue, or purple frame surrounds the gray-level images. At full saturation, the CIE coordinates of the colors are (0.666, 0.334), (0, 0), and (0.572, 0), respectively. The purple "distracter" color was chosen such that its hue component (0.875) in HSV space is equidistant from the hue components of the two "target" colors: orange (0.083) and blue (0.667). The brightness (value) of all three colors was adjusted for perceptual equiluminance, using a technique based on minimizing flicker between colors at 14 Hz (Wagner and Boynton 1972). During training, the saturation of the colors was decreased to make the task more difficult. The brightness was adjusted for each saturation level such that perceptual equiluminance between all three colors was maintained. Typically, saturation was decreased to about 0.15 during training, which corresponds to CIE coordinates of (0.360, 0.333) for orange, (0.315, 0.315) for blue, and (0.355, 0.303) for purple.

7.2.4 Experimental Paradigm

The design of our stimuli allows us to define two groups of two tasks each that are as unrelated as possible, while still coinciding spatially. The first group of tasks (IMG tasks) consists of detecting whether an animal is present in the image (cued by the word "Animal") or whether a means of transport is present (cued by "Transport"). This has been shown to be possible without color information (Delorme et al. 2000; Fei-Fei et al. 2005). The second task group (COL tasks) relates to the color of the frame, namely detecting an orange frame ("Orange") or a blue frame ("Blue") around the image. Stimuli are displayed at a random location with fixed eccentricity from fixation to avoid spatial attention effects.

To compare reaction times in situations with and without task switching, two kinds of blocks are used: single task blocks (48 trials) and mixed blocks (96 trials). In single task blocks the task for the

