Perceptual switching, eye movements, and the bus paradox

Junji Ito, Andrey R Nikolaev, Marjolein Luman, Maartje F Aukes, Chie Nakatani,
Cees van Leeuwen¶
Laboratory for Perceptual Dynamics, Riken BSI, 2-1 Hirosawa. Wako-shi, Saitama, 351-0198 Japan;
e-mail: ceesvl@brain.riken.go.jp; http://www.brain.riken.go.jp

Abstract. According to a widely cited finding by Ellis and Stark (1978 *Perception* **7** 575 – 581), the duration of eye fixations is longer at the instant of perceptual reversal of an ambiguous figure than before or after the reversal. However, long fixations are more likely to include samples of an independent random event than are short fixations. This sampling bias would produce the pattern of results also when no correlation exists between fixation duration and perceptual reversals. When an appropriate correction is applied to the measurement of fixation durations, the effect disappears. In fact, there are fewer actual button-presses during the long intervals than would be expected by chance. Moving-window analyses performed on eye-fixation data reveal that no unique eye event is associated with switching behaviour. However, several indicators, such as blink frequency, saccade frequency, and the direction of the saccade, are each differentially sensitive to perceptual and response-related aspects of the switching process. The time course of these indicators depicts switching behaviour as a process of cascaded stages.

1 Introduction

Our visual system normally provides us with reliable, unambiguous information about the world. Yet numerous displays could be created, especially pictorial ones, for which two or more alternative interpretations are frequently reported. These *ambiguous patterns*, amongst them the widely known Necker cube, play an important role in the study of human visual perception. Brief expositions of these stimuli reveal preferences for one of several possible alternative interpretations. In ambiguous dot patterns, quantitative laws regulate the preferences as a function of relations between its components such as proximity in space (Kubovy et al 1998), proximity in space and time (Gepshtein and Kubovy 2000), or contrast (Ohtani et al 1995).

With prolonged viewing of an ambiguous display, the perceiver tends to reverse the preference. Basic aspects of this mechanism are still unknown. An ongoing debate, for instance, is whether reversals occur spontaneously, as traditionally assumed (Attneave 1968, 1971), as a result of neural mechanisms such as fatigue (Köhler 1940; Köhler and Wallach 1944), stochastic fluctuation (Borsellino et al 1972), or intrinsic metastability (Kelso et al 1995; van Leeuwen et al 1997); or alternatively are induced by top – down influences (Horlitz and O'Leary 1993; Leopold and Logothetis 1999); or perhaps are the product of both spontaneous and top – down mechanisms working in concert (Long and Toppino 1994).

With viewing prolonged over a further extended period of time, preference reversals occur repeatedly, resulting in the well-known phenomenon of perceptual switching (Borsellino et al 1972; Hock et al 1997). Hock et al (1997) studied the effects of possible neural fatigue, and found that such a mechanism is not sufficient to explain the time distribution and that stochastic fluctuations are necessary. Borsellino et al (1972) showed that a gamma distribution provides a good fit for the interval between two switches. Such a distribution is obtained theoretically if switches are independent,

¶ Author to whom all correspondence should be addressed.

Poisson-distributed events (Baddeley 1999). This seems to imply that switching is essentially a stochastic process (De Marco et al 1977). However, it was recently shown that deterministic processes at neural level could give rise to the switching distributions (van Leeuwen et al 1997).

Neurophysiological studies suggest that switching consists of activity modulation within distributed, cooperative brain structures. Switching corresponds to increased activity in higher-order visual areas, in combination with reduced activity in early visual areas, including the primary visual cortex and the pulvinar (Kleinschmidt et al 1998).

The involvement of the pulvinar suggests a possible relation with eye movement. The pulvinar belongs to the secondary visual system known to be involved in calibrating the visual image with respect to movement of the eyes. Whereas there is enhanced activity before and after the saccade, reduced activity in the pulvinar has been observed in monkey during saccades (Benevento and Port 1995). Saccades are known also to lead to sensory suppression (Volkmann et al 1980; Uchikawa and Sato 1995; Ridder and Tomlinson 1997), leading in turn to reduced activity in the early visual cortical areas. The suppression of activities in these areas in relation to saccades parallels that in switching, and this further suggests a relationship between the two.

Our current study is aimed at determining the relationship of switching behaviour and eye movements. Eye-movement recording is one of the most suitable techniques for on-line monitoring of perception. Indices based on eye movements have been used for human cognitive processes, such as text comprehension, object identification, and scene viewing (Kowler 1990; Rayner 1998; Underwood 1998). In high-level mental tasks, a response often takes more than 1 s to complete. Thus, the response latency is a sum of processing times of various processes over 1 s. On the other hand, eye-movement measures, such as saccade measures, give higher temporal resolution (about 3 saccades s^{-1}) and more dimensions (eg fixation duration, fixation location, saccade length, and saccade direction). This offers great opportunities for using this technique with ambiguous figures.

In spite of these advantages, eye movements have not frequently been studied in ambiguous displays. An important exception is found in Ellis and Stark (1978). The authors showed Necker cubes as ambiguous figures for extended periods. Participants responded by button-press to the occurrence of perceptual switches. The authors studied fixation durations (the time interval between two saccades) for these figures. They reported longer fixation durations for intervals in which a switch occurred as indicated by a button-press, compared to surrounding fixations. The effect sizes they reported were in the order of 200 ms.

The observed relation between eye movement and perceptual processes may be considered interesting in its own right. In addition, an effect of this size would provide a reliable behavioural indicator for a mental event (the reversal of a perceptual interpretation) unaccompanied by changes in external stimulation. This is a rare and valuable asset in the research of mental processes in, for instance, a psychophysiological context, where each change in stimulation typically triggers a cascade of neural responses. It is, therefore, surprising that this result, although widely cited, has not been widely used in this domain of research.

We suspect, however, that the effects as reported are the result of an illustrative statistical sampling error, one that is likely to be made in other domains of eye-movement literature as well. The error has become known through the bus paradox (Baddeley 1999). This paradox deals, among others, with the issue how a bus with a frequency of, say, 10 min may take even longer, on average, than 10 min to appear at the bus stop. If the busses arrive exactly every 10 min, your expected waiting time would be 5 min. But if the frequency fluctuates, this time is increased. (For similar reasons you tend to be waiting in the slowest cue at the supermarket.) This is not because of Murphy's

law, but simply because the probability of your arrival at the bus stop per time unit is independent of that of the busses, and therefore equal *per time unit* for short and long intervals. As a result, you are more likely to arrive, and hence be waiting, in a long interval between two busses than in a short one. In extreme cases, such as when the busses clog up in clusters of three, your expected waiting time will be at least 15 min. This may illustrate that fluctuation determines the extent of the effect. But given a certain amount of fluctuation, independence of bus arrival and your arrival is sufficient to enhance your waiting time.

The occurrence of a perceptual switch corresponds to your arrival at the bus stop. Fixation duration—the interval between two saccades—corresponds to the interval between two busses. Let these intervals fluctuate, for instance by a gamma distribution (Borsellino et al 1972). Let switches be independently distributed events. Then this would be sufficient, in absence of any correlation between fixation duration and switches, to produce the effect observed by Ellis and Stark (1978).

In their analyses, the authors compared the length of a fixation in which the switch occurs to that of the surrounding fixations. This means that they compared the length of intervals sampled with the bias resulting from the bus paradox to intervals sampled without the bias. They do not report using any control or correction for the sampling bias. It is possible, therefore, that what they reported is entirely based on the bias.

For these reasons, the first specific aim of the present study is to replicate Ellis and Stark (1978), using measures to control for the bias. The second one is, if necessary, to propose alternative eye-movement-based indicators for perceptual switching and provide our own account of the switching process based on these measures.

