Perceptual asynchronies for biological and non-biological visual events

Christelle Aymoz

Faculty of Psychology and Educational Sciences
University of Geneva

Paolo Viviani

Faculty of Psychology and Educational Sciences
University of Geneva
and
Faculty of Psychology
UHSR University Milan

Correspondence to: Dr. Paolo Viviani, Faculty of Psychology and Educational Sciences
UNI MAIL 40 Boulevard du Ponts d'Arve, CH-1211 Geneva 4, Switzerland
e-mail: viviani@pse.unige.ch

Abstract

Four experiments investigated the hypothesis that different attributes of a visual scene are processed by independent channels working asynchronously. Experiments 1 and 2 considered the attributes of colour, form, and movement of simple geometrical configurations. In each of three conditions, two of these attributes switched simultaneously between two fixed values (Green/Red, Circle/Square, Fixed/Moving). Participants indicated which of the two attributes changes was closer in time to a sound signal. Response probabilities varied as a function of the time of occurrence of the sound, showing that the processing of the movement channel is delayed with respect to the other two. A smaller but significant difference was also detected between the processing times for colour and form. Comparing Experiments 1 and 2 showed that movement velocity does not affect the delay with which movement onset is perceived. Experiment 3 contrasted colour and movement in the perception of a biological movement. The stimuli were video clips of a coloured ball being lifted by a hand. The colour of the ball changed a variable amount of time before or after the ball started moving. Participants indicated which of the two changes had occurred first. We found that, unlike in Experiments 1 and 2, movement perception no longer lagged colour perception. Experiment 4 tested the hypothesis that the disappearance of the asynchrony is due to perceptual anticipation. We discuss the implications of the results *vis-à-vis* current theories on perceptual binding and on the coding of dynamic events.

Key words: Perception - Asynchrony – Biological movement – Colour – Form

Introduction

Converging evidence suggests that attributes of the visual world such as colour, shape, depth, and movement, are analysed by autonomous functional modules in the brain. For instance, Julesz's random-dot stereograms demonstrate that depth information is extracted reliably even though there is no identifiable figural detail in the scene that can be matched binocularly (Julesz, 1960). The existence of functionally autonomous modules dedicated to the analysis of specific attributes is also suggested by clinical cases in which the perception of one attribute is either selectively impaired, or selectively spared (review in Zeki, 1990; Zeki, 1991). Moreover, the modular hypothesis is consistent with anatomical and physiological evidence suggesting that different attributes are extracted in specialized and geographically distinct visual areas outside the primary visual cortex (Zeki, 1978; Livingstone & Hubel, 1988; Felleman & van Essen, 1991; Zeki, 1993; Leventhal, Thompson, Liu, Zhou, & Ault, 1995; Gegenfurtner, Kiper, & Fenstemaker, 1996; Bullier, 2001).

Recently, the modular hypothesis has been generalised by the further assumption that the processing modules are also *perceptual modules,* making their own autonomous contribution to the genesis of conscious, perceptual experience (Zeki & Bartels, 1998 a,b). If so, it is possible that these modules reach their perceptual end-point at different times. This, in fact, seems to be the case. Recent behavioural experiments suggested that simultaneous changes in the colour, form, position, and movement attributes of a stimulus are not perceived at the same time (Moutoussis & Zeki, 1997 a,b; Arnold, Clifford, & Wenderoth, 2001; Viviani & Aymoz, 2001). In all cases, the processing of movement information was found to be more time-consuming than that of either colour or form information, despite the widely held belief that the system underlying movement perception can follow more rapid events than the system underlying colour and form perception (Livingstone & Hubel, 1987). Instead, there is disagreement about the speed of colour processing relative to the processing of

geometrical attributes such as orientation, form, and position. Moutoussis and Zeki (1997 b) reported that colour outpaced orientation by as much as 63 ms. Viviani and Aymoz (2001), with a different experimental technique found no significant difference between colour and form. Pisella, Arzi, and Rossetti (1998) studying the affects of changing colour and position of the target at the onset of a manual pointing, concluded that colour processing is slower than position processing by about 80 ms. Finally, it should be noted that the temporal frequency cut-off for equiluminant chromatic change is much lower than for luminance change (Ives, 1923), suggesting that the colour system is slower than the luminance system.

In spite of its relative popularity, the modular doctrine of vision is not the only theorical framework available for integrating neurophysiological and behavioural evidence. Based on an extensive reviews of the literature, Lennie (1998) has recently challenged the modular doctrine suggesting that the various attributes of an image are not parcelled out to separate areas, but that their analysis remains intimately couples at all stages of analysis (Burr, 1999). Clearly, the presence of asynchronies in the perception of the various visual attributes is far more in keeping with the modular doctrine than with the opposite view advocated by Lennie that through all stages of analysis all dimensions of the image remain intimately couples. Therefore it seems desirable to verify that inclued perceptual asynchrony do exist.

Four experiments were conducted to investigate further the relative timing with which changes in colour, form and movement are processed.

Experiment 1

If confirmed, differences among the duty cycles of different perceptual channels would pose the challenging problem of why the corresponding asynchronies are not perceptually conspicuous under most real life conditions. Nishida and Johnston (2002), noting that the problem arises only insofar as simultaneity illusions are interpreted in terms of neural-

processing delays, argued, that such an interpretation is not logically compelling on the sole basis of experiments involving sequences of rapidly alternating stimuli (Moutoussis & Zeki, 1997 a,b). In particular, the large delay in motion perception measured with this technique may reflect the coupling of heterogeneous temporal landmarks (colour *transitions,* with movement *turning points,* Nishida & Johnston, 2000 a,b, Johnston & Nishida, 2001), rather than true duty-cycle differences. This alternative explanation was supported by two experiments (Nishida & Johnston, 2002) showing that the apparent perceptual motion delay depends on the rate of alternation, and disappears in the case of a single transition between two colours and two motion directions. This latter observation contradicts directly the results of our previous experiment (Viviani & Aymoz, 2001). Although we had eluded Nishida and Johnston's criticism by testing only single transitions, motion perception was again found to be delayed with respect to colour and form perception, as originally claimed by Moutoussis and Zeki (1997 a,b). Thus, the first goal of Experiment 1 was to confirm that, at least as far as the movement channel is concerned, differences in processing time are real, not an artefact of the measuring technique. Second, we wanted to try and adjudge the unsettled issue of whether geometrical and colour information is processed at the same speed. To pursue these two goals, we introduced a new technique for measuring the relative timing with which we perceive changes in the colour, form and movement attributes of a visual display.

Method

Participants. Twenty University of Geneva students (17 female and 3 male; age range: 18 to 29 years) volunteered for the study. All participants reported normal or corrected-to-normal acuity and had no known deficiency in colour perception. Participants were naive about the purpose of the experiment. They gave their informed consent and were paid 20 to 40 Swiss Francs, according to their performance. The Ethical Committee of the University of Geneva approved the experimental protocol.

Apparatus and Stimuli. The experiment was conducted in a quiet room kept in dim light. Participants seated at a distance of about 50 cm in front of a computer monitor (HP mod. D8901; resolution 800 × 600 pixels; vertical refresh rate: 84 Hz; CIE chromatic coordinates: Red: [x = 0.6116, y = 0.3418], Green: [x = 0.2922, y = 0.5974] and Blue: [x = 0.1456, y = 0.0688]). A graphic software (Authorware) controlled the generation of the stimuli and the recording of the responses. The stimuli were two geometrical figures: a solid square (side: 8.87 deg) and a solid circle (radius = 4.96 deg) with the same surface (at the viewing distance, 1 cm subtends approximately 1 degree of visual angle). The figures were defined only by their colour (green or red; no outlines). The stimuli were made isoluminant with the help of a DPT92 Monitor Calibrator (X-Rite Inc. Grandville MI, US). The CIE X, Y, Z coordinates for green and red were: [X = 9.86, Y = 18.94, Z = 4.32] and [X = 32.69, Y = 18.17, Z = 3.09], respectively. Stimuli were presented against a grey background (X = 28.15, Y = 30.33, Z = 44.88). A sound stimulus (1 ms square impulse with an intensity of 43 dB) was also delivered in conjunction with the presentation of the visual stimuli (see later).

In addition to the colour ([C]: Green/Red) and form ([F]: Circle/Square) attributes, the stimuli were also defined their state of motion ([M]: Fixed/Moving). In the "Fixed" modality, the figures remained at the centre of the screen. In the "Moving" modality, they moved from the centre toward the upper right corner of the screen with a constant velocity of 9.5 deg/s.

Experimental procedure and task. There were three conditions, one for each pairing of the stimuli attributes: *Form/Colour* (FC), *Colour/Movement* (CM) and *Form/Movement* (FM). In each condition, a trial comprised the following steps: a fixation point (a solid black circle, 1 deg in diameter) appeared at the centre of the screen for 1 s. After its disappearance, one of the two geometrical figures in one of the two colours (the *initial stimulus*, for example, Red Square) was displayed at the centre of the screen. After 1 s, the two attributes selected for comparison switched simultaneously to their other value (the *final stimulus*). For instance, in

condition FC a Green Square became a Red Circle, and in condition CM a Red Square turned into a Green Square, and, at the same time, started moving. The final stimulus (either still or moving) was displayed for 1 s, making the entire sequence last 2 s (Figure 1). A sound stimulus occurred before, at the same time, or after the attribute changes with an asynchrony

---*Figure 1 about here*---

(SOA) Δ varying between 0 to 300 ms in 6 steps of 50 ms. The timing of the sound relative to the attribute changes was controlled by the Authorware program with a 1 ms accuracy. By convention, the sign of the asynchrony was set as positive (negative) when the sound occurred before (after) the change. Altogether, there were 13 SOA values. After the final stimulus had disappeared, the screen was filled with the uniform background. Participants were told that the sound stimulus would occur near the transitions, but ignored that actually both stimulus attributes switched simultaneously. Their task was to indicate (forced-choice) which attribute change had occurred closer in time (either before or after) to the sound stimulus. The response was entered soon after the disappearance of the final stimulus by using three keys in the upper row of the keyboard (F5 for answer C, Esc for answer F, and F10 for answer M). A new trial started immediately after entering the response.

In condition FC there were four possible transitions between the initial and final stimulus. In conditions CM and FM there were only two relevant transitions, because we wanted the movement to start always at the centre of the screen. Thus, the movement attribute was always Fixed for the initial stimulus and Moving for the final stimulus. However, we added two additional transitions by combining also the irrelevant attribute (Form for CM, and Colour for FM) in the definition of the initial stimulus. Thus, there were 4 [transitions] × 13 [SOA] = 52 different sequences for each condition (Table 1). Each sequence was presented 10

---*Table 1 about here*---

times for a total of 560 trials (sequences with $\Delta = 0$ were presented twice). The sample size for computing individual responses was 40 for $\Delta \neq 0$ and 80 for $\Delta = 0$. Conditions were blocked and administered in separate sessions. Within each condition, the order of presentation of the sequences was randomised for each participant. The order in which conditions were tested was counterbalanced across participants. Sessions lasted approximately one hour. At the participant's request, the experiment could be interrupted for a short rest. Before each session the experimenter provided the general instructions, and demonstrated the experimental conditions of the session by 10 practice trials.