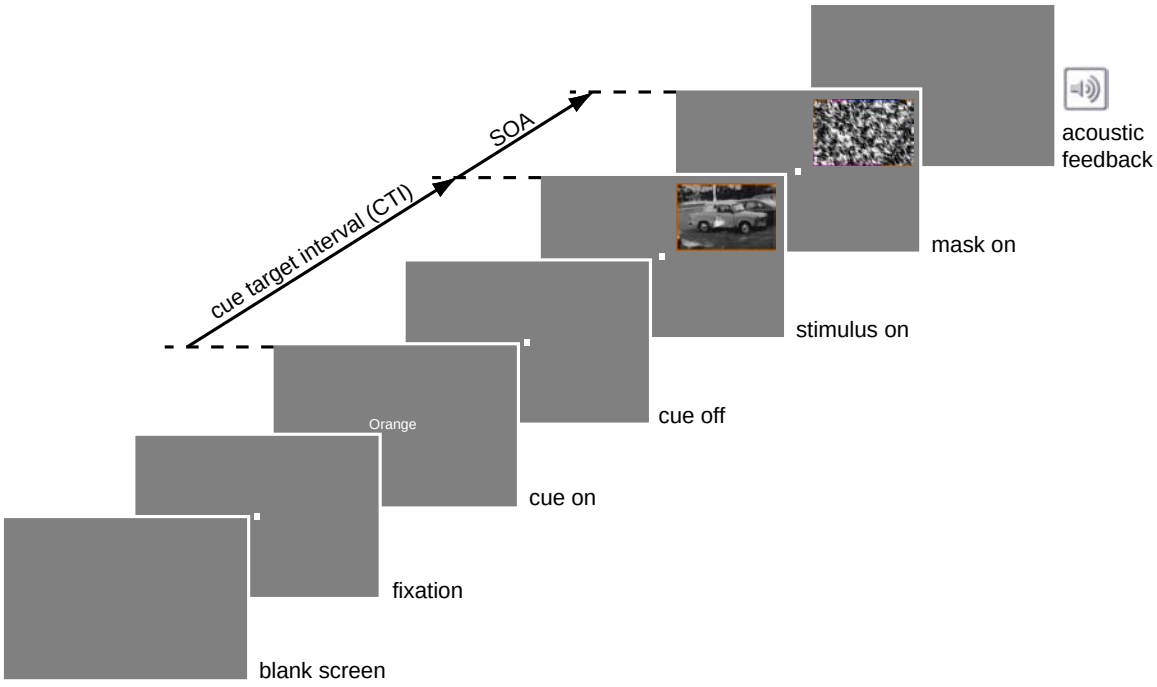

Figure 7.2: Experimental set-up. Each trial starts 1300 ms before target onset with a blank gray screen. At 650 ± 25 ms before target onset, a white fixation dot ($4.1' \times 4.1'$) is presented at the center of the display. At a variable cue target interval (CTI) before target onset, a word cue (0.5° high, between 1.1° and 2.5° wide) appears at the center of the screen for 17 ms (two frames), temporarily replacing the fixation dot for CTIs less than 650 ms. At 0 ms, the target stimulus, consisting of a gray-level photograph and a color frame around it, is presented at a random position on a circle around the fixation dot such that the image is centered around 6.4° eccentricity. After a stimulus onset asynchrony (SOA) of 200–242 ms, the target stimulus is replaced by a perceptual mask. The mask is presented for 500 ms, followed by 1000 ms of blank gray screen to allow the subjects to respond. In the case of an error, acoustic feedback is given (pure tone at 800 Hz for 100 ms), followed by 100 ms of silence. After this, the next trial commences.

entire block is indicated by an instruction screen preceding the block. For mixed blocks subjects are instructed to solve two out of the four possible tasks, either the two IMG tasks, the two COL tasks, or one task from each group. Within each mixed block, equal numbers of trials for the two tasks are shuffled randomly to make their order unpredictable. This procedure results in a statistically equal number of trials with the same task as the preceding trial (repeat) and with the other task (switch). The task for each trial is indicated by a word cue presented at the center of the screen at CTIs of 50 ms, 200 ms, and 800 ms before target onset (figure 7.2). For each block, one CTI is used throughout. For consistency, word cues are presented in both types of blocks, even though they serve no purpose in single task blocks.

On each trial, the probability of seeing a positive (i.e., as cued) example is 50 %, the probability for an example of the non-cued class is 25 %, and the probability for a distracter (non-target) example is 25 %. This is illustrated further in table 7.1. The probabilities for target frame colors

Table 7.1: Stimulus probabilities depending on task.

Task	Animal images	Transport images	Distracter images
"Animal"	50 %	25 %	25 %
"Transport"	25 %	50 %	25 %

orange and blue and the distracter color purple are distributed in an analogous manner.

Subjects are instructed to hold the left mouse button pressed with the index finger of their right hand throughout the block, and to only briefly release it as soon as they detect the cued target property for a given trial. If no response is given within 1500 ms of mask onset, a negative response is assumed (speeded go/no go response). Reaction time is measured for correct positive responses as the time passed between the onset of the target stimulus and the registration of the mouse button release event. If subjects give a positive response, i.e., release the mouse button, the 1000 ms waiting period after the mask is cut short. In case of an error, acoustic feedback is given.

Subjects were trained on single task blocks for 2–3 hours (40–60 blocks of 48 trials). For each image class, a randomly chosen subset of 80 images was set aside and re-used repeatedly for training, but not for testing. During training, the SOA was adjusted in a staircase procedure based on the performance in IMG blocks, starting with an initial 400 ms. A stable target performance between 88 % and 92 % was achieved with SOAs between 200 ms and 242 ms (average 214 ms). The same SOA was also used for COL blocks. To achieve the same level of difficulty, the saturation of the colors was decreased in a staircase procedure, starting with 1 down to between 0.098 and 0.185 (average 0.151). At the end of training, SOA and saturation were fixed for each subject.

The eight one-hour test sessions for each subject consisted of 10 mixed blocks (96 trials) interleaved with 5 single task blocks (48 trials). All positive IMG trials were done with images that the subjects had seen at most once before, thus avoiding overtraining on individual images. The order of blocks with different CTIs and task combinations was randomized within each session and counter balanced across sessions.