In our replication, we took another ambiguous display instead of the Necker cube. Both alternative interpretations of the Necker cube are three-dimensional. Three-dimensional interpretations of two-dimensional line drawings take considerable time to complete (Sekuler and Palmer 1992) and thus might contribute irrelevant variance to the switching process. We used ambiguous figures that do not invoke a third dimension.

A version of Attneave's (1968) triangles was used in the present experiment. These are ipsilateral triangles which have their angles pointing in the directions: 3 o'clock, 11 o'clock, and 7 o'clock. Attneave observed that these triangles are perceived alternatively to point in any of these three directions with approximately equal preference. The preferences for these directions have been investigated in relation to motion (Bucher and Palmer 1985), texture (Palmer and Bucher 1982), and configuration (Palmer and Bucher 1981). Spontaneous switching between these directions has been reported for prolonged viewing conditions (Attneave 1968, 1971). We used a modified version of the original stimulus, which has filled instead of outline triangles. It consists of seven black triangles in a set composition on a gray background (figure 1). In all conditions the same stimulus was used.

In our experiment, we will be looking at fixations in which a switch occurs, and compare their durations with those of flanking fixations after a bias correction. Following the method in Ellis and Stark (1978), one fixation before and one after a switch response were taken as the baseline fixations. This removes the need for separate experimental treatment and baseline conditions. Instead, treatment conditions were used to evaluate the robustness of the effect. There are, however, no a priori hypotheses about how treatment conditions will affect the uncorrected or corrected scores.

One factor is *instruction*. Often instructions relating to switches implicitly encourage reporting switches. This implies a set bias to report these switches. We would like to observe effects independently of such bias. For this reason, instruction either encouraged or discouraged participants to report switches. We expect that more switches will be reported in the former condition than in the latter, which could affect the eye-movement distributions in a variety of ways.

The other factor is *response*. One of the problems in switching studies is that the timing of the switch is always indicated by a behavioural response, as in Ellis and Stark (1978) a button-press. Obviously this response can interfere with any subtle measures that are derived from eye movement. (A similar problem occurs with electrophysiological studies of switching.) There is no principled solution to this problem, because we cannot simply subtract a manual response component. Our proposal is that at least we should observe independence of the observed phenomenon of the difficulty of the manual response. We compared compatible and incompatible response conditions.

2 Method
2.1 Participants
Twenty-two participants took part in this experiment: five men and seventeen women. The mean age was 20.5 years with range of 19 to 30 years. All participants had normal or corrected-to-normal vision. Participants received 1000 yen per hour for taking part in the experiment.

2.2 Stimuli
The stimulus was the display of figure 1. The diameter of a circle containing the seven black triangles was 25 cm. The entire stimulus was a 600 by 800 pixel bitmap image. Stimulus size was 27.4 cm by 20.6 cm, 90 cm from the participant, yielding approximately 17 deg of visual angle. The cluster of triangles can be seen as pointing in one of three directions: 3 o'clock, 7 o'clock, or 11 o'clock.
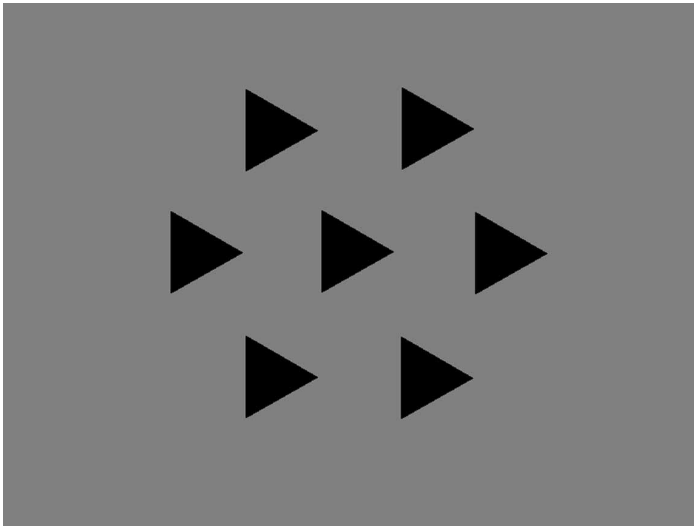
Figure 1. Seven black triangles on a gray background used as stimuli in our experiment.

2.3 Apparatus
The experiment was controlled by a SMI Eyelink system that consisted of two PCs (one for stimulus presentation, the other for data recording), a headset for eye-movement recording and a response-button box. The system presented the stimulus, and recorded eye movements and responses (button-presses) during the task. The stimulus was presented on a 21-inch CRT display with an 85 Hz refresh rate. The eye position was sampled every 4 ms. The button box for collecting responses had four buttons and three of them were used (figure 2).
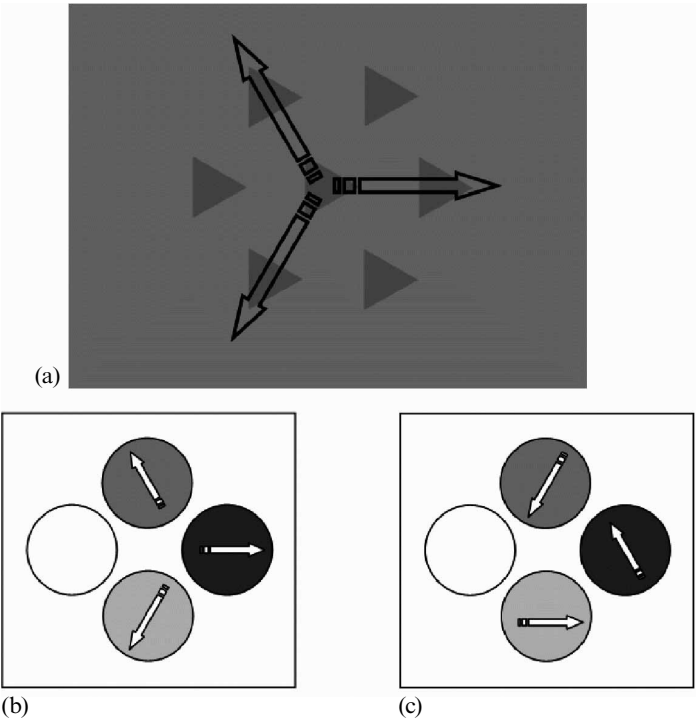
Figure 2. (a) Possible directions of perceptual switches; locations of buttons of the button box in (b) compatible conditions; and (c) incompatible conditions.

2.4 Design

Two instructions about perceptual switching and two response mappings were employed. In the *switch* condition, the participants were asked to report spontaneous switches of the perceived direction of the triangle cluster (not individual triangles) by a button-press. In the *non-switch* condition, the participants were asked to try to refrain from switching as much as possible, but still to report perceptual switches by a button-press. Two alternative versions of the response mapping were applied to both switch and non-switch conditions. In the *compatible* condition, the locations of the response buttons were made approximately compatible with the pointing-direction of the triangles (figure 2b). In the *incompatible* condition, the locations of the response buttons are shifted clockwise; thus the location of the buttons and perceived direction was made incompatible (figure 2c). The switch versus non-switch instruction and two response mappings made four conditions: switch/compatible, switch/incompatible, non-switch/compatible, and non-switch/incompatible conditions. The four conditions were counterbalanced over participants by means of a Latin square.