Results

The results are presented in the form of psychometric functions which, for any two attributes A_1 and A_2, describe the relationship between the SOA and the relative frequency $P(A_1)$ with which the sound stimulus was perceived closer in time to the switching of attribute A_1 than to the switching of attribute A_2 (recall that the sign of the SOA indicates whether the sound occurred before [$\Delta > 0$], after [$\Delta < 0$], or simultaneously [$\Delta = 0$] with the transition). The data points in the three panels in Figure 2 describe the psychometric functions for the

---*Figure 2 about here*---

indicated contrasts (smoothed averages over all participants). The continuous curves through the data points are the predictions of a psychophysical model to be introduced later. Perceptual biases were estimated by the response frequency at $\Delta = 0$ ($P_{\Delta=0}$, Table 2). For all contrasts $P_{\Delta=0}$ was significantly larger than .5 as demonstrated by the .99 confidence intervals (FC: [.563 - .726], CM: [.581 - .716], FM: [.609 – 772]).

-- *Table 2 about here* ---

In condition FC, response probabilities were almost symmetrical with respect to $\Delta = 0$, where P(Form) was maximum. The sound stimulus was likely to be processed faster than both colour and form changes (average processing time: 36 ms, Pöppel, 1988). Therefore, the fact

that P(Form) was significantly higher than chance level for small absolute values of the SOA implies that the interval between the perception of the sound and that of a form change was shorter than the interval between the perception of the sound and that of a colour change. In other words, form change were perceived before colour changes. P(Form) was lower than chance level when the sound stimulus was delivered either long before, or long after the attribute switching. This seemingly paradoxical inversion is well captured by the psychophysical model (see interpolating lines in Figure 2). We shall return to this specific prediction when the model is introduced.

The psychophysical functions for the two comparisons involving movement (CM and FM) were asymmetric and fairly similar. Both P(Colour) and P(Form) remained well above chance level when the sound was occurred either before, or up to 100 ms after the switching. This suggests that both colour and form changes were perceived well before movement onset. The probability of perceiving the sound closer to movement onset than to the change in colour or form is maximum for $\Delta \sim 200$ ms. Thus, the asynchrony between the movement channel and the other two channels should be roughly 200 ms

Differences among conditions were tested statistically by considering the average (non-smoothed) frequencies for all SOA values (separate one-way ANOVAs after applying the Freeman and Tukey variance-stabilising arcsin transformation, Table 3). Significant differences emerged for most negative, and for two positive values of the SOA.

-- *Table 3 about here* ---

The foregoing qualitative analysis of the psychophysical functions relative to the three comparisons between attributes confirmed that, although attributes switched simultaneously, their changes reached consciousness at different times. We used a psychophysical model to estimate quantitatively the asynchrony among the corresponding channels. The model (Figure 3) assumes that the visual attributes and the sound stimulus are processed independently. The

delays with which changes in attributes A_1 and A_2, and the sound stimulus S are perceived are random variables $\mathbf{X_1}$, $\mathbf{X_2}$ and $\mathbf{S}$ with Gaussian probability density function (pdf) d_1, d_2, and d_s, respectively. The relative position of the three pdf's on the time axis depends jointly on the SOA and on the average processing times. The differences $\mathbf{V} = \mathbf{X_1} - \mathbf{S}$ and $\mathbf{W} = \mathbf{X_2} - \mathbf{S}$ determine the response through a deterministic rule. Suppose that at least one of the two variables $\mathbf{V}$ and $\mathbf{W}$ is positive and smaller than a threshold T. Then, if $\mathbf{V} > \mathbf{W} \rightarrow$ answer "A_2"; if $\mathbf{V} < \mathbf{W} \rightarrow$ answer "A_1". The rule incorporates the intuition that no consistent response strategy exists if either 1) $\mathbf{V} > T$ and $\mathbf{W} > T$, i.e. the moments when attribute changes are perceived are too far away from the moment when the sound is perceived for a reliable discrimination of time differences, or 2) $\mathbf{V} < 0$ and $\mathbf{W} < 0$, i.e. attribute changes are both perceived *before* the sound. In the first case, we supposed a decay of the memory traces of the time of perception for the attribute changes. In the second case, the reason for not having a consistent response strategy is supposed to be the lack of the reference landmark when attribute changes are perceived. In either case, the model assumes that participants respond at random with $P(A_1) = P(A_2) = 0.5$. Figure 3C shows how this response rule partitions the (V,W) plane into domains corresponding to an answer "A_1" (marked $[X_1]$), "A_2" (marked $[X_2]$), or to a random answer (marked $[X_1/X_2]$). The density function of $\mathbf{V}$ is the convolution of the pdf's $d_s(-t)$ and $d_1(t)$. Likewise, the density function of $\mathbf{W}$ is the convolution of the pdf's $d_s(-t)$ and $d_2(t)$. The response probabilities are then computed by integrating the joint density function d(V,W) (Figure 3B) over the corresponding regions of the (V,W) plane (Figure 3C). Specifically, let $p_1 = \int_{[X_1]} d(V,W)$ and $p_2 = \int_{[X_1/X_2]} d(V,W)$, then $P(A_1) = p_1 + 0.5\ p_2$.

The model has 7 parameters, the averages (μ_1, μ_2, μ_s) and standard deviations (σ_1, σ_2, σ_s) of the pdf's, and the threshold T. However, because we were interested only in time

differences, scales were fixed by setting arbitrarily the parameters for the sound stimulus: $\mu_s =$ 100 ms, $\sigma_s = 1$ ms. Moreover, we imposed the constraint that the threshold T and the averages (μ_C, μ_F, μ_M) for attributes C, F, and M assume the same values for all three pair-wise contrasts. Instead, the standard deviations were allowed to depend on the attributes being contrasted.

The model was fitted simultaneously to the 3 [conditions] $\times$ 13 [SOA] = 39 average probabilities. Table 4 reports the least-square estimates of the model parameters calculated by a standard Simplex minimization routine. The model predictions corresponding to these

-- *Table 4 about here* ---

estimates (continuous lines in Figure 2) interpolate quite precisely the average response probabilities (even more accurate fittings were obtained by allowing the threshold and the attribute means to vary slightly across conditions). In summary, relative to the arbitrary sound reference ($\mu_s = 100$ ms), the processing of form, colour, and movement was estimated to take an additional time of 61 ms, 89 ms, and 252 ms, respectively. Therefore, the delay between colour and form was 28 ms, the delay between movement and colour was 163 ms, and the delay between movement and form was 191 ms. Note (Table 4) that in the Form/Colour contrast the colour variance was twice as large as the form variance. Thus, although the distribution for colour was delayed with respect to the distribution for form, there were instances where colour change was in fact perceived before form change. Therefore, the model (correctly) predicts the counterintuitive observation that the probability of answering "Colour" exceeds the probability of answering "Form" for large negative values of Δ (Figure 2).

Experiment 2

In Experiment 1, the velocity of the stimuli (9.5 deg/s) in the *Colour/Movement* and *Form/Movement* conditions was the same as that used in Viviani & Aymoz (2001), so as to be

able to compare the results across different tasks. It is known (Honsbein & Mateeff, 1992) that reaction times to both motion onset and offset depend on velocity. By analogy, it is possible that also the perceptual asynchronies estimated in Experiment 1 depend on stimulus velocity. To test this hypothesis, we conducted a control experiment using a different stimulusvelocity. The new velocity has been selected by taking into account the experimental conditions of Experiment 3 (see later) in which the stimuli are moved by a natural grasping movement whose approach phase, as displayed on the screen, was at 20 deg/s. By doing so, we will be able to compare also the results from this experiment with those involving the display of a biological movement.

Method

Participants. We tested again five of the twenty individuals who had participated to Experiment 1 (5 female; age range: 21 to 29 years). Participants were naïve about the purpose of the experiments. They gave their informed consent and were paid 10 to 20 Swiss Francs, according to their performance. The Ethical Committee of the University of Geneva approved the experimental protocol.

Apparatus, Stimuli, Experimental Procedure and Task. The apparatus and the general experimental conditions were the same as in Experiment 1. The only difference was the velocity of the movement (20 deg/s instead of 9.5 deg/s). We tested only the contrast *Colour/Movement* (condition CM).

Results

The results are presented in Figure 4 with the same format as Figure 2. As in Experiment 1, P(Colour) remained above chance level when the sound was occurred either before, or up to 50 ms after the switching between attributes, suggesting that colour changes

were perceived well before movement onset. Over and above the minor numerical differences between the two set of data, and the higher values of the variances, which reflect also the smaller sample size (5 vs 20 participants), it is clear that doubling the velocity of the stimuli

------------------------------------- *Figure 4 about here* ---

did not alter significantly the asynchrony between colour and movement. This is also demonstrated by the interpolation of the data points by the model. For this fitting we fixed all the parameters of the model to the same values obtained in Experiment 1, with the exception of the variances for colour and movement. The new optimal values of the variances were very close to those already computed before, and the interpolation was equally satisfactory.

Individual perceptual biases were again estimated by the response frequency at $\Delta = 0$ (Table 2). Because the participants to this experiment also served for the previous one, statistical analysis was conducted on the individual values of $P_{\Delta=0}$. The difference between conditions was not significant (t-test for paired samples: t = -0.253, p=0.812). The coefficient of linear correlation between individual values was fairly high (r=.692), indicating a good degree of consistency across experiments.

Experiment 3

The perceptual modularity hypothesis is intimately related to the so-called *binding problem*. If indeed the various attributes of a visual event were processed independently, it would seem necessary to posit a mechanism for tagging the outputs of the different channels, so that, ultimately, they are all attached to one and the same event. Treisman and co-workers (Treisman & Gelade, 1980; Kahneman & Treisman, 1984; Kahneman, Treisman, & Gibbs, 1992) have suggested that visual attention plays this binding role. Specifically, when a visual event draws attention, a selective enhancement of the event's features would be followed by an integration of the features into a unified representation. The nature of such an hypothetical integration mechanism, however, remains elusive. In particular, the evidence summarized

above that feature processing can be asynchronous is damaging for the only detailed account of this mechanism put forward so far, namely, that integration is achieved by the temporal synchronization of neural activities that, in different networks, signal the processing of different attributes of the same event (von der Maslburg & Schneider, 1986; Singer & Gray, 1993).