7.2.5 Data Analysis

Reaction time was recorded for correct positive trials. After discarding the first trial of each block, trials with a reaction time more than four standard deviations above the mean (above 995 ms) or below 200 ms were discarded as outliers (1 % of the data, see figure 7.3). Error rates and RTs were pooled separately for switch and repeat trials in mixed blocks and over all trials in single task blocks. These block results were pooled separately for each CTI value, and, for some analyses, for each task combination over all sessions for all five subjects (35 sessions in total), and the standard error of the mean (s.e.m.) was computed.

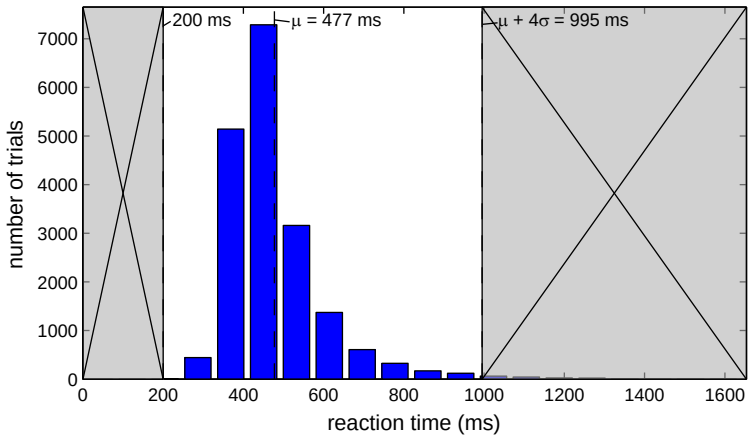

Figure 7.3: Histogram of the reaction times of all trials. Trials with reaction times below 200 ms and more than four standard deviations above the mean (above 995 ms) were discarded as outliers (1 % of the data).

Mixing and switch costs for RT are computed as

$$C^{\mathrm{RT}}_{\mathrm{mix}} = \langle RT_{\mathrm{repeat}} \rangle - \langle RT_{\mathrm{single}} \rangle \text{ and} \tag{7.1}$$

$$C^{\mathrm{RT}}_{\mathrm{switch}} = \langle RT_{\mathrm{switch}} \rangle - \langle RT_{\mathrm{repeat}} \rangle, \tag{7.2}$$

where $\langle \cdot \rangle$ denotes the mean over sessions and subjects. Their standard errors (s. e.) are derived as follows:

$$\mathrm{s.\,e.}(C^{\mathrm{RT}}_{\mathrm{mix}}) = \sqrt{\frac{\mathrm{var}(RT_{\mathrm{repeat}})}{N_{\mathrm{repeat}}} + \frac{\mathrm{var}(RT_{\mathrm{single}})}{N_{\mathrm{single}}}} \text{ and} \tag{7.3}$$

$$\mathrm{s.\,e.}(C^{\mathrm{RT}}_{\mathrm{switch}}) = \sqrt{\frac{\mathrm{var}(RT_{\mathrm{switch}})}{N_{\mathrm{switch}}} + \frac{\mathrm{var}(RT_{\mathrm{repeat}})}{N_{\mathrm{repeat}}}}. \tag{7.4}$$

Analogous formulae are used to compute the mixing ($C^{\mathrm{Err}}_{\mathrm{mix}}$) and switch costs ($C^{\mathrm{Err}}_{\mathrm{switch}}$) for error rate and their standard errors.

The significance of mixing and switch costs is determined by testing whether the two constituent RT or error rate samples are drawn from populations with different means using an unmatched t-test. Mixing and switch costs are further analyzed using N-way ANOVAs. Throughout this chapter, alpha levels of 0.05 (*), 0.01 (**), and 0.005 (***) are reported in figures and tables.