2.5 Procedure

Participants were seated in front of a monitor in a room with dimmed lights. Instructions were read to the participant, while showing the triangles on the screen. After the instruction of one of the four conditions, a 1 min practice session was given to the participants without eye-movement recording. Participants were asked to learn to use the button box, for which a sheet showing which response each button indicates was available. Also, they were instructed to hold the button box face up in their right hand, and hold the box parallel to the sagittal line of their body. Buttons were pressed with the right thumb only, so participants had to move their thumb onto the appropriate button for reporting each switch. In all conditions, participants were instructed

to find a stable direction of the triangle cluster at the beginning of a session, then to start reporting the perceptual switching of the pointing direction of the cluster. After practice, participants put on the headset for the eye-movement recording and completed a session that took approximately 25 min. Each session consisted of one of the four above-mentioned conditions. Before each session, eye positions were calibrated by using nine calibration points that covered the stimulus display. To further ensure the accuracy of the eye-position recording, a session was divided into twenty 1 min segments after each of which loss of vertical accuracy was corrected by drift correction. Also, a recalibration was applied whenever necessary and at least twice per session. A break of 20 min was taken between the sessions in which participants were released from the headset. The total time for the experiment was approximately 4 h.

3 Results and discussion
Eye events, such as fixation duration, fixation position, saccade duration, and blinks, were extracted from the eye-position record, for which an eye-event filtering program routine of the Eyelink system was used. The response-button presses and their timing were extracted and superimposed on the eye-event sequence.

Analyses of the data consist of three parts: the first provides the scores for how the patterns were preferably perceived. Second, the effect of switching on the mean fixation duration was analysed. Third, we analysed the time series of eye events for possible alternative behavioural indicators of perceptual switching.

3.1 Preferences
Perceptual switches among the three directions were reported in the task. Repeated-measures ANOVA on button-press frequency corresponding to each of three directions in twenty-two participants in the compatible condition revealed a significant effect of preferred pointing direction ($F_{2,42} = 6.4$, $p = 0.004$). The 3 o'clock and 11 o'clock directions were more preferred than the 7 o'clock pointing direction (a posteriori $p = 0.002$ and $p = 0.02$, respectively; see figure 3).
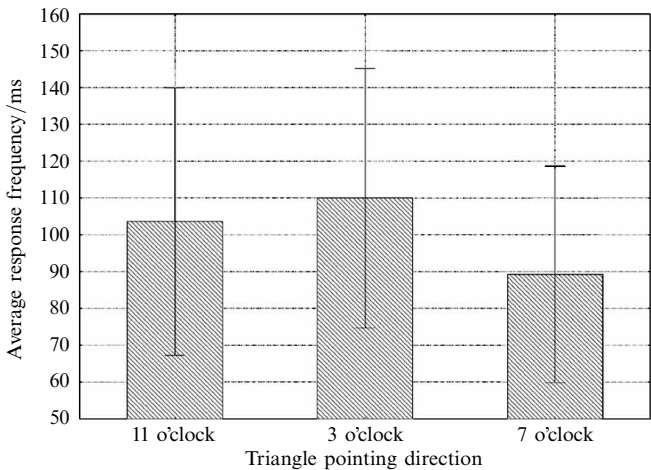
Figure 3. Means and confidence intervals of response frequency distribution of triangle pointing direction by button-presses in the compatible condition.

3.2 Mean fixation durations
Six participants were omitted from this analysis; four for having exceptionally long blinks—according to self-reported evidence due to sleepiness, one for poor recording, and one for too small a number of button-presses. Button-presses were used as a marker for perceptual switches. We calculated the mean fixation durations for all conditions

in the fixations where button-presses occurred and in their flanking fixations. These were initially compared without bias correction, to see if this procedure replicates Ellis and Stark (1978)—if the mean fixation duration at the button-press is about 200 ms longer than those of flanking fixations, the same effect reported in Necker-cube reversal is replicated for the triangles. Subsequently, bias corrections were applied by using surrogate series of eye events and button-presses. If the increase of fixation duration is due to the sampling bias, the results of the surrogate and actual fixation series will be identical.

Mean fixation duration was computed for each participant and condition, to remove outliers. For every participant, the fixations that deviated more than 2.5 SDs. from its mean were excluded from analysis (2.7% of all fixations). The mean number of switches reported was 256.2 for the switch/compatible condition, 220.6 for the switch/incompatible condition, 61.7 for the non-switch/compatible condition, and 62.7 for the non-switch/incompatible condition. Only button-presses that occurred during a fixation were used in the analyses. Thus, button-presses that fell between two fixations (during saccades or blinks) were excluded from the analysis (18.5%). The durations of the fixations that precede, contain, and succeed the button-press are referred to as reorganisation fixation '0' and temporally flanking fixations '−1' and '1' as in Ellis and Stark (1978).

An ANOVA [$3 \times 2 \times 2$; three sequential positions (−1, 0, and 1), by two switching instructions, by two response mappings] was performed on fixation durations. The main effect of the sequential position was significant ($F_{2, 14} = 29.7$, $p = 0.000$). Averages over all conditions are shown in table 1. In the series of eye events uncorrected for sampling error, fixations containing the button-press (the '0' fixation) were longer on average (543 ms, SD 210 ms) than those of the flanking fixations (412 ms, SD 141; and 404 ms, SD 150; respectively for the '−1' and '1' fixation). No effects of instruction or response conditions were found significant in this analysis. The results for the uncorrected data replicate the Ellis and Stark (1978) study, although the effect sizes are somewhat smaller.

Table 1. Mean fixation durations.

Condition	Fixation			Average
	−1	0	1	
Switch	400.2	509.6	391.0	433.6
Non-switch	425.6	559.3	417.1	467.3
Compatible	412.2	528.3	419.7	453.4
Incompatible	413.6	540.6	388.5	447.5
Average	412.9	534.5	404.1	450.5

3.3 Surrogate series
Three different types of surrogate series were used to eliminate the sampling error. In the surrogate series, pseudo button-presses were generated and replaced the originals. The number of the pseudo button-presses was the same as that of the original. In the *inverse* surrogate series, the sequence of the original button-presses was inverted; the time index of the original button-presses, t, was transformed by $t_{surrogate} = 20$ min (ie total task time) $- t$. The inverse pseudo button-press sequence was inserted into the original eye-event sequence, after making sure that no linear trend applied to the data. The inverse surrogate series preserves in reversed order all the dependences of the original button-press series. In the *Poisson* surrogate series, pseudo button-presses were randomly inserted in the time series, with a Poisson distribution the parameters of which were estimated from the original series. This surrogate series uses the theoretical distribution

of button-presses as random events. In the *uniform* surrogate series, pseudo button-presses were placed at uniform distances across the time continuum with an interval equal to the average time interval between the real button-presses in the original series. This surrogate series minimises fluctuation in the pseudo button-press series.

In all three of the surrogate series, durations for fixations preceding '−1', containing '0', and following '1' the pseudo button-presses were marked. As in the original scores, pseudo button-presses falling within a saccade or blink were again not taken into account. The averages based on the pseudo button-presses have the same sampling biases as the original series. A corrected score for each interval was obtained by taking the difference between the real and the surrogate series. According to Ellis and Stark (1978), we would expect a positive score for the '0', compared to '−1' and '1' fixations.

For the 'inverse', 'Poisson', and 'uniform' correction method, three separate ANOVAs (the three sequential positions, −1, 0, and 1 of fixation, by two switching instructions, by two response compatibilities) were performed on the difference score (see table 2). The main effect of position was significant in all three methods (for the 'inverse' method, $F_{2, 14} = 4.8$, $p = 0.026$; for the 'Poisson' method, $F_{2, 14} = 7.8$, $p = 0.005$; and for the 'uniform' method $F_{2, 14} = 3.9$, $p = 0.028$). After the sampling bias was corrected with the 'inverse', 'Poisson', or 'uniform' method, the scores of the '0' fixation, the one containing the button-press, turned out to be *lower* than those of the '−1' and '1' fixations surrounding the button-press, and assumed negative values.