Zeki and his co-workers (Zeki & Bartels, 1998 a,b; Aleksander & Dunmall, 2000) have attempted to bypass the difficulty by taking issue with the very idea that visual consciousness arises from feature binding (Crick & Koch, 1990; Dennett & Kinsbourne, 1992). Instead, according to Zeki's version of the modular hypothesis, each specialised feature-processing system yields its computational output at a specific node of the visual network. The activity at each node is perceptually explicit, giving rise to a separate state of micro-consciousness. Thus, rather than preceding or facilitating visual consciousness, binding actually brings together the activity of multiple micro-consciousnesses, and it is no longer necessary to posit a final stage of processing where conscious visual perception is represented. Despite this claim, however, it is not obvious that Zeki's view of perceptual modularity does actually reconcile the possibility that visual attributes are processed asynchronously with the fact we have a coherent perception of visual events.

The nature of the binding mechanism might become more understandable if we were able to demonstrate that differences in processing rates are contingent upon some specific characteristics of the visual event. There are reasons to suspect that one such characteristics is the fact that the event involves a human gesture. Indeed, biological movements, *qua* visual stimuli, have peculiar properties. Following Johansson (Johansson, 1973; Johansson, 1976; Johansson, 1977) several authors have shown that bodily movements of the body have a remarkable perceptual saliency. Even sketchy descriptions of the joint positions afford a vivid image of the actual movement identity of the agent (Kozlowski & Cutting, 1977; Cutting,

Proffitt, & Kozlowski, 1978), to extract the symbolic meaning of a gesture (Poizner, Bellugi, & Lutes-Driscoll, 1981), and even to discriminate his own walking pattern from that of others (Beardworth & Buckner, 1981). Even facial expressions can be perceived from the movements of a few point-light (Bassili,1978). Because a preference for biological movements emerges as early as 3 to 5 months of age (Fox & McDaniel, 1982; Bertenthal, Proffitt, & Cutting, 1984), this sharp perceptual tuning is likely to reflect inborn properties of the underlying neural machinery (Oram & Perrett, 1994).

At the functional level, several studies have suggested that the saliency of biological movements, as well as a number of perceptual effects, originate from the interplay between sensory data and the implicit competence shared by all humans about the peculiar characteristics of their movements (for a review, see Viviani, 2002). At the neurophysiological level, the hypothesis that motor competence enters into the perception of dynamic events involving voluntary action is supported by the recent discovery of neurons in the monkey premotor cortex – the so-called "mirror neurons" that respond both when the animal performs a specific manual action, and it watches the same action performed by an external agent (Di Pellegrino, Fadiga, Fogassi, Gallese & Rizzolatti, 1992; Fadiga, Fogassi, Gallese, & Rizzolatti, 2000).

Based on this evidence for a special perceptual status of human gestures, Experiment 3 was designed to test the hypothesis that the considerable lag with which the movement aspect is perceived relative to form and colour disappears when the motion results overtly from the intervention of a human agent.

Method

Participants. Twenty University of Geneva students (13 female and 7 male; age range: 21 to 31 years) volunteered for the study. None of them had served for the first experiment. All participants reported normal or corrected-to-normal acuity and had no known deficiency in

colour perception. Participants were naive about the purpose of the experiment. They gave their informed consent, and were paid 20 to 40 Swiss Francs according to their performance. The Ethical Committee of the University of Geneva approved the experimental protocol.

Apparatus and Stimulus. The apparatus and the general experimental conditions were the same as in Experiment 1. We recorded several video clips of a hand gesture with a digital camera (Sony DCR-TRV900E). At the beginning of the clip (Figure 5), the hand was laying palm-down on a horizontal surface, at about 10 cm to the right of a ball placed at the centre of the scene. Then, the hand moved forward, seized the ball with the natural five-finger grip, and lifted it outside the scene. The entire sequence lasted 3 s. The size of the ball on the screen was the same as that of the circular stimulus in Experiment 1 (radius = 4.96 cm). All spatial references were suppressed by covering the surface with dark uniform fabric (CIE coordinates: X = 2.00, Y = 1.98, Z = 3.19). After visual inspection, we selected two clips with

-- *Figure 5 about here* --

the criterion that 1) the ball remained clearly visible throughout the sequence; 2) the approach phase of the hand movement was similar, but not identical, in the two clips; 3) the trajectory of the lifting phase followed as closely as possible the diagonal line directed to the upper right corner of the screen, as the moving stimuli in Experiment 1. The average velocity of the hand during the lifting was approximately 20 deg/s. By editing the video clip with a standard software (Adobe Premiere), we could change the hue of the ball while leaving the rest of the scene unchanged (note that changing the hue did not alter the original shading of the ball's surface). From each selected clip, we generated 36 stimulus sequences. In half of them, the ball was red (CIE coordinates X = 25.42, Y = 14.76, Z = 4.08) at the beginning of the clip, and became green (CIE coordinates X = 7.87, Y = 14.95, Z = 3.24) in the course of the sequence (because the shading of the ball's surface was not uniform, the indicated CIE coordinates are averages of a number of measures performed with the monitor calibrator). In

the other half, the colour switched in the opposite direction. In both cases, the ball remained globally isoluminant (14.76 vs 14.95). Each sequence within the two subgroups of 18 was characterized by the rank order K of the frame in which the colour switched with respect to frame in which the ball began to move. In 9 sequences, the colour switched before (K = +8, +7, … +1) or at (K = 0) movement onset. In the remaining 9 the switch occurred after (K = -1, -2, … -8) or at (K = 0) movement onset. At the recording speed of 25 frames/s, the asynchrony Δ between the switching of the Colour (C: Green/Red) and Movement (M: Fixed/Moving) attributes ranged between –320 ms and +320 ms in steps of 40 ms. By convention, the sign of the SOA was set as negative when the ball left the ground before the change of colour and positive in the complementary case. By crossing the two original clips

-- *Table 5 about here* --

with all combinations of attributes and SOA, we obtained 72 different stimuli (Table 5).

Experimental conditions, procedure and task. Aside from the nature of the stimuli, and the fact that we tested only the Colour/Movement contrast, the procedure was exactly as in the experiment reported previously (Viviani & Aymoz, 2001). Briefly, trials began with a fixation point (a solid black circle 1 deg in diameter) lasting 1 s at the centre of a uniform gray background (CIE coordinates X = 28.15, Y = 30.33, Z = 44.88). Immediately afterward, one video clip was displayed at the centre of the screen. At the end, the screen was filled with the background and remained so until the participant initiated a new trial by entering the response. The task was to indicate (forced-choice) which attribute had changed first, i.e. whether the colour of the ball had changed before (answer C), or after (answer M) the ball had started moving. Responses were entered by using two keys in the upper row of the keyboard (Esc for M and F12 for C). Each of the 72 different sequences was presented 10 times for a total of 720 trials. The selection of the sequence was randomised for each participant with the constraint that the same sequence was never presented twice in successive trials. Sequences

with $\Delta = 0$ occurred twice. Thus, the sample size for computing individual response frequencies was 40 for $\Delta \neq 0$, and 80 for $\Delta = 0$. An experimental session lasted approximately one hour, including a brief rest period that participants could allow themselves. Each session was preceded by a period of adaptation to the dim ambient light, a verbal description of the task, and a familiarisation phase of 10 trials.

Results

The results are presented in the form of a psychometric function describing the relationship between the SOA (Δ), and the relative frequency P(Colour) with which colour switching was perceived before movement onset. The middle panel in Figure 6 shows the psychometric function obtained by averaging the relative frequencies over all participants.

-------------------------------------- *Figure 6 about here* ---

The z-transform of the raw (un-smoothed) response frequencies could not be fit rather adequately by a linear regression, indicating that the psychometric function deviated significantly from a cumulative Gaussian distribution (lower panel of Figure 6). At the population level, there was no obvious perceptual bias in favour of either attribute (median = -0.024) and a fairly low differential threshold (JND = 0.123).

For all participants, the response probabilities P(Colour) increased monotonously toward as a function of the SOA, levelling at a value close to 1. Individual differences were estimated by three parameters of the psychometric function, namely the response frequency at $\Delta = 0$ ($P_{\Delta=0}$), the SOA for which responses are at chance level ($\Delta_{p=0.50}$), and the Just-Noticeable-Difference JND = ($\Delta_{p=0.75}$ - $\Delta_{p=0.25}$) /2. The values $\Delta_{p=0.50}$, $\Delta_{p=0.75}$, and $\Delta_{p=0.25}$ were computed by a spline interpolation of the un-smoothed response frequencies. There was some variability in $P_{\Delta=0}$, which was based on just one measure (Table 6). By contrast, individual estimates of $\Delta_{p=0.50}$ and of the JND were quite homogeneous, and their population averages

were very close to the corresponding values computed from the average psychometric function. Moreover, $\Delta_{p=0.50}$ did not differ significantly from 0 (two-tailed t-test, t = 0.479).

-- *Table 6 about here* --

We estimated the distribution of the duration of the perceptual processing for movement with the same method devised in our previous study (Viviani & Aymoz, 2001). Briefly, the analysis is based on the assumption that colour and movement changes are processed independently. The processing times are described by random variables, $\mathbf{t}_C$ and $\mathbf{t}_M$, with probability density functions (pdf) d_C and d_M (upper panel of Figure 7). The response is supposed to be dictated by the difference $\delta = \mathbf{t}_M - \mathbf{t}_C$ through the deterministic rule: if $\delta > 0 \rightarrow$ Answer C, if $\delta < 0 \rightarrow$ Answer M. The pdf d_δ of δ is the convolution of $d_C(\text{-}t)$ and $d_M(t)$, and

-- *Figure 7 about here* --

the effect of varying the SOA is simply to shift d_δ along the time axis (lower left panel of Figure 7). Given the assumed response rule, the psychometric function is the cumulative distribution of δ. Its shape and position along the SOA axis depends only on the variances (σ_C, σ_M) of d_C and d_M, and on the difference of their means $\mu_\delta = \mu_M - \mu_C$ (lower right panel of Figure 7). Because we were interested only in the relative duration of the perceptual processing, we assumed $\mathbf{t}_C$ to have a Gaussian pdf with $\mu_C = 0$ and $\sigma_C = 30$ ms. The pdf of $\mathbf{t}_M$ was then computed by estimating d_δ as the derivative of the psychometric function with respect to Δ, and solving numerically the convolution equation $d_\delta(t) = d_C(\text{-}t) * d_M(t)$. The result (upper panel of Figure 6) shows that also d_M is very nearly Gaussian. The estimated mean μ_M is almost 0. Thus, the processing times for colour and movement are indeed almost identical as suggested also by the data of Table 6.