7.3 Results

Figure 7.4 shows RTs and error rates for single task blocks, repeat trials, and switch trials in mixed blocks for the three values of CTI. For single task blocks, RT is independent of CTI. RTs for task

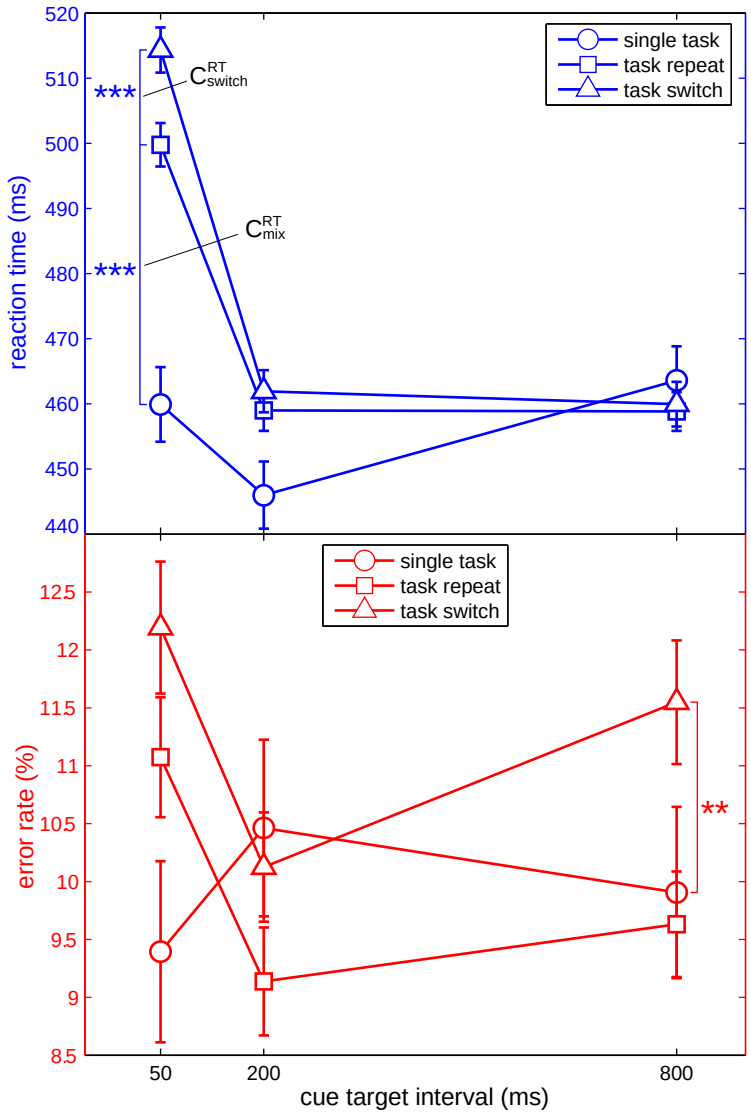

Figure 7.4: Reaction times (top, blue) and error rates (bottom, red) for single task blocks, task repeat trials, and task switch trials in mixed blocks for $n = 5$ subjects. Error bars are s.e.m. For RT, both mixing and switch cost are significant at a CTI of 50 ms, but not at CTIs of 200 ms and 800 ms ($p > 0.05$, t-test). The drop of the single task RT at 200 ms compared to 50 ms and 800 ms is not significant ($p > 0.05$, t-test). For error rate, only switch cost at a CTI of 800 ms is statistically significant. There are no other significant effects for error rate.

repeat and task switch trials are statistically the same as for single task blocks for CTI values 200 ms and 800 ms. At the shortest CTI of 50 ms, there is a significant mixing cost of 39.9 ± 6.6 ms ($p < 10^{-7}$, t-test) and a significant switch cost of 14.5 ± 4.8 ms ($p < 0.005$, t-test). There is a statistically significant switch cost of 1.9 ± 0.7 % ($p < 0.01$, t-test) in error rate at CTI = 800 ms, but no systematic effect is discernible for mixing or switch cost in error rate.

A 3-way analysis of variance (ANOVA) of mixing cost reveals significant main effects for the factors *CTI*, *task group* (COL or IMG), and *subject* for both RT and error rate (table 7.2). Table 7.2

Table 7.2: 3-way analysis of variance with interactions for mixing cost.