The negative scores for the '0' fixations allow us to conclude that the effects reported in Ellis and Stark (1978) could only be replicated for the observations uncorrected by sampling bias. When sampling bias was corrected, fixation durations at button-presses indicating a switch appear to be shorter than expected under the null hypothesis.

Table 2. Differences of means from those of surrogate series obtained by inverse, Poisson, and uniformly distributed methods.

Condition	Fixation			Average
	−1	0	1	
Inverse				
Switch	−4.9	−30.5	−15.1	−16.8
Non-switch	−5.5	−75.3	−30.1	−37.0
Average	−5.2	−52.9	−22.6	−26.9
Compatible	−12.9	−70.7	−14.9	−32.8
Incompatible	2.5	−35.0	−30.3	−20.9
Average	−5.2	−52.9	−22.6	−26.9
Poisson				
Switch	5.6	−36.7	−14.7	−15.3
Non-switch	−0.8	−64.1	−12.0	−25.6
Average	2.4	−50.4	−13.3	−20.4
Compatible	−4.1	−64.4	2.9	−21.9
Incompatible	8.9	−36.4	−29.5	−19.0
Average	2.4	−50.4	−13.3	−20.4
Uniform				
Switch	−0.2	−30.7	−13.4	−14.8
Non-switch	3.4	−43.5	−18.9	−19.7
Average	1.6	−37.1	−16.1	−17.2
Compatible	−4.3	−65.6	3.9	−22.0
Incompatible	7.5	−8.5	−36.2	−12.4
Average	1.6	−37.1	−16.1	−17.2

The result therefore is opposite to what is predicted by Ellis and Stark (1978). We may, therefore, conclude that fluctuations in the switching intervals, rather than actually longer fixations, are responsible for their effect.

Instruction was shown to have a large influence on this effect. As shown in table 2, scores for the '0' fixations are more pronouncedly negative in the non-switch than in the switch condition, in the 'inverse' method -75.3 versus -30.5 ms ($F_{1, 60} = 3.9$, $p = 0.05$). The non-switch condition contains fewer button-presses (about one-third of the number). If we consider that in this condition the response threshold for switches is high, the few ones that do exceed this threshold will be accompanied by particularly powerful events. It seems therefore plausible that their effect is comparably larger. For response compatibility, the interaction effect between the fixations and response condition reached significance in the inverse method ($F_{2, 14} = 2.1$, $p = 0.016$), the Poisson method ($F_{2, 14} = 6.2$, $p = 0.012$), and in the uniformly distributed method ($F_{2, 14} = 4.7$, $p = 0.028$), see table 2. All three methods resulted in a greater reduction of the '-1' and '0' fixations for the compatible than for the incompatible conditions, but an opposite trend for the '1' fixation (table 2). No other effects approached significance.

The analyses for the corrected scores show that lengthening of fixation durations around the button-press does not occur in any of the response conditions. The observations that fluctuations in switching intervals produce the effect, therefore, is independent of instruction or response conditions. The differences observed between these conditions, however, allow for the possibility that these conditions affect the fluctuations, thereby indirectly affecting the scores.

From the negative scores for the '0' fixations we might be tempted to infer that fixations containing button-presses are actually shortened compared to others. Such an effect could occur as a consequence of the disruption of the fixation by the button-press. Or perhaps it could occur because the perceptual switch interrupts the current scanning strategy. In both cases it would be difficult to explain why shortening is also obtained for the fixation after button-press. Disruptions of this kind would lead to an increase in the duration of the fixation following the button-press, as recomputation of the scanning path will be needed. Alternatively, an increase in the rate of eye movements is known to lead to sensory suppression (Volkmann et al 1980; Uchikawa and Sato 1995; Ridder and Tomlinson 1997). Sensory suppression may destabilise the percept, which would increase the relative impact of mechanisms that suggest an alternative interpretation.

Drawing this conclusion, however, would be premature, as it is based on corrected scores. The correction was made on the assumption of independence between switches and eye movements. For illustrative reasons, an actual reduction of '0' fixation duration is not necessary for obtaining negative scores. Note that in the bus paradox, the larger the fluctuation in intervals, the larger the extent of the biases. Hence, what looks like shortening may, for instance, equally well be a reduction in the *fluctuation* of fixation durations in relation to the switching. All we can conclude, therefore, is that fixation and switch are not independent. In order to find out whether the dependence involves reduction of fixation duration, a more fine-grained analysis is needed, to which we turn next.

3.4 Time series
Intervals between button-presses varied widely between participants. From the original twenty-two participants those participants were selected who had mean intervals between button-presses of more than 3 s, leaving out seven. One participant was excluded owing to poor recording. The data of the remaining fourteen participants were analysed. Time series analyses were performed on a subset of the conditions that had enough measurement points to permit this (specified for each analysis individually).

We performed moving-window analyses on *blink* and *saccade frequency*, as well as on *saccade directions*. Blink frequency is known to decrease when concentrated effort is needed to perform a task or 'to uptake' information (eg De Jong and Merckelbach 1990; Zangemeister et al 1995; Veltman and Gaillard 1998). Blink frequency is used in the present study as a measure of effort. If any moment relating to a switch (or its manual response) involves effort, this should be observed in the reduction of blink frequency. Saccade frequency analysis is the time-series counterpart of looking at fixation durations. A moving-window analysis will allow us, amongst others, to decide whether fixation durations are actually shorter around button-presses. A saccade direction is the angle of the vector that traces the movement from one fixation location to another. The angle is measured relative to an arbitrary origin 0, for which we used the direction 3 o'clock. We investigated if a systematic dependence could be found between the angle of a saccade and the perceived direction indicated by the button-press.

3.4.1 *Blink frequency.* To investigate the temporal properties of the blink time series around button-press we used a moving-window with a length of 200 ms and a step of 10 ms within the interval between 10 s before and after each button-press. The obtained single-blink time series were superimposed with the button-press used as origin and averaged. The averages were multiplied by 5 to obtain values per second. The mean blink frequencies were calculated for the overall 20 s interval, and also for 1500 ms before and after button-press.

Figure 4 shows that before the response the blink frequency is reduced; it reaches its minimum approximately $700 - 600$ ms before button-press. Around the button-press a compensatory increase of blink frequency occurred. A Wilcoxon paired test showed that the mean blink frequency in the interval $< -1500, 0 >$ ms was significantly lower than the overall mean ($Z = 2.04$, $p = 0.04$), and the mean blink frequency in the interval $< 0, 1500 >$ ms was significantly higher than the overall mean ($Z = 2.35$, $p = 0.02$).