Experiment 4

The disappearance of the colour-movement latency in Experiment 3 could be due to the fact that the approach phase of the hand movement provides anticipatory cues about the

moment when the ball is grasped and lifted. In other words, the delay with which movement onset is perceived would be offset by estimating the time-to-contact before the contact actually takes place. Thus, before claiming any role for the natural character of the dynamic event in suppressing the perceptual asynchrony, it was necessary to demonstrate that anticipatory clues *per se* do not affect movement perception. To this end, we designed a control experiment in which the movement of the ball is prompted by the impact of a non-biological stimulus whose motion provides the same timing cues as the approaching hand.

Method

Participants. Ten University of Geneva students (3 male and 7 female; age range: 23 to 30 years) who had already participated to Experiment 3 served also for this experiment. They were paid 20 to 40 Swiss Francs, depending on the performance. The Ethical Committee of the University of Geneva approved the experimental protocol.

Apparatus and Stimuli. The apparatus and the general experimental conditions were the same as in Experiment 3. A graphic software (Authorware) controlled the generation of the stimuli and the recording of the responses. There were three stimuli. The first two were a solid square (side: 8.87 deg) and a solid circle (radius: 4.96 deg) with the same surface. These stimuli, defined only by their colour (green or red), were made isoluminant with the help of a DPT92 Monitor Calibrator (X-Rite Inc. Grandville MI, US). Their CIE X, Y, Z coordinates were respectively (X = 9.86, Y = 18.94, Z = 4.32) and (X = 32.69, Y = 18.17, Z = 3.09). Stimuli were presented against a grey background (X = 28.15, Y = 30.33, Z = 44.88). At the beginning of a trial, one of the two figures was displayed at the centre of the screen. The third stimulus, a solid black circle (CIE coordinates: X = 2.00, Y = 1.98, Z = 3.19) was also presented in conjunction with the central geometrical figure (see later). At the onset of a trial, the black circle moved with a straight trajectory from the upper left corner of the screen toward the geometrical figure with a constant velocity of 20 deg/s.

Experimental procedure and task. Only the *Colour/Movement* (CM) condition was tested. Trials began with the display of a central fixation point (a solid black circle, 1 cm in diameter) lasting 1s. The moving black circle stopped as soon as it touched the central stimulus (one of the two geometrical figures in one of the two colours, for example, Red Square). At the same time, one of the attribute (colour or movement) switched between its two possible values. Then, the second attribute (movement or colour) also switched with an asynchrony Δ (SOA) varying between 0 ms to 320 ms in steps of 40 ms. By convention, the sign of the asynchrony interval Δ was set as positive when colour changed before movement and negative in the opposite case. In all cases, the central stimulus started moving toward the upper right corner with the same constant velocity as the black circle (20 deg/s). The entire sequence lasted 3 s. Thereafter, the screen was filled with the background, and remained so until the participant initiated a new trial by entering the response. As in Michotte's displays, the scene elicited the very vivid perceptual illusion of a mechanical interaction between two real objects. The task was to indicate (forced choice) which attribute changed first by using two keys in the upper row of the keyboard (F5 for answer C and F12 for answer M).

There were 4 [transitions] × 17 [SOA] = 72 different sequences (Table 7). Each sequence was presented 10 times for a total of 720 trials (sequences with Δ = 0 were presented twice). Thus, as in Experiment 3, the sample size for computing individual response frequencies was 40 for Δ ≠ 0 and 80 for Δ = 0. Sessions lasted approximately 30 min and were preceded by a period of adaptation to the dim ambient light, a verbal description of the task, and a familiarisation phase of 10 practice trials.

Results

The results are presented in Figure 8 with the same format of Figure 6. As in Experiment 3, the probability of answering "Colour" increased monotonously when the SOA varied from −320 ms to +320 ms (lower panel). The z-transform of the raw (un-smoothed)

response frequencies show that the psychometric function deviated significantly from a cumulative normal distribution (bottom panel). More importantly, the psychometric function was clearly asymmetric with respect to $\Delta = 0$ ($P_{\Delta=0} = 0.75$), indicating that colour changes are processed faster than movement onset. The model introduced in Experiment 3 was fitted to the data points (continuous line) yielding an estimated asynchrony between the two processes of about 45 ms (upper panel in Figure 8). The results demonstrate that, in this condition, movement onset is perceived with a delay with respect to colour change, as it was the case in Experiments 1 and 2. The asynchrony was smaller than that observed before, but was quite similar to the one estimated in Viviani and Aymoz (2001).

-- *Figure 8 about here* --

The comparison between the two experimental conditions was carried out also at the individual level by considering the three parameters of the psychometric function defined above, namely the response frequency at $\Delta = 0$ ($P_{\Delta=0}$), the SOA for which responses are at chance level ($\Delta_{p=0.50}$), and the Just-Noticeable-Difference JND = ($\Delta_{p=0.75} - \Delta_{p=0.25}$) /2. Table 6 reports for comparison the values of these parameters for the individuals who participated in both Experiment 3 and 4. In all cases there was a significant difference between the corresponding parameters in the two conditions (t-test for paired samples; $P_{\Delta=0}$: t = -3.750, p = .005; $\Delta_{p=0.50}$: t = 3.270, p = .014; JND: t = 2.376, p = .041).

General Discussion

Experiments 1 and 2 confirmed an earlier suggestion that changes in the colour, form and movement attributes of a visual configuration are processed at different rates. As in previous experiments, the movement channel was found to be slower than both the colour and form channel, the delay being of about 160 ms and 190 ms, respectively. Moreover, form processing outpaced colour processing by about 28 ms. The only difference between

Experiment 1 and 2 was the velocity of the moving stimuli. The fact that the results in both conditions were well interpolated by the model, with essentially the same parameters, demonstrated that velocity is not a critical factor for determining the relative asynchronies. However, movement delays were considerably longer than those reported recently both by ourselves (Viviani & Aymoz, 2001), and others (Moutoussis and Zeki, 1997 a,b; Arnold, et al., 2001). A likely explanation for the increased difference is that, unlike previous experiments, participants had to process almost simultaneously an acoustic and a visual stimulus. Thus, the click sound that provided the reference for judging the relative timing of the attribute changes may well have interfered with the processing of the visual information.

As regards the Form/Colour contrast, the estimated delay is close to the minimum threshold for temporal discrimination (Hirsch & Sherrick, 1961; von Steinbüchel, Wittman, & Pöppel, 1996). However, there is a clear disagreement between our finding that form changes reach consciousness *before* colour changes, and the results of Moutoussis and Zeki (1997 b) indicating an opposite delay of 63 ms in favour of colour. A direct comparison between the two studies may not be entirely warranted because we tested actual shape changes, whereas Moutoussis and Zeki only tested changes in edge orientation. A full geometrical shape turning into another shape generates a richer array of changes than just an edge switching from one orientation to another. This, in turn, might result into a faster access to consciousness, either because individual changes reinforce mutually, or because some changes are processed faster than others, and the perception of just one of them is sufficient to signal the switching between forms. Note that, although the Form-Colour delay measured in our previous experiment (10 ms) was too small to reach statistical significance, it was in the same direction of the one measured here. More importantly, a greater saliency of form over colour changes emerged clearly, quite independently of the specific model we used to estimate the mean processing time. In fact, when the sound occurred at the same time of the colour and form

changes ($\Delta = 0$), the probability of indicating the latter as being closest to the sound was significantly higher than .5 (Table 2). Note also that the results obtained with our new technique are consistent with impulse-response (Burr and Morrone, 1993), simultaneity judgments (Bowen, 1981), and evoked potential (Fiorentini, Burr, and Morrone, 1991) results indicating that the channel involved in detecting luminance transients has a larger bandwidth than the channel involved in chromatic transients.

Both by Moutoussis and Zeki (1997 a) and Nishida and Johnston (2002) used as stimuli sequences of alternations between colour and movement direction, varying the phase between alternations until one has the impression that both attributes change simultaneously. Phase differences may be interpreted as processing delays, and this is indeed what Moutoussis and Zeki (1997 a) did. However, a different account of the same data is also available (Johnston & Nishida, 2001). The neural processing of the attribute changes takes a finite amount of time. Moreover, the transition from one colour to another is a first-order temporal change, whereas the transition between opposite movement directions (turning points) is a second-order temporal change. Thus, accurate estimates of subjective simultaneity would require pairing events of different type. It could be, however, that observers erroneously link colour changes with transition in the motion sequence, i.e. with events of the same type. In this case, the phase difference that nulls the subjective delay would simply reflect this erroneous marker correspondence, rather than a true neural delay. Thus, results obtained with sequences of transitions are bound to be ambiguous. Such an ambiguity could not have arisen in Experiments 1 and 2 where - as in our previous study (Viviani & Aymoz, 2001) - trials involved single transitions between the two attributes of the stimuli, not sequences of transitions. This seems to bar the possibility of explaining the observed asynchronies as the result of an erroneous temporal judgment. Indeed, Nishida and Johnston (2002) designed their Experiment 2 precisely for the purpose of demonstrating that no delay is present in the single

transition condition, and, therefore, that phase differences measured with alternating stimuli should not be interpreted as processing delays. The question remains of why their results for the Colour/Movement contrast (Nishida & Johnston, 2002, Figure 2B), are at variance with ours, and with similar results obtained by Arnold with the after-effect technique (Arnold et al., 2001). The crucial difference might be that both in our Experiments 1 and 2, and in our previous study (Viviani & Aymoz, 2001) there were no turning points in the movement, just a sudden transition from a steady position into uniform motion. Actually, the movement delay seems to depend on the angular difference between directions at the turning point (Arnold & Clifford, 2001). Be as it may, in the absence of turning points, our observers could not have made the "biased cross-attribute linkage of time markers of the same temporal type" that, according to Nishida and Johnston (2002, p. 360), has been mistakenly interpreted as a processing delay.

Taken together, the results of the three contrasts tested in Experiments 1 and 2 confirmed that, when the stimuli are abstract geometrical forms, and the motion has no obvious cause, the outcome of the colour, form and movement channels reach consciousness at different times. This conclusion provides further support to the modular doctrine of vision. By contrast, the very presence of asynchronies in the perception of different attributes of the image is clearly inconsistent with the alternative view advocated recently by Lennie (1998) that most single neurons in the visual brain represent a value on a multidimensional vector that describes simultaneously all aspects of the image locally.

Experiment 3 demonstrated instead that the large asynchrony between colour and movement disappears when the stimulus is part of a naturalistic scene, and the motion is the consequence of a human gesture. A direct comparison is possible with the results of the Colour/Movement contrast in Viviani and Aymoz (2001) because the task was the same. The velocity of the moving stimuli was not the same as in our previous study. However,

comparing Experiment 1 and 2 showed that this experimental parameter does not have a significant influence on the asynchronies. Thus, the only major difference between the two studies was the intervention of a human hand performing the lifting action.