Source		**Mixing cost in reaction time**				**Mixing cost in error rate**			
	d.f.	*Mean square*	*F*	*p*		*Mean square*	*F*	*p*	
CTI	2	8261	67.43	$3{\cdot}10^{-13}$	***	50	10.18	$3{\cdot}10^{-4}$	***
task group	1	4490	36.65	$5{\cdot}10^{-7}$	***	56	11.41	0.002	***
subject	4	2735	22.32	$2{\cdot}10^{-9}$	***	13	2.67	0.047	*
CTI * task group	2	864	7.05	0.003	***	43	8.75	0.001	***
CTI * subject	8	350	2.86	0.014	*	17	3.44	0.005	***
task group * subject	4	1048	8.56	$5{\cdot}10^{-5}$	***	7	1.32	0.3	

Table 7.3: 4-way analysis of variance with interactions for switch cost.

Source		**Switch cost in reaction time**				**Switch cost in error rate**			
	d.f.	*Mean square*	*F*	*p*		*Mean square*	*F*	*p*	
CTI	2	915	3.79	0.03	*	6.5	1.06	0.4	
task group	1	168	0.70	0.4		0.2	0.03	0.9	
switch condition	1	1914	7.93	0.009	**	3.9	0.63	0.4	
subject	4	418	1.73	0.2		20.3	3.34	0.02	*
CTI * task group	2	221	0.92	0.4		6.7	1.10	0.3	
CTI * switch condition	2	727	3.01	0.06		5.6	0.92	0.4	
CTI * subject	8	226	0.94	0.5		10.1	1.65	0.2	
task group * switch condition	1	863	3.57	0.07		21.0	3.45	0.07	
task group * subject	4	225	0.93	0.5		8.1	1.32	0.3	
switch condition * subject	4	124	0.52	0.7		7.6	1.25	0.3	

also shows that all two-way interactions reach significance as well, with the exception of *(task group * subject)* for error rate.

As shown in figure 7.5, mixing cost in RT for CTI = 50 ms is significantly higher for IMG (52.5 ± 8.6 ms) than for COL (29.6 ± 5.3 ms) tasks ($p < 0.04$, t-test). On the other hand, mixing cost in error rate is significantly higher for COL (4.2 ± 0.6 %) than for IMG (0.4 ± 1.0 %) tasks ($p < 0.004$, t-test), suggesting a speed-accuracy trade-off.

For analyzing switch cost, we use *switch condition* (within or between task groups) as a fourth variable for the ANOVA and obtain a significant effect for it as well as for *CTI* in RT (table 7.3). The factors *task group* and *subject* are not significant for RT, while *subject* is significant for error rate. None of the two-way interactions is significant.

Note that while mixing cost is significantly affected by all factors (table 7.2), RT switch cost is only dependent on *CTI* and *switch condition* (table 7.3). In particular, the ANOVA for RT switch cost does not show a significant effect for subject identity, indicating that subject-dependent effects are absorbed by mixing cost, keeping switch cost subject independent.

While the dependence of RT switch cost on CTI is apparent from figure 7.4, figure 7.6 illustrates its dependence on the task switch condition for CTI = 50 ms. No significant switch cost was found for switching from IMG to IMG (2.8 ± 8.5 ms, $p > 0.05$, t-test) or from COL to COL (5.1 ± 7.4 ms,