No unique moment in the blink frequency curve can be identified as the occurrence of the switch. Rather, an extended wave is manifested in the blink response, a wave
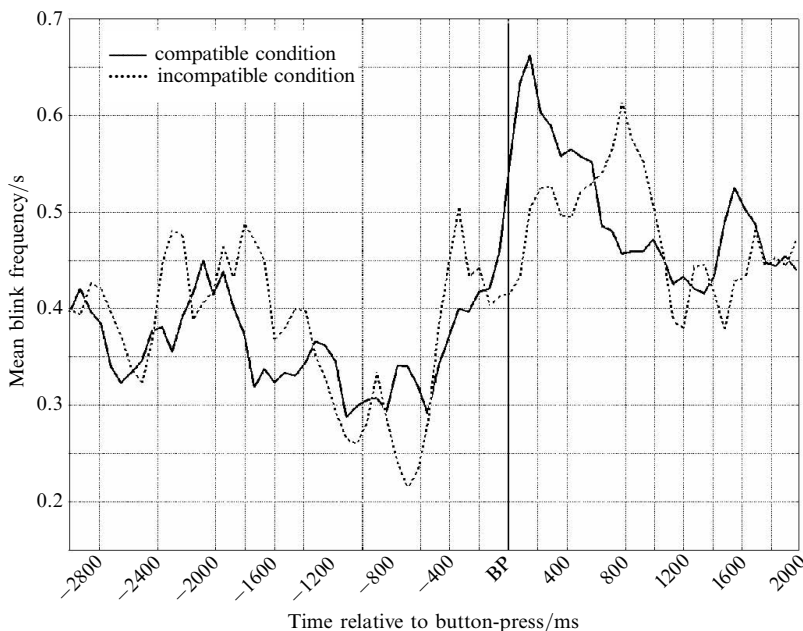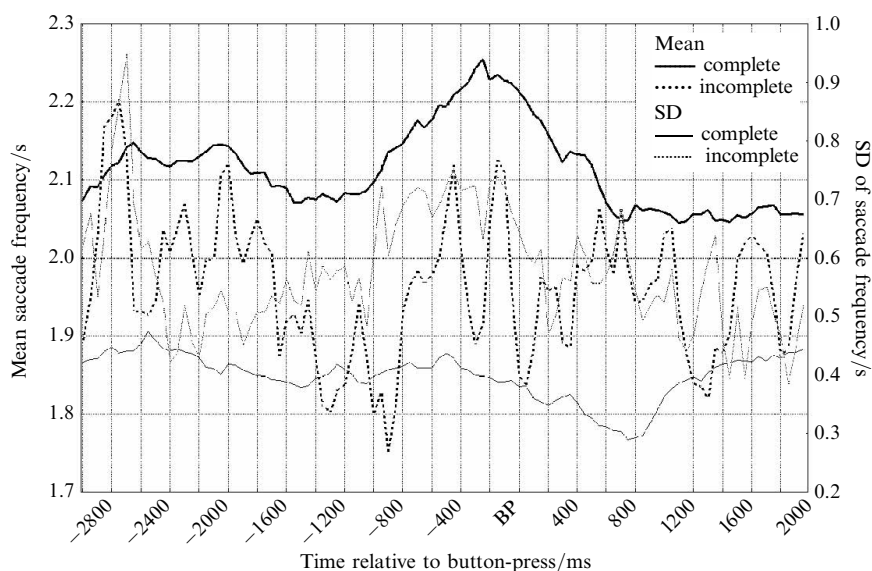

Figure 4. Mean blink frequency in the incompatible and the compatible conditions in the interval $< - 3000, + 2000 >$ ms about button-press (BP) (fourteen participants).

that extends from 1100 ms prior to 600 ms after the button-press. Reduced blinks have been associated with effort, and thus we may view the wave as a slow fluctuation in the effort needed to complete the task. Such an extended period is likely to contain perceptual as well as response preparation components of the switch. Effort may be required, in principle, by attentional demands involved in perceptual reorganisation, as well as response preparation. The transitory increase in blink frequency after the button-press can be interpreted as relaxation after the completion of the response.

In the compatible and incompatible conditions, the blink frequency curves in the interval about the button-press $<-1000, 500>$ ms are very similar ($r = 0.72$). In the incompatible condition, the curve seems to start rising slightly earlier and the peak around the button-press seems less pronounced than in the compatible condition (figure 4). However, a Wilcoxon paired test on blink frequency maxima for fourteen participants revealed that neither their timing ($Z = 1.5$, $p = 0.13$), nor amplitude ($Z = 0.91$, $p = 0.36$) differed significantly. Response incompatibility seems to have no substantial effect on the blink frequency curve. The curve appears to be equally time-locked to button-press in both conditions. If most of the effort had gone into the perceptual components of the task, there would have been greater temporal jitter in the incompatible than in the compatible condition. We may therefore conclude that the blink frequency curve represents the attentional demands on the response-related aspects of the task.

3.4.2 *Saccade frequency.* Saccade frequency is inversely related to fixation duration (the time interval between two saccades). We analysed saccade frequency in the same way as blink frequency: using a moving window with a width of 200 ms and a step size of 10 ms (figure 5). In the compatible condition, saccade frequency shows a clearly marked peak at -250 ms before button-press. This finding constitutes a direct proof that fixations around button-press are actually *shorter* than surrounding fixations.

The saccade frequency in the incompatible condition shows high amplitude fluctuations, which are almost of the same range as its SD. This observation may be explained by jitter of the button-press response. So the saccade frequency curve, unlike the blink frequency curve, is not strictly time-locked to the response. Saccade frequency

Figure 5. Means and standard durations (SDs) of saccade frequency in the incompatible and the compatible conditions in the interval $<-3000, +2000>$ ms around button-press (BP) (fourteen participants).

(and thereby fixation duration) appears to be more closely related to the perceptual reorganisation aspects of the task. The relation between fixations and perceptual reversal, however, is opposite to that suggested by Ellis and Stark (1978).

In our analyses involving surrogate series, we proposed that the most likely account of these effects was associated with increased saccade frequency resulting in destabilisation of the percept. Saccades are associated with suppression of sensory activity. This suppression will reduce the relative impact of mechanisms that suggest alternative interpretations. Sensory suppression has been associated with both saccades and blinks (Volkmann et al 1980; Uchikawa and Sato 1995; Ridder and Tomlinson 1997). Saccadic gaze shifts are often accompanied by blinks (von Cranach et al 1969; Fogarty and Stern 1989; Evinger et al 1994), and are time-locked to them (Zuber and Stark 1966; Rottach et al 1998). This leads us to investigate the temporal relationship between the respective time courses of saccade and blink frequencies.

Cross correlations were calculated between the respective curves for each participant in the compatible condition. Cross correlations that are greater than twice their respective standard errors were considered as significant. In eight of fourteen participants a positive cross-correlation was observed. In six participants saccade frequency curve precedes the blink frequency curve with a mean lag of 483 ± 147 ms. In two participants the saccade frequency curve remained behind the blink frequency (100 and 600 ms, respectively). The observed cross-correlations are in line with the previously observed relationships. The lack of consistent time order across participants, however, contradicts their time-locked character as observed in the literature (Zuber and Stark 1966; Rottach et al 1998). The current observation in our view does not invalidate the interpretation of saccade and blink frequency as related, respectively, to the input and output aspects of the task. It does, however, rule out a simple serial model, in which perceptual reorganisation strictly precedes the preparation of the response.
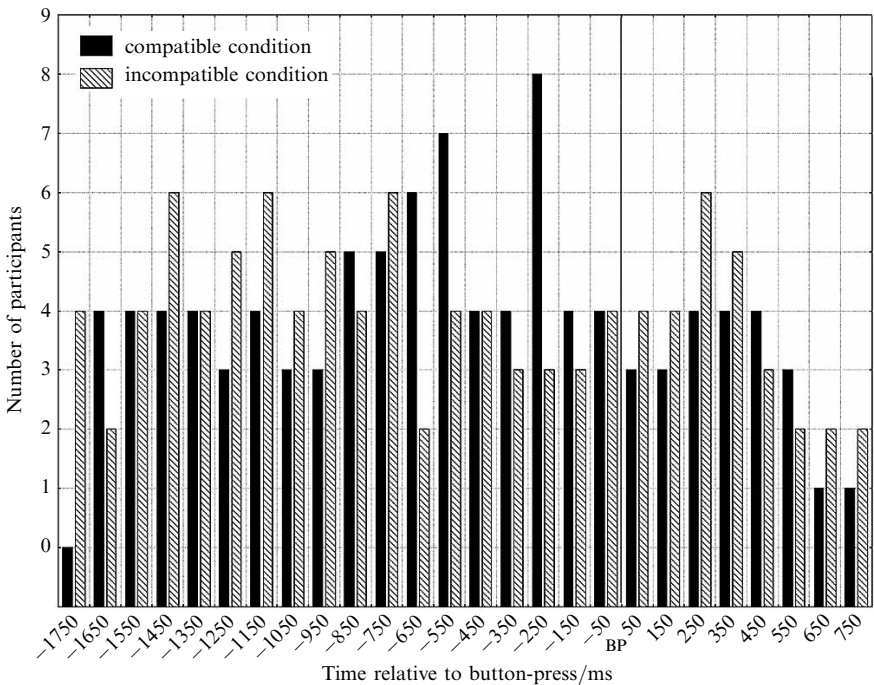
3.4.3 *Saccade directions.* A saccade direction is defined as the angle of the vector that traces the movement from one fixation location to another. The angle is measured relative to an origin 0, for which the SMI Eyelink system uses the direction 3 o'clock; 11 o'clock is $+120°$ and 7 o'clock is $-120°$ (figure 2). We investigated whether any systematic dependence could be obtained between saccade directions and the pointing directions as indicated by button-press responses. First the saccade direction distribution was tested against the uniform distribution in three 1 s time windows by a chi-square procedure (figure 6). The distribution of saccade directions was not uniform in most participants in all conditions. The exception is in the $<-1000, 0>$ ms interval before button press in the non-switch/compatible condition. This is the condition where switches are discouraged. The absence of effect in this condition could indicate efforts to suppress a switch by making random eye movements.