One may attempt to explain why the moving hand suppresses perceptual asynchrony by invoking visual factors only. Whereas the display provides no clue about the colour switching time, the moment when the ball is going to move is anticipated by watching the hand during the approach phase. Thus, expectation of the impending movement may replace the actual stimulus in generating the corresponding time marker. If so, the absence of subjective asynchrony would imply that expectation sets the movement time marker about 60 ms earlier than the end of the duty cycle of the movement channel that processes the true onset of the movement (recall that in our previous experiment colour outpaced movement by the 60 ms). Alternatively, expectation may have a priming affect, by enhancing the dynamics of the movement channel. The control Experiment 4 in which the hand was replaced by an artificial agent dispelled both hypotheses. Indeed, all sources of purely visual information that might have affected the asynchrony in Experiment 3 were present also in Experiment 4. Yet, virtually the same asynchrony reported previously reappeared when the motion was not generated by a biological agent.

A different explanation of the role of the moving hand calls into play the interaction between visual and motor factors. As noted before, neurons in the monkey ventral premotor cortex are activated both when it performs a hand gesture, and when it watches the same action being performed by someone else. It has been proposed (Gallese, Fadiga, Fogassi, & Rizzolatti, 2002) that these neurons are part of a network, which includes also area 7b of the inferior parietal lobule, matching action observation and execution. A number of studies (Grèzes, Costes, & Decety, 1998; Hari, Forss, Avikainen, Kirverskari, Salenius, & Rizzolatti, 1998) strongly suggest that such a "mirror matching system" exists also in humans, where its

primary function would be to allow us to understand the behaviour of others. It can then be supposed that hand gestures perceived visually are interpreted by activating covertly the same motor structures involved in their actual execution.

We argued (Viviani, 2002) that such a recognition-by-resonance mechanism, which is reminiscent of the one invoked in Liberman's theory of speech perception (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967), is implicated in the motor-perceptual interactions that occur when watching natural gestures. In particular, it has been shown that the kinematics of these gestures is perceived (Viviani & Stucchi, 1992), reproduced (Viviani, Campadelli, & Mounoud, 1987; Viviani, Baud-Bovy, & Redolfi, 1997), and anticipated (Orliaguet, Kandel, & Böe, 1997, and Kandel, Orliaguet, & Viviani, 2000) better than that of non-biological motions. The naturalistic scene displayed in Experiment 3 – unlike the artificial stimuli used both in our previous experiment (Viviani & Aymoz, 2001), and in Experiment 4 might have activated the mirror matching system. If so, one can speculate that the suppression of the movement asynchrony is one by-product of this activation. To the extent that the mirror matching system is instrumental for generating a reliable representation of a human action, it may provide a binding mechanism whereby all attributes of the relevant events in the visual scene are set into strict time register before reaching consciousness. It should be stressed that we are not suggesting that such a temporal binding is vital for coherent perception. Many visual scenes that do not involve human gestures, and for which the mirror matching system is presumably inactive, are nevertheless perceived in what appears to be a unitary fashion. In such cases, differential delays of the order of tens of milliseconds among the various visual channels, which are detected experimentally, may well go unnoticed in everyday life. What we are suggesting, however, is that a synchronization of the various visual channels is an ingredient of the high perceptual tuning that we seem to have for human gestures.

Further research is needed to validate the hypothesis that motor factors are instrumental for suppressing the delay with which movement onset is perceived. However, the very fact that the asynchrony is present in some conditions (Experiments 1, 2 and 3), and compensated in others (Experiment 4) indicates that, whatever the underlying process, binding is not a mandatory operation. Perhaps, an adequate characterization of this process is provided by the notion of a flexible buffer receiving the output of specialized visual channels, and maintaining active the incoming information until the various pieces of information have coalesced. The degree of synchronization afforded by the buffer need not to be strict. It may actually be rather loose - of the order of a few tens of milliseconds - unless the stimuli have a peculiar salience, or the focussing of attention forces a stricter temporal alignment.

References

Aleksander, I. & Dunmall, B. (2000). An extension to the hypothesis of the asynchrony of visual consciousness. *Proceeding of the Royal Society of London, Serie B, 267,* 197-200.

Arnold, D.H., & Clifford, C.W.G. (2002). Determinants of asynchronous processing in vision. *Proceeding of the Royal Society of London, Serie B,* 268, 7381-7385.

Arnold, D.H., Clifford, C.W.G., & Wenderoth P. (2001). Asynchronous processing in vision: Color leads motion. *Current Biology, 11,* 596-600.

Bassili, J.N. (1978). Facial motion in the perception of faces and of emotional expressions. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 373-379.

Beardworth, T. & Buckner, T. (1981). The ability to recognize oneself from a video recording of one's movements without seeing one's body. *Bulletin of the Psychonomic Society, 18 (1),* 19-22.

Bertenthal, B. I., Proffitt, D. R., & Cutting, J. E. (1984). Infant sensitivity to figural coherence in biomechanical motions. *Journal of Experimental Child Psychology, 37,* 213-230.

Bowen, J.J. (1981). Latencies for chromatic and achromatic visual mechanisms. *Vision Research, 21,* 1457-1466.

Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews, 36,* 96-107.

Burr, D:C: (1999). Vision: Modular analysis -or not?. *Current Biology, 9*, R90-R92.

Burr, D. C. & Morrone, M. C. (1993). Impulse-response functions for chromatic and achromatic stimuli. *Journal of the Optical Society of America, 8,* 1706-1713.

Crick, F. & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences, 2,* 263-275.

Cutting, J. E., Proffitt, D. R., & Kozlowski, L. T. (1978). A biomechanical invariant for gait perception. *Journal of Experimental Psychology: Human Perception and Performance, 4 (3),* 357-372.

Dennett, D. C. & Kinsbourne, M. (1992). Time and the observer: The where and when of consciousness in the brain. *Behavioral and Brain Science, 15,* 183-247.

Di Pellegrino, G., Fadiga, L., Fogassi, V., Gallese, V. & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research, 91,* 176-180.

Fadiga, L., Fogassi, L., Gallese, G., & Rizzolatti, G. (2000). Visuomotor neurons: Ambiguity of the discharge or "motor" perception? *International Journal of Psychophysiology, 35,* 165-177.

Felleman, D.J. & van Essen, D.C. (1991). Distribued hierarchical processing in primate cerebral cortex. *Cerebral Cortex, 1,* 1-47.

Fiorentini, A., Burr, D.C., & Morrone, C. (1991). Spatial and temporal characteristics of colour vision. In: A. Valberg and B.B. Lee (Eds) *From pigment to perception: Advances in understanding visual processing* (pp. 139-150). New York: Plenum.

Fox, R. & McDaniel, C. (1982). The perception of biological motion by human infants. *Science, 218,* 486-487.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. In W.Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention and performance* (pp. 247-266). Oxford: Oxford University Press.

Gegenfurtner, K. R., Kiper, D. C., & Fenstemaker, S. B. (1996). Processing of color, form, and motion in macaque area V2. *Visual Neuroscience, 13,* 161-172.

Grèzes, J., Costes, N., & Decety, J. (1998). Top-down effect of strategy on the perception of human biological motion: A PET investigation. *Cognitive Neuropsychology, 15,* 553-582.

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academic Science of the USA, 95,* 15061-15065.

Hirsh, I.J. & Sherrick, C.E. (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology, 62,* 423-432.

Hohnsbein, S. & Mateeff, S. (1992). The relation between the velocity of visula motion and the reaction time to motion onset and offset. *Vision Research, 32,* 1789-1791.

Ives, H.E. (1923). A chart of the flicker photometer. Journal of the Optical Society of America, Rev. Sci. Instrum.*7,* 363-365.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics, 14 (2),* 201-211.

Johansson, G. (1976). Spatio-temporal differenciation and integration in visual motion perception. *Psychological Research, 38,* 379-393.

Johansson, G. (1977). Studies on visual perception of locomotion. *Perception, 6,* 365-376.

Johnston, A. & Nishida, S. (2001). Time perception: Brain time or event time ? *Current Biology, 11,* 427-430.

Julesz, B. (1960). Binocular depth perception of computer-generated patterns. *Bell System Technical Journal, 39,* 1125-1162.

Kahneman, P. & Treisman, A. (1984). Changing view of attention and automaticity. In R.Parasuraman, R. Davies, & J. Beatty (Eds.), *Varieties of attention.* London: Academic Press.

Kahneman, P., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24,* 175-219.

Kandel, S., Orliaguet, J-P., & Viviani, P. (2000). Perceptual anticipation in handwriting: The role of implicit motor competence. *Perception & Psychophysics, 62,* 706-716.

Kozlowski, L. T. & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics, 21,* 575-580.

Lennie, P. (1998). Single units and visual cortical organization. *Perception, 27,* 889-935.

Leventhal, A. G., Thompson, K. G., Liu, D., Zhou, Y., & Ault, S. J. (1995). Concomitant sensitivity to orientation, direction, and color of cells in layers 2, 3, and 4 of monkey striate cortex. *The Journal of Neuroscience, 15,* 1808-1818.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of speech code. *Psychological Review, 74,* 431-461.

Livingstone, M. S. & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *The Journal of Neuroscience, 7,* 3416-3468.

Livingstone, M. S. & Hubel, D. H. (1988). Segregation of Form, Color, Movement, and Depth: Anatomy, Physiology, and Perception. *Science, 240,* 740-749.

Maslburg van der, C. & Schneider, W. (1986). A neural cocktail-party processor. *Biological Cybernetics, 54*, 29-40.

Moutoussis, K. & Zeki, S. (1997a). A direct demonstration of perceptual asynchrony in vision. *Proceeding of the Royal Society of London, Serie B, 264,* 393-399.

Moutoussis, K. & Zeki, S. (1997b). Functional segregation and temporal hirerarchy of the visual perceptive systems. *Proceeding of the Royal Society of London, Serie B, 264,* 1407-1414.

Nishida, S. & Johnston, A. (2000a). Dissociation of perceived temporal synchrony and response latency for changes in colour and motion. *Investigative Ophtalmology and Visual Science, 41*: S712.

Nishida, S. & Johnston, A. (2000b). Perceptual delay for rapid direction alternations: A new account in terms of the dichotomy of first-order and second-order temporal changes. *Perception, 29 (Suppl)*: 27.

Nishida, S. & Johnston, A. (2002). Marker correspondence not processing latency determines temporal binding of visual attributes. *Current Biology, 12,* 359-368.

Oram, M. W. & Perrett, D. I. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to "biological motion" stimuli. *Journal of Cognitive Neuroscience, 6,* 99-116.

Orliaguet, J-P., Kandel, S., & Böe, L. J. (1997). Visual perception of cursive handwriting: Influence of spatial and kinematic information on the anticipation of forth coming letters. *Perception, 26,* 905-9012.

Pisella, L., Arzi, M., & Rossetti, Y. (1998). The timing of color and location processing in the motor context. *Experimental Brain Research, 121,* 270-276.