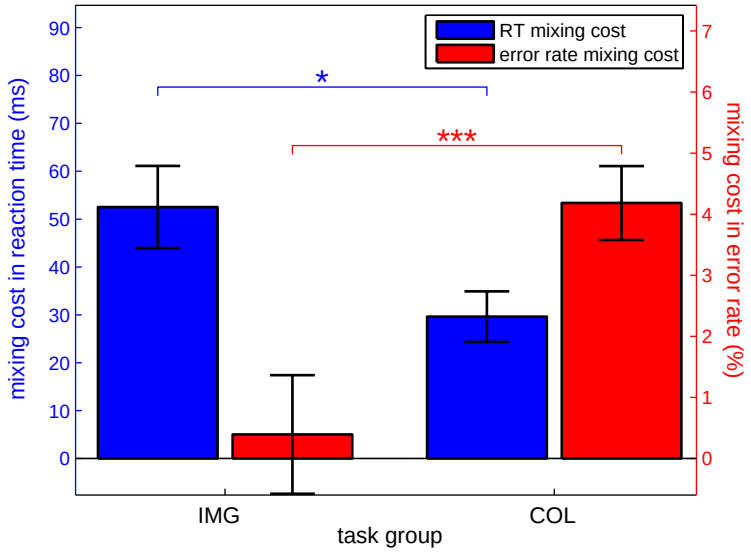

Figure 7.5: Mixing cost in RT (blue) and error rate (red) for all subjects for CTI = 50 ms, plotted by task group. While mixing cost in RT is significantly higher for IMG than for COL tasks, mixing cost in error rate is significantly higher for COL than for IMG tasks.

$p > 0.05$, t-test), but switch cost is significant for switching from COL to IMG (20.0 ± 8.6 ms, $p < 0.05$, t-test) and from IMG to COL (28.4 ± 6.9 ms, $p < 10^{-4}$, t-test).

7.4 Discussion

We set out to find the difference in switch cost of between-group and within-group task switches in order to shed light on the cost of having to shift attention from one stimulus attribute to another. We did indeed find significant RT switch costs of 20 ms (COL to IMG) and 28 ms (IMG to COL) for between-group switches, but no significant switch cost for within-group switches. The only difference between these two switch modes is that between-group switching requires the shift of attention to another stimulus attribute, while within-group switching does not. We conclude that the RT cost of having to shift attention in our fast detection paradigm is 20–28 ms.

Both mixing and switch cost in RT are significant only for a CTI of 50 ms, but not for 200 ms or longer. This agrees with the results in visual search by Wolfe et al. (2004) who found that cueing becomes fully effective within 200 ms from cue onset. This means that a CTI of 200 ms is sufficient to perceive the cue and shift attention to the cued stimulus attributes without incurring a reaction time penalty compared to perceiving the cue and not having to shift attention. Thus, 200 ms is an upper bound on the time it takes to shift attention. Presenting the cue with a CTI of 50 ms still allows subjects to perform the task (no significant switch cost in error rate), but at a penalty of 20–28 ms in RT if an attention shift is required.

What happened to the other contributors to switch cost, in particular remapping of the motor

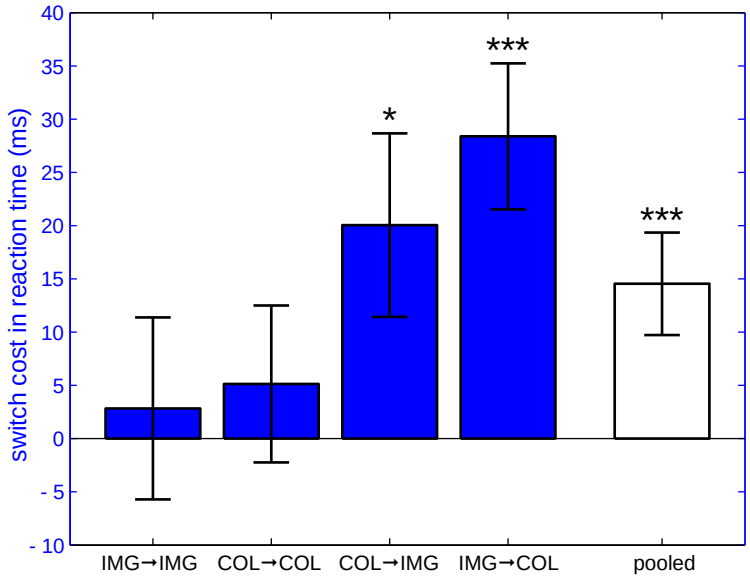

Figure 7.6: Switch cost in RT at a CTI of 50 ms for different switch conditions (blue) and pooled over all conditions (white). The white bar corresponds to the difference labeled as $C^{\mathrm{RT}}_{\mathrm{switch}}$ in figure 7.4. Error bars are standard errors as defined in eqs. 7.3 and 7.4. Switch cost is only significant when switching between the IMG and COL task groups, but not when switching within the groups.