To determine more precisely the time course of changes of saccade directions the latter were compared with a Kruskal–Wallis ANOVA with factors response (3 levels, the triangle pointing directions indicated by button-presses) in moving-window analysis with a width of 500 ms and a step of 100 ms in the interval between -2000 and $+1000$ ms relative to the button-press. This was done for each participant separately. Only one participant had no significant difference in any of these time windows. The number of participants with a significant difference in a certain time interval is shown in figure 7.

Duration of intervals with congruent saccade directions varied substantially between individual participants: 907 ± 412 ms (compatible condition); 821 ± 387 ms (incompatible condition), but revealed no difference between conditions ($Z = 0.71$, $p = 0.48$). For the compatible condition, the number of participants with saccade directions congruent to the triangle pointing direction peaked in the $<-650, -250>$ ms interval before
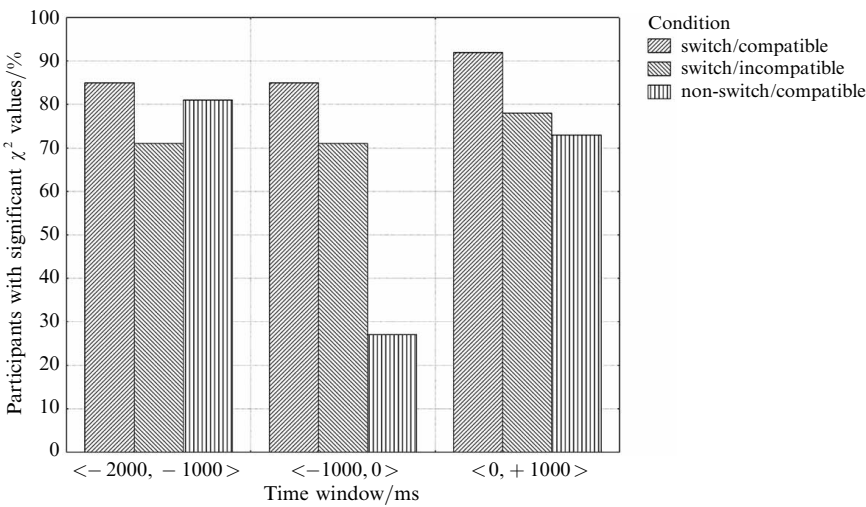
Figure 6. Percentages of participants with significant chi-square values when testing the observed distribution of saccade directions against the uniform distribution in three time windows for switch/compatible, switch/incompatible, and non-switch/compatible conditions.

Figure 7. Number of participants with significant difference in saccade directions congruent to pointing directions of the triangles on the display.

button-press. In the incompatible condition, these differences were more diffuse in time and tended to be shifted to earlier intervals, reaching a maximum in the interval $<-1500, -750>$ ms before button-press.

Importantly, a Wilcoxon paired test showed that individual participants show the largest systematic difference in saccade directions at the same time as the peak in saccade frequency ($Z = 0.28$; $p = 0.78$). In other words, when saccade frequency is at its highest, they are also maximally congruent with the subsequently reported pointing direction.

This result confirms our interpretation of saccade frequency as related to the perceptual reorganisation stage of the switch.

Moreover, by the same reasoning as for the saccade frequencies, the increased diffuseness of the maximum in the incompatible condition suggests that the observations are time-locked to the perceptual reorganisation process, rather than to the response. The delay in the peak in this condition suggests delay in response, in accordance with what has been found in the literature on finger response compatibility (Stins and Michaels 2000). The extra time needed after a switch for making an incompatible motor response diffuses the peak of the effect, but does not seem to affect the strength of the relation of eye movement and perceptual switch direction.

The saccade directions at the intervals of highest significant difference, as previously obtained in the Kruskal–Wallis ANOVA, were determined for each participant separately. The obtained angles were averaged for each participant and for each button. Repeated-measures ANOVA on these mean angles with factor of switch direction (three buttons corresponding to response directions) was performed for a group of fourteen participants. The tested effect was strong ($F_{2,26} = 15.5$, $p = 0.00004$, figure 8) and, moreover, a posteriori Duncan test revealed differences between *all* three switching directions. Notably the direction of eye movements corresponds to the perceived pointing directions indicated by the button-presses. So, around a switch, participants predominantly move their eyes in the direction that is congruent with the newly perceived pointing direction of the triangles. We might expect such a preference for saccade directions that are *congruent* with the perceived pointing direction of the triangles based on general considerations of stimulus–response compatibility.
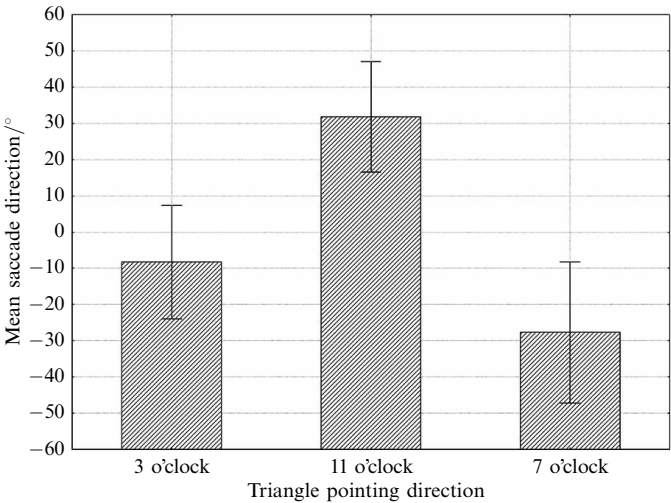
Figure 8. Means and confidence intervals of saccade directions corresponded to pointing directions of the triangles, ie certain button-presses.