Poizner, H., Bellugi, U., & Lutes-Driscoll, V. (1981). Perception of american sign language in dynamic point-light displays. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 430-440.

Pöppel, E. (1988). Mind works. Time and Conscious experience. NewYork: Harcourt Brace Jovanovich.

Singer, W. & Gray, C. M. (1993). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience, 18,* 555-586.

Steinbüchel von, N., Wittmann, M. & Pöppel, E. (1996). in M.A. Pastor & J. Artieda (Eds.). *Time, internal clocks, and movements* (pp 281-304). Amsterdam: Elsevier.

Treisman, A. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97-136.

Viviani, P. (2002). Motor competence in the perception of dynamic events. In W.Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention and performance* (pp 406-442). Oxford: Oxford University Press.

Viviani, P. & Aymoz, C. (2001). Colour, form, and movement are not perceived simultaneously. *Vision Research, 41,* 2909-2918.

Viviani, P., Baud-Bovy, G., & Redolfi, M. (1997). Perceiving and tracking kinesthetic stimuli: Further evidence of motor-perceptual interactions. *Journal of Experimental Psychology: Human Perception and Performance, 23, 4,* 1232-1252.

Viviani, P., Campadelli, P., & Mounoud, P. (1987). Visuo-manual pursuit tracking of human two-dimensional movements. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 62-78.

Viviani, P. & Stucchi, N. (1992). Biological movements look uniform: Evidence of motor-perceptual interactions. *Journal of Experimental Psychology: Human Perception and Performance, 18 (3),* 603-623.

Zeki, S. (1978). Functional specialisation in the visual cortex of the rhesus monkey. *Nature, 274,* 423-428.

Zeki, S. (1990). A century of cerebral achromatopsia. *Brain, 113,* 1721-1777.

Zeki, S. (1991). Cerebral akinetopsia visual motion blindness. *Brain, 114,* 811-824.

Zeki, S. (1993). *A vision of the brain.* London: Blackwell.

Zeki, S. & Bartels, A. (1998a). The asynchrony of consciousness. *Proceeding of the Royal Society of London, Serie B, 265,* 1583-1585.

Zeki, S. & Bartels, A. (1998b). The autonomy of the visual systems and the modularity of conscious vision. *Philosophical Transactions of the Royal Society of London, Serie B, 353,* 1911-1914.

Table 1

Experiment 1: Sequences of attribute changes in the three experimental conditions

FC (Form/Colour)	CM (Colour/Movement)	FM (Form/Movement)
Δ = -300, -250, -200, -150, -100, -50, 0 (ms)		
1 SR*F* → CG*F*	*S*RF → *S*GM	S*R*F → C*R*M
2 SG*F* → CR*F*	*S*GF → *S*RM	S*G*F → C*G*M
3 CR*F* → SG*F*	*C*RF → *C*GM	C*R*F → S*R*M
4 CG*F* → SR*F*	*C*GF → *C*RM	C*G*F → S*G*M
Δ = 300, 250, 200, 150, -100, 50, 0 (ms)		
5 SR*F* → CG*F*	*S*RF → *S*GM	S*R*F → C*R*M
6 SG*F* → CR*F*	*S*GF → *S*RM	S*G*F → C*G*M
7 CR*F* → SG*F*	*C*RF → *C*GM	C*R*F → S*R*M
8 CG*F* → SR*F*	*C*GF → *C*RM	C*G*F → S*G*M

Note- C: Circle, S: Square, G: Green, R: Red, F: Fixed, M: Moving. Values of irrelevant attributes in each sequence are set in italic. Each sequence in the two blocks was paired with the indicated values of the SOA (Δ) with which the click sound occurred relative to the attribute changes. By convention, SOAs are negative (positive) in the first (second) block.

Table 2

Experiments1 and 2 : Perceptual bias for each participant

	$P_{\Delta=0}$			
	Experiment 1			**Experiment 2**
	FC	CM	FM	CM
	P(Form)	P(Colour)	P(Form)	P(Colour)
Participant				
1	0.873	0.831	0.874	
2	0.584	0.668	0.597	
3	0.861	0.545	0.490	
4	0.541	0.665	0.620	
5	0.476	0.581	0.611	
6	0.850	0.797	0.653	
7	0.707	0.614	0.722	
8	0.893	0.722	0.933	0.587
9	0.788	0.842	0.838	0.987
10	0.588	0.531	0.622	
11	0.464	0.604	0.708	
12	0.514	0.639	0.680	
13	0.717	0.556	0.580	
14	0.542	0.432	0.422	0.462
15	0.607	0.925	0.969	
16	0.496	0.535	0.591	
17	0.504	0.815	0.602	
18	0.527	0.528	0.927	
19	0.504	0.575	0.506	0.475
20	0.873	0.584	0.861	0.987
Average	0.645	0.649	0.691	0.699

Note - $P_{\Delta=0}$: Response probability when sound and attribute changes occur simultaneously.

Table 3

Experiment 1 : Summary of the analysis of Variance for each SOA

SOA	F	df	MS	p
-300	13.763	2	0.033	<.001
-250	14.009	2	0.032	<.001
-200	24.408	2	0.028	<.001
-150	13.494	2	0.013	<.001
-100	8.613	2	0.006	<.001
-50	1.146	2	0.0076	.329
0	1.347	2	0.0013	.272
50	0.368	2	0.0004	.695
100	0.498	2	0.0005	.612
150	14.003	2	0.061	<.001
200	12.954	2	0.046	<.001
250	1.839	2	0.0045	0.173
300	0.988	2	0.0033	0.382

Table 4

Experiments 1 and 2 : Best fitting parameters of the model

Condition	T	μ_C	σ_C	μ_F	σ_F	μ_M	σ_M (ms)
Experiment 1							
Form/Colour	300	189	257	161	126	×	×
Colour/Movement	300	189	212	×	×	352	107
Form/Movement	300	×	×	161	201	352	129
Average			234		163		118
Experiment 2							
Colour/Movement	300	189	229	×	×	352	97

Table 5

Experiment 2: Sequences of attributes changes

CM (Colour/Movement)

Δ = -320, -280, -240, -200, -160, -120, -80, -40, 0 (ms)

1	$C_1\text{RF} \rightarrow C_1\text{GF} \rightarrow C_1\text{GM}$
2	$C_1\text{GF} \rightarrow C_1\text{RF} \rightarrow C_1\text{RM}$
3	$C_2\text{RF} \rightarrow C_2\text{GF} \rightarrow C_2\text{GM}$
4	$C_2\text{GF} \rightarrow C_2\text{RF} \rightarrow C_2\text{RM}$

Δ = 320, 280, 240, 200, 160, 120, 80, 40, 0 (ms)

5	$C_1\text{RF} \rightarrow C_1\text{GF} \rightarrow C_1\text{GM}$
6	$C_1\text{GF} \rightarrow C_1\text{RF} \rightarrow C_1\text{RM}$
7	$C_2\text{RF} \rightarrow C_2\text{GF} \rightarrow C_2\text{GM}$
8	$C_2\text{GF} \rightarrow C_2\text{RF} \rightarrow C_2\text{RM}$

Note – C_1: video Clip 1, C_2: video Clip 2, G: Green, R: Red, F: Fixed, M: Moving. Values of the irrelevant attribute in each transition are set in italic. Each sequence in the two blocks was paired with the indicated nine values of the SOA (Δ) between the first and the second attribute change. By convention SOAs are negative (positive) in the first (second) block.

Table 6

Experiments 3 and 4 : Perceptual bias, constant error and differential limen (JND) for

each participant

	Experiment 3			Experiment 4		
Participant	$P_{\Delta=0}$	$\Delta_{P=0.5}$	JND	$P_{\Delta=0}$	$\Delta_{P=0.5}$	JND
1	0.419	0.012	0.155			
2	0.338	0.014	0.085	0.652	-0.010	0.064
3	0.944	-0.083	0.121	0.717	-0.029	0.122
4	0.728	-0.052	0.214			
5	0.692	-0.019	0.096	0.980	-0.080	0.071
6	0.301	0.023	0.112	0.674	-0.029	0.155
7	0.197	0.063	0.194	0.863	-0.029	0.060
8	0.707	-0.015	0.069	0.769	-0.023	0.076
9	0.601	-0.013	0.126	0.945	-0.052	0.065
10	0.509	-0.001	0.175	0.905	-0.053	0.090
11	0.507	-0.001	0.163			
12	0.283	0.032	0.154			
13	0.148	0.064	0.156			
14	0.717	-0.038	0.163			
15	0.132	0.057	0.126			
16	0.563	-0.010	0.163	0.844	-0.058	0.131
17	0.724	-0.050	0.206			
18	0.641	-0.027	0.178			
19	0.463	0.003	0.106			
20	0.730	-0.044	0.179	0.977	-0.095	0.084
Average	0.519	-0.003	0.178	0.817	-0.046	0.118

Note - $P_{\Delta=0}$: Frequency of answer « Colour » when colour and movement attributes changed simultaneously ; $\Delta_{P=0.5}$: Asynchrony for which responses were at level chance; JND: Differential limen estimated by the semi-interquartile range.

Table 7

Experiment 4: Sequences of attributes changes

CM (Colour/Movement)	
1	CRF $\rightarrow$ CRM $\rightarrow$ CGM
2	CGF $\rightarrow$ CGM $\rightarrow$ CRM
3	SRF $\rightarrow$ SRM $\rightarrow$ SGM
4	SGF $\rightarrow$ SGM $\rightarrow$ SRM
Δ = 320, 280, 240, 200, 160, 120, 80,-40, 0 (ms)	
5	CRF $\rightarrow$ CGF $\rightarrow$ CGM
6	CGF $\rightarrow$ CRF $\rightarrow$ CRM
7	SRF $\rightarrow$ SGF $\rightarrow$ SGM
8	SGF $\rightarrow$ SRF $\rightarrow$ SRM

Note − C: Circle, S: Square, G: Green; R: Red, F: Fixed, M: Moving. Values of irrelevant attributes in each sequence are set in italic. Each sequence in the two blocks was paired with the indicated values of the SOA (Δ) between the first and the second attribute change, By convention SOAs are negative (positive) when movement onset preceeds the change of colour.

Figure legends

Figure 1. Experiment 1: Timing of the events within trials. Stimuli were defined by two attributes A_1 and A_2, each of which could take two values. The initial stimulus (in the example A_{12}-A_{22} lasted 1 s, and then changed into the final stimulus (A_{11}-A_{21}), which lasted again 1 s. A 1ms click sound (width not in scale) occurred either shortly before, or shortly after the change. The asynchrony (SOA: Δ) ranged from -300 ms to 300 ms in 50 ms steps. Although both attributes always changed simultaneously, participants had to indicate (forced choice) which attribute change was closer in time to the click sound.