response? The lack of a significant within-group switch cost suggests that there is no significant cost for this, which appears to contradict the results of Meiran (2000), who found significant contributions of both stimulus and response switching to the switch cost. This discrepancy may be explained by differences in the experimental design. Meiran (2000) as well as the majority of task switching studies (e.g., Rogers and Monsell 1995; Meiran 1996; De Jong 2000; Kleinsorge 2004) use a two-alternative forced choice design, typically instructing subjects to operate two keys with different hands and thus requiring coordination of the motor response across both hemispheres of the brain. Compared to these designs, our go/no go response by releasing a mouse button with only the right hand is rather simple and may not require as much time to be re-assigned, thus accounting for the absence of switch cost for within-group switches.

In their recent ERP and fMRI studies, Rushworth et al. (2005, 2001) used a design where they required subjects to pay attention to the color or the shape of one of two presented stimuli in order to detect a rare target. While they did find significant switch cost in RT, Rushworth et al. (2005, 2001) only considered switches between stimulus attributes, but they did not compare them with switches within attributes.

Switch cost is believed to arise when residual activity in the neural cicuitry for the previous task interferes with performance in the current task. Using face and word discrimination tasks that activate well-known brain regions, Yeung et al. (2006) were able to demonstrate this process in prefrontal cortex using fMRI. Their results imply competitive interactions between areas responsible

for the two conflicting tasks, akin to biased competition in the model of selective attention by Desimone and Duncan (1995).

Unlike other task switching studies (e.g. Sohn et al. (2000); Kimberg et al. (2000); Altmann (2004), but see Rogers and Monsell (1995)), we did not find a residual switch cost for long CTIs. There is an ongoing debate in the task switching literature about the factors that affect residual cost.

Although we tried to equalize the difficulty of the two task groups during training, there is a significantly higher RT mixing cost for IMG (53 ms) than for COL (30 ms) tasks, which might be compensated by an opposite effect in error rate mixing cost (0.4 % for IMG, 4.2 % for COL) in a speed-accuracy trade-off. Switch cost, on the other hand, does not depend on task group. The paradigm of distinguishing mixing and switch cost allows us to catch such variations in the mixing cost while keeping switch cost unaffected by them.

We see a similar effect in inter-subject variability. While RT mixing cost shows significant dependence on subject identity, there is no such dependency for RT switch cost. This suggests that the processes involved in shifting attention are stereotypical among individuals, implying an automated process with fixed processing duration. Mixing cost, on the other hand, shows more individual differences in processes such as cue perception and interpretation and additional effort for having to potentially solve two tasks instead of one.

What do our results mean for the top-down control of visual perception? Due to its short processing time, fast object detection in natural scenes of the sort shown by Thorpe et al. (1996) is assumed to be possible in a purely feed-forward, hierarchical model of the ventral pathway (Thorpe et al. 2001; Riesenhuber and Poggio 1999b). Switching top-down attention to a different feature value within the same stimulus attribute requires biasing feed-forward connections in such a hierarchical system at fairly high levels (see chapter 4 for a computational model), e.g., in inferotemporal cortex (IT) or even the connections from IT to prefrontal cortex (PFC) for object categories (Freedman et al. 2003). For example, two different classifiers, possibly located in the PFC, would access the same data in IT to decide whether or not an animal or a vehicle is present in the image (Hung et al. 2005).

When switching to a different stimulus attribute processed by a different visual area, one would assume that task-specific biasing of neural activity has to happen at an earlier stage in the hierarchy, before specialization of processing streams takes place. In the case of switching between color and object detection, this could be area V4, V2, or even V1. Our current results indicate a higher cost in RT for task switches between attributes, i.e., for biasing at an earlier stage, than within attributes, i.e., biasing at a later, more specialized stage. This finding agrees with ideas of a reverse hierarchy put forward by Hochstein and Ahissar (2002).