3.5 *Tentative interpretation*

Given what could be concluded about the time course of individual variables, we may construe an interpretation of the switching process, based on these curves combined. Figure 9 combines the time course of the variables blink frequency, saccade frequency (Z-transformed), and number of participants with congruent saccade directions for the compatible condition. A flash animation of the eye-movement dynamics revealed in figure 9 may be found in the 'demonstrations' section of our webpage at http://pdl.brain.riken.go.jp.
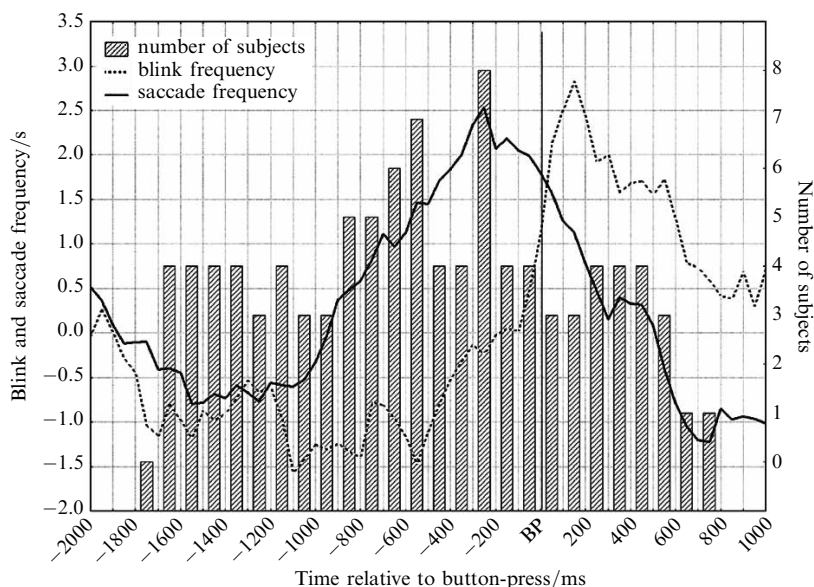
Figure 9. Time lines of saccade and blink frequencies (*Z*-transformed) and saccade directions corresponded to pointing directions of the triangles in the compatible condition. The vertical bars show the number of subjects.

Each of these curves shows an extended, slow wave of modulation. Note the negative peak in all these variables located approximately 1100 ms prior to the button-press, the gradual increase, in saccade frequency, and saccade congruence resulting in a positive peak 250 ms prior to the button-press, as well as the initial suppression of blink frequency at 1100 ms prior to the button-press, followed by a gradual increase reaching a peak after the response. Also note the clear synchronisation of saccade and blink frequency outside this interval. Correlation coefficients (CCs) between mean blink and saccade curves before, within, and after this interval were: $<-3000, -1100>$ ms, CC = 0.45 ($p = 0.04$); $<-1100, +300>$ ms, CC = 0.01 ($p = 0.96$); $<+300, +2000>$ ms , CC = 0.72 ($p = 0.002$).

The fact that blink and saccade frequencies are modulated over a long interval suggests that no unique moment could be identified as the point in time at which a switch occurs. Instead, switching an interpretation may be considered a relatively prolonged process, which involves both perceptual reorganisation and response-related process components. As blink, but not saccade, frequency appears to be time-locked to the response, the former may be more strongly linked to the perceptual reorganisation component than the latter.

Modulation of both curves start at 1100 ms prior to button-press, suggesting that this moment marks the initiation of the switch. This is also the moment where the two curves, previously closely time-locked, get desynchronised. There is a simultaneous decrease in blink frequency and an increase of saccade frequency. The decrease in blink frequency was thought to be due to response-related effort. It seems plausible to assume, therefore, that top–down processes mediate the initiation of the switch. After this moment, saccade frequency increases, and saccade directions increasingly become congruent with the newly perceived pointing direction of the triangles. These observations suggest a process of perceptual reorganisation and construal of a new interpretation. We may assume that this process terminates at or after the peak, 250 ms prior to response. In the interval $<-1100, -250>$ ms prior to the response interval, the increased frequency of the saccades may, in turn, contribute to the perceptual reorganisation.

For instance, the quick saccades result in destabilisation of sensory information in the cortex—and so facilitate switches. Saccades are associated with suppression of sensory sensitivity. The same is reported for blinks (Volkmann et al 1980; Uchikawa and Sato 1995; Ridder and Tomlinson 1997). There is evidence for a linkage between the circuitry producing saccades and the neural circuits subserving blinks (Evinger et al 1994).

Sensory suppression starts before and is maximal at the start of a saccade and outlasts saccades (Ross et al 2001). For this reason, the simultaneous occurrence of blinks and saccades is an efficient strategy for minimising visual suppression (Fogarty and Stern 1989). As reported in section 3.4.2, the zero-lag time-locking of saccades and blinks frequently observed in the literature (Zuber and Stark 1966; Rottach et al 1998) is not found here. It is, however, obtained in the first time interval, before the moment 1100 ms prior to the button-press, where we tentatively located the initiation of the switch, as well as after the switch response has been completed. During the period where we locate perceptual reorganisation, the blink frequency curve follows the saccade frequency curve with a constant time lag of 400 ms. As a result of both increased saccade frequency and the time lag with blinks, sensory suppression may be stronger than normal during the perceptual reorganisation process.

Sensory suppression increases the relative impact of mechanisms for suggesting alternative interpretations. For this reason the increased saccade frequency as well as the transient desynchronisation of saccade and blink curves may facilitate the occurrence of the switch. It was observed that the blink frequency curve was more strictly time-locked to the response than the saccade frequency curve. Consequently the saccade curve was associated with perceptual reorganisation and the blink curve with response preparation. Even if there is normally zero-lag synchronisation between the saccade and blink curves, during the switch the saccade curve temporarily precedes the blink curve. This suggests that the process component that is associated with saccade frequency temporarily exerts control. This implies that the mechanism that generates the alternative resides within the perceptual system, rather than in the system that prepares the response.

The time-lag of 400 ms implies that the perceptual reorganisation and the response preparation are cascaded stages. This is not self-evident, given that the mechanisms that initiate switching are of a top–down nature. On the other hand, cascaded processing allows for top–down modulation of perceptual reorganisation. The 400 ms lag leaves sufficient time for such influences to occur after the initial stage.

At −250 ms the perceptual reorganisation is complete and saccade frequency decreases to the level before switching. The time–blink curve follows with a constant lag. The blink curve could be associated with increased mental effort and subsequent relaxation. That mental effort seems needed to process the decision before the perceptual process is completed, suggests that the second process is engaged in actively accumulating evidence for alternative responses using information collected from the perceptual information system.

4 Conclusion

The effect reported by Ellis and Stark (1978), that fixations are longer when perceivers switch from one interpretation of an ambiguous pattern to another, could not be replicated when sampling bias was corrected for. On the contrary, increased saccade frequency was obtained during the switching process. This result still constitutes evidence for a crucial role of eye movements in perceptual switching, albeit in a different manner than previously believed. No single eye-movement indicator identifies a unique moment where the switch occurs. Instead, in using a moving-window analysis, several eye-movement indicators can be combined to obtain a detailed image of the time course of a perceptual switch as a slow process with cascaded stages.

The stages involved in switching could be labeled perceptual reorganisation and response preparation. Saccade-frequency and blink-frequency dynamics appear to be plausible indicators of these respective stages. According to these indicators, the initiation of the switch may be a top–down, response-related event; the subsequent perceptual organisation, however, proceeds largely independently of such late feedback, although the overlap between the cascaded stages allows for top–down modulation in the completion of the perceptual reorganisation.

Such an account of switching would fit what is known about brain activity in perceptual switching, as well as in eye movement. The combination of increased activity in higher-order visual areas and reduced activity in the primary visual cortex and the pulvinar (Kleinschmidt et al 1998) could be understood as the result of top–down mediated initiation of the switch, together with sensory suppression. Suppressed activity in the primary visual cortical areas as well as in the pulvinar are known to occur during saccades (Benevento and Port 1995), and their occurrence in perceptual switching may well be the result of enhanced saccade frequency. In this respect, saccades and blinks have the same effect. Usually, saccades and blinks are time-locked, which appears an elegant way to minimise sensory suppression. During the whole process that leads to the switch response, these two are no longer strictly synchronised. This will increase sensory suppression, which may in turn facilitate the switch.