Figure 2. Experiment 1: Response frequencies as a function of the asynchrony (SOA) between the attribute changes and the click sound (psychometric functions). Data pooled over all participants for the indicated pairings of the attributes. Average frequencies computed from 20 [participants] $\times$ 40 [repetitions] = 800 responses for $\Delta \neq 0$ and 20 [participants] $\times$ 80 [repetitions] = 1600 responses for $\Delta = 0$. Bars around data points are the .99 confidence intervals of the mean (exact binomial theory). The continuous lines through the data points are predictions of the model used to translate response frequencies into delay estimates (see Figure 3).

Figure 3. Experiment 1: Model for translating response frequencies into estimates of the processing delays. A: Both the transition between the two values of the attributes (stimuli onset), and the sound click are perceived after completing a processing requiring a random amount of time ($\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{S}$). In this example the sound precedes the attribute changes ($\Delta <$ 0). The probability density functions (pdf) of the processing times (d_1, d_2, and d_S) are all Gaussian. The mean processing time for the sound (μ_S) is supposed to be shorter than that for both visual attributes (μ_1 and μ_2). Participants are asked to compare the differences $\mathbf{V} = \mathbf{X}_1$ - $\mathbf{S}$

and $\mathbf{W} = \mathbf{X_2} - \mathbf{S}$. The pdf's of $\mathbf{V}$ and $\mathbf{W}$ are the convolution of $d_s(-t)$ and $d_1(t)$ and the convolution of $d_s(-t)$ and $d_2(t)$, respectively. B: The joint pdf $d(V,W)$ of $\mathbf{V}$ and $\mathbf{W}$. C: The response rule maps into a partition of the (V,W) plane. The probability of answer A_1 (A_2) is the integral of $d(V,W)$ over the regions marked $[X_1]$ ($[X_2]$) plus half the integral over the regions marked $[X_1/X_2]$.

Figure 4. Experiment 2. Results for a display in which the velocity of the moving stimulus was 20 deg/s rather than 9.5 deg/s as in Experiment 1. Response probabilities are virtually unaffected by this increase in velocity. Same format as in Figure 2.

Figure 5. Experiment 3. Upper panels: four frames from one of the video clips used as stimuli. Actual clips lasted 3 s. Lower panel: timing of the events within trials. F (Fixed) and M (Moving) were the two values of the attribute Movement. G (Green) and R (Red) were the two values of the attribute Colour. In this example the initial, intermediate, and final values of the stimulus were (RF), (GF), and (GM) respectively. The asynchrony (Δ) between attribute changes ranged between -320 ms and 320 ms in 40 ms steps. Participants had to indicate which attribute had changed first.
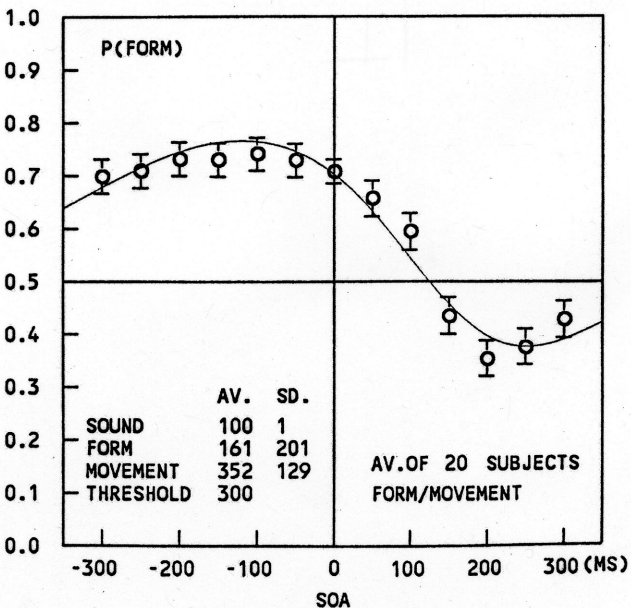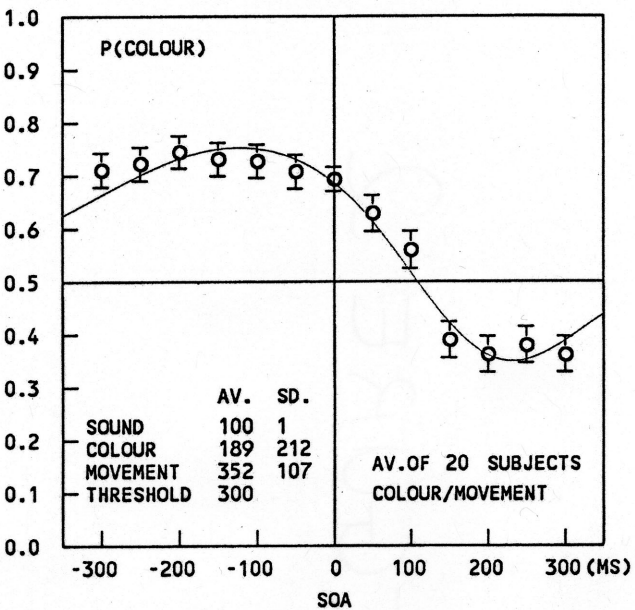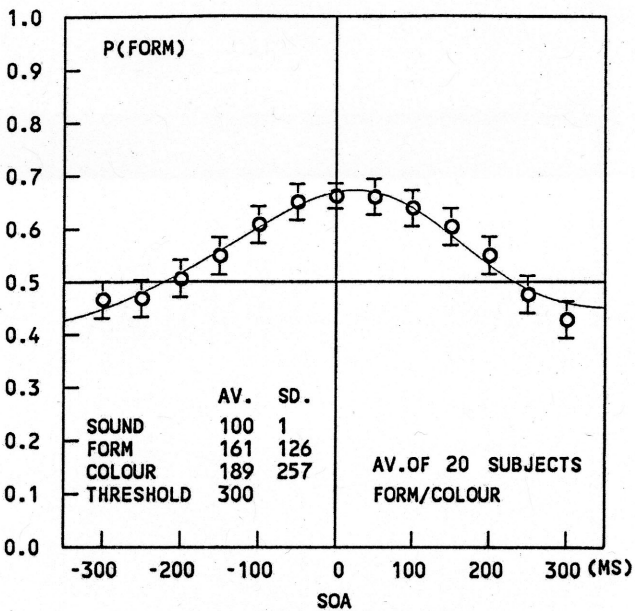
Figure 6. Experiment 3. Upper panel: probability density function (pdf) of the total processing times for Colour and Movement, and pdf of the difference between the total processing times (d_δ). The distributions were computed as described in the text. Middle panel: Response frequencies as a function of the asynchrony (SOA) between the change of colour and movement attributes(psychometric function). Data pooled over all participants. Bars around data points are the 0.99 confidence intervals (exact binomial model). The continuous line through the data point is a gaussian fit. Lower panel: Z-transform of the psychometric
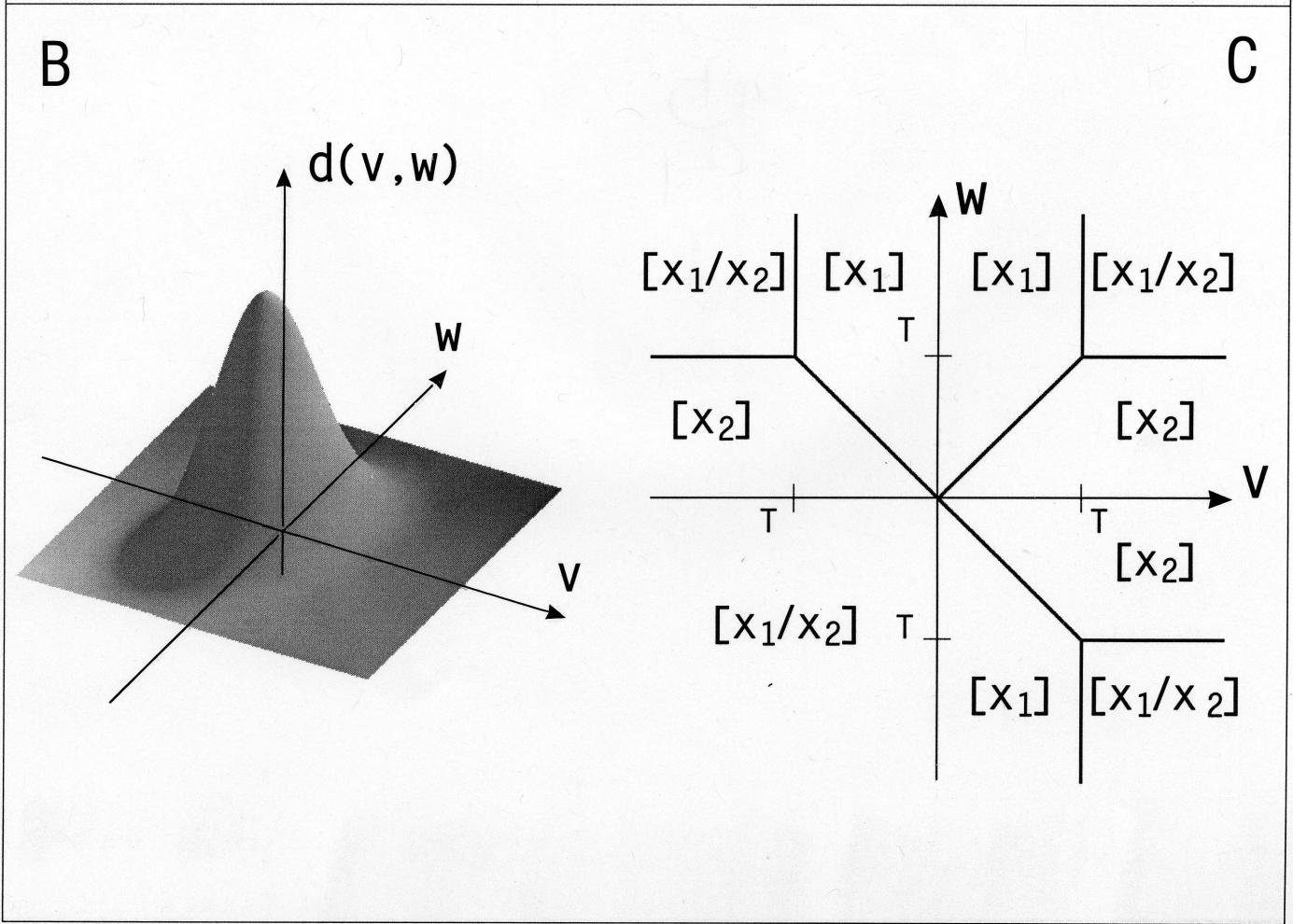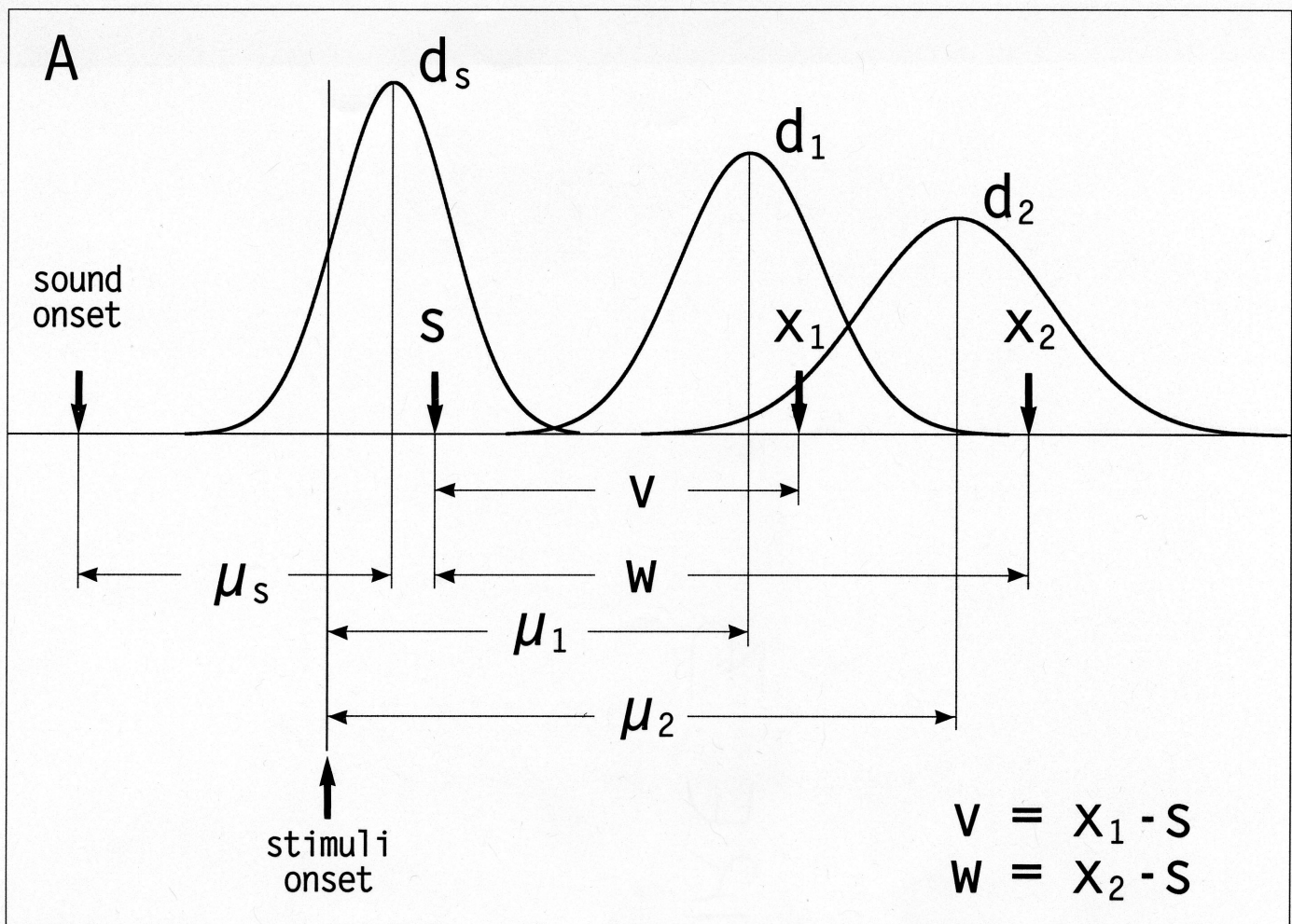
function. The raw (un-smoothed) data points are interpolated by a linear regression (heavy line and associated 0.95 confidence parabolae).
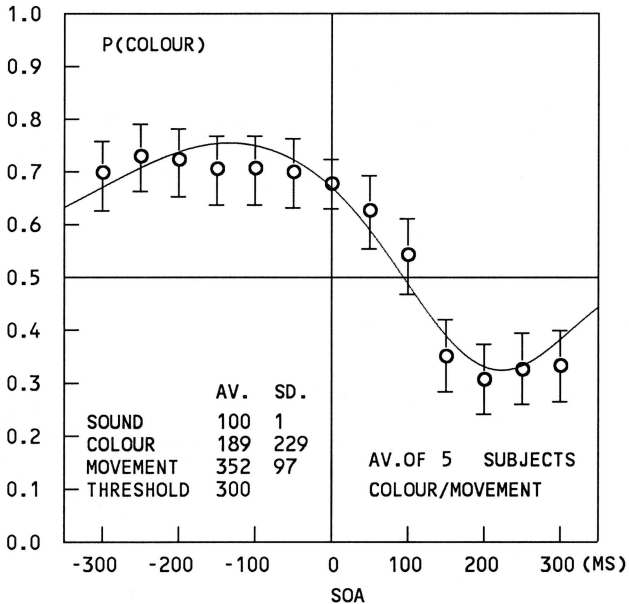