Despite the fact that this account of switching was inferred from our data, and in spite of being in accordance with the available neural evidence, a direct test of our account will require more detailed study of the neurodynamics of both eye movement and switching, currently beyond the horizon of our abilities. Lacking such an opportunity to test our account, we propose it as a tentative working hypothesis, and present our eye-movement indicators for future use as behavioural indicators of perceptual switching in combination with the study of brain activity.

References

Attneave F, 1968 "Triangles as ambiguous figures" *American Journal of Psychology* **81** 447–453

Attneave F, 1971 "Multistability in perception", in *Readings from Scientific American* **63** 143–152

Baddeley A, 1999 "A crash course in stochastic geometry", in *Stochastic Geometry: Likelihood and Computation* Eds O E Barndorff-Nielsen, W S Kendall, H N N van Lieshout (London: Chapman and Hall) pp 1–35

Benevento L A, Port J D, 1995 "Single neurons with both form/color differential responses and saccade-related responses in the nonretinotopic pulvinar of the behaving macaque monkey" *Visual Neuroscience* **12** 523–544

Borsellino A, De Marco A, Allezetta A, Rinesi S, Bartolini B, 1972 "Reversal time distribution in the perception of visual ambiguous stimuli" *Kybernetik* **10** 139–144

Bucher N M, Palmer S E, 1985 "Effects of motion on perceived pointing of ambiguous triangles" *Perception & Psychophysics* **38** 227–236

Cranach M von, Schmidt R, Vogel M W, 1969 "The relationship between gaze movement and eye blink under various conditions" *Biologische Forschung* **33** 68–78 (in German)

De Jong P J, Merckelbach H, 1990 "Eyeblink frequency, rehearsal activity, and sympathetic arousal" *International Journal of Neuroscience* **51** 89–94

De Marco A, Penengo P, Trabucco A, Borsellino A, Carlini F, Riani M, Tuccio M T, 1977 "Stochastic models and fluctuations in reversal time of ambiguous figures" *Perception* **6** 645–656

Ellis S R, Stark L, 1978 "Eye movements during the viewing of Necker cubes" *Perception* **7** 575–581

Evinger C, Manning K A, Pellegrini J J, Basso M A, Powers A S, Sibony P A, 1994 "Not looking while leaping: the linkage of blinking and saccadic gaze shifts" *Experimental Brain Research* **100** 337–344

Fogarty C, Stern J A, 1989 "Eye movements and blinks: their relationship to higher cognitive processes" *International Journal of Psychophysiology* **8** 35 – 42

Gepshtein S, Kubovy M, 2000 "The emergence of visual objects in space-time" *Proceedings of the National Academy of Science of the USA* **14** 8186 – 8191

Hock H S, Schöner G, Voss A, 1997 "The influence of adaptation and stochastic fluctuations on spontaneous perceptual changes for bistable stimuli" *Perception & Psychophysics* **59** 509 – 522

Horlitz K L, O'Leary A, 1993 "Satiation or availability? Effects of attention, memory, and imagery on the perception of ambiguous figures" *Perception & Psychophysics* **53** 668 – 681

Kelso J A S, Case P, Holroyd T, Horvath E, Raczaszek J, Tuller B, Ding M, 1995 "Multistability and metastability in perceptual and brain dynamics", in *Ambiguity in Mind and Nature. Multistable Cognitive Phenomena* Eds P Kruse, M Stadler (Berlin: Springer) pp 159 – 184

Kleinschmidt A, Buchel C, Zeki S, Frackowiak R S, 1998 "Human brain activity during spontaneously reversing perception of ambiguous figures" *Proceedings of the Royal Society of London, Series B* **265** 2427 – 2433

Köhler W, 1940 *Dynamics in Psychology: Vital Applications of Gestalt Psychology* (New York: Liveright)

Köhler W, Wallach H, 1944 "Figural after-effect: An investigation of visual process" *Proceedings of the American Philosophical Society* **88** 269 – 357

Kowler E (Ed.), 1990 *Eye Movements and Their Role in Visual and Cognitive Processes* (Amsterdam: North Holland/Elsevier)

Kubovy M, Holcombe A O, Wagemans J, 1998 "On the lawfulness of grouping by proximity" *Cognitive Psychology* **35** 71 – 98

Leopold D A, Logothetis N K, 1999 "Multistable phenomena: changing views in perception" *Trends in Cognitive Sciences* **3** 254 – 264

Long G M, Toppino T C, 1994 "Adaptation effects and reversible figures: a comment on Horlitz and O'Leary" *Perception & Psychophysics* **56** 605 – 610

Ohtani Y, Ido K, Ejima Y, 1995 "Effects of luminance contrast and phase difference on motion assimilation for sinusoidal gratings" *Vision Research* **35** 2277 – 2286

Palmer S E, Bucher N M, 1981 "Configural effects in perceiving pointing of ambiguous triangles" *Journal of Experimental Psychology: Human Perception and Performance* **7** 88 – 114

Palmer S E, Bucher N M, 1982 "Textural effects in perceiving pointing of ambiguous triangles" *Journal of Experimental Psychology: Human Perception and Performance* **8** 693 – 708

Rayner K, 1998 "Eye movements in reading and information processing: 20 years of research" *Psychological Bulletin* **124** 373 – 422

Ridder W H 3rd, Tomlinson A, 1997 "A comparison of saccadic and blink suppression in normal observers" *Vision Research* **37** 3171 – 3179

Ross J, Morrone M C, Goldberg M E, Burr D C, 2001 "Changing in visual perception at the time of saccades" *Trends in Neurosciences* **24** 113 – 121

Rottach K G, Das V E, Wohlgemuth W, Zivotofsky A Z, Leigh R J, 1998 "Properties of horizontal saccades accompanied by blinks" *Journal of Neurophysiology* **79** 2895 – 2902

Sekuler A B, Palmer S E, 1992 "Perception of partly occluded objects: a microgenetic analysis" *Journal of Experimental Psychology: General* **121** 95 – 111

Stins J F, Michaels C F, 2000 "Stimulus – response compatibility for absolute and relative spatial correspondence in reaching and in button pressing" *Quarterly Journal of Experimental Psychology A* **53** 569 – 589

Uchikawa K, Sato M, 1995 "Saccadic suppression of achromatic and chromatic responses measured by increment-threshold spectral sensitivity" *Journal of the Optical Society of America A* **12** 661 – 666

Underwood G (Ed.), 1998 *Eye Guidance in Reading and Scene Perception* (Amsterdam: North Holland/Elsevier)

van Leeuwen C, Steyvers M, Nooter M, 1997 "Stability and intermittency in large-scale coupled oscillator models for perceptual segmentation" *Journal of Mathematical Psychology* **41** 319 – 344

Veltman J A, Gaillard A W, 1998 "Physiological workload reactions to increasing levels of task difficulty" *Ergonomics* **41** 656 – 669

Volkmann F C, Riggs L A, Moore R K, 1980 "Eyeblinks and visual suppression" *Science* **207** 900 – 902

Zangemeister W H, Sherman K, Stark L, 1995 "Evidence for a global scanpath strategy in viewing abstract compared with realistic images" *Neuropsychologia* **33** 1009 – 1025

Zuber B L, Stark L, 1966 "Saccadic suppression: elevation of visual threshold associated with saccadic eye movements" *Experimental Neurology* **16** 65 – 79