Figure 7. Experiment 3: Scheme for transforming the response frequencies into an estimate of the processing times. Upper panel: the transition between the two possible values of the attributes is perceived only after completing a processing requiring a random amount of time. The probability density functions (pdf, d_C and d_M) of the processing times may be different. In this example the mean processing time is longer for Movement than for Colour , ($\mu_M > \mu_C$), and Movement changes after Colour (SOA > 0). Lower left panel: distributions (d_δ) of the difference between the processing time for Movement and Colour for three values of SOA. Participants answer "C" ("M") when the difference is positive (negative). Because $\mu_M > \mu_C$, the probability of answer "C" is greater than that of answer "M" even at SOA = 0. Lower right panel: Psychometric functions relating the probability of answer "C" as a function of SOA for three values of the mean difference $\mu_\delta = \mu_M - \mu_C$.

Figure 8. Experiment 4. Results for a dynamic display in which the motion and the changes in colour of the central stimulus is prompted by the contact with a moving geometrical figures rather than the contact with a real hand as in Experiment 3. Same format as in Figure 6.
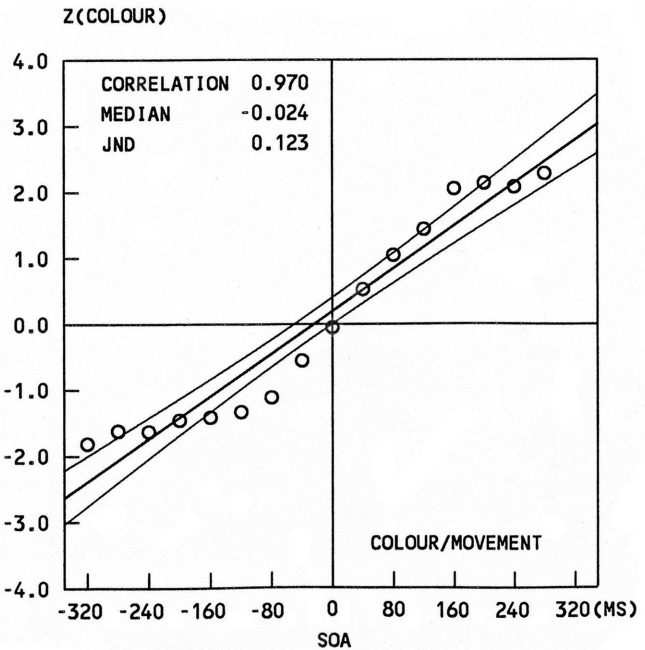
FIXATION POINT

ATTRIBUTE A$_1$

ATTRIBUTE A$_2$

SOUND

RESPONSE

A_{12}
A_{11}
A_{22}
A_{21}

Δ

1s 1s 1s RT

Top graph:

P(FORM)

	AV.	SD.
SOUND	100	1
FORM	161	126
COLOUR	189	257
THRESHOLD	300	

AV.OF 20 SUBJECTS
FORM/COLOUR

SOA

Middle graph:

P(COLOUR)

	AV.	SD.
SOUND	100	1
COLOUR	189	212
MOVEMENT	352	107
THRESHOLD	300	

AV.OF 20 SUBJECTS
COLOUR/MOVEMENT

SOA

Bottom graph:

P(FORM)

	AV.	SD.
SOUND	100	1
FORM	161	201
MOVEMENT	352	129
THRESHOLD	300	

AV.OF 20 SUBJECTS
FORM/MOVEMENT

SOA

A

sound onset

d_s

S

d_1

x_1

d_2

x_2

μ_s

v

w

μ_1

μ_2

stimuli onset

$v = X_1 - S$
$w = X_2 - S$

B

$d(v,w)$

w

v

C

w

$[x_1/x_2]$ $[x_1]$ $[x_1]$ $[x_1/x_2]$
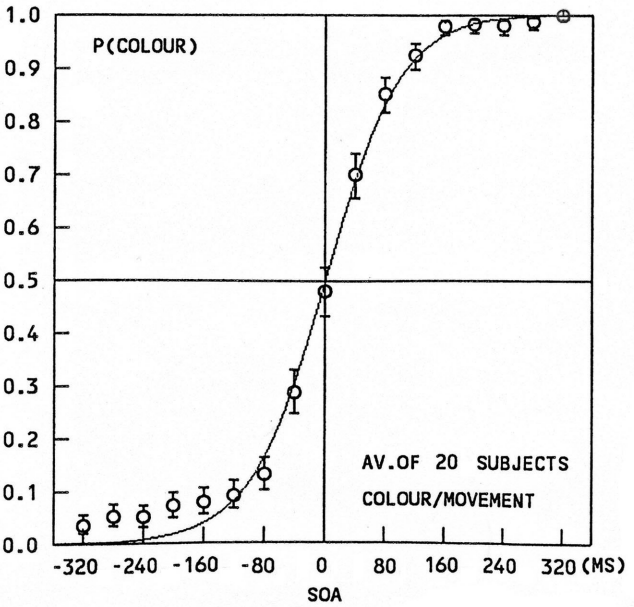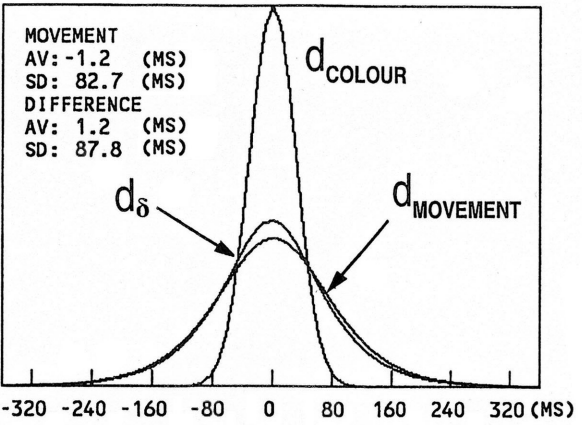
$[x_2]$ $[x_2]$

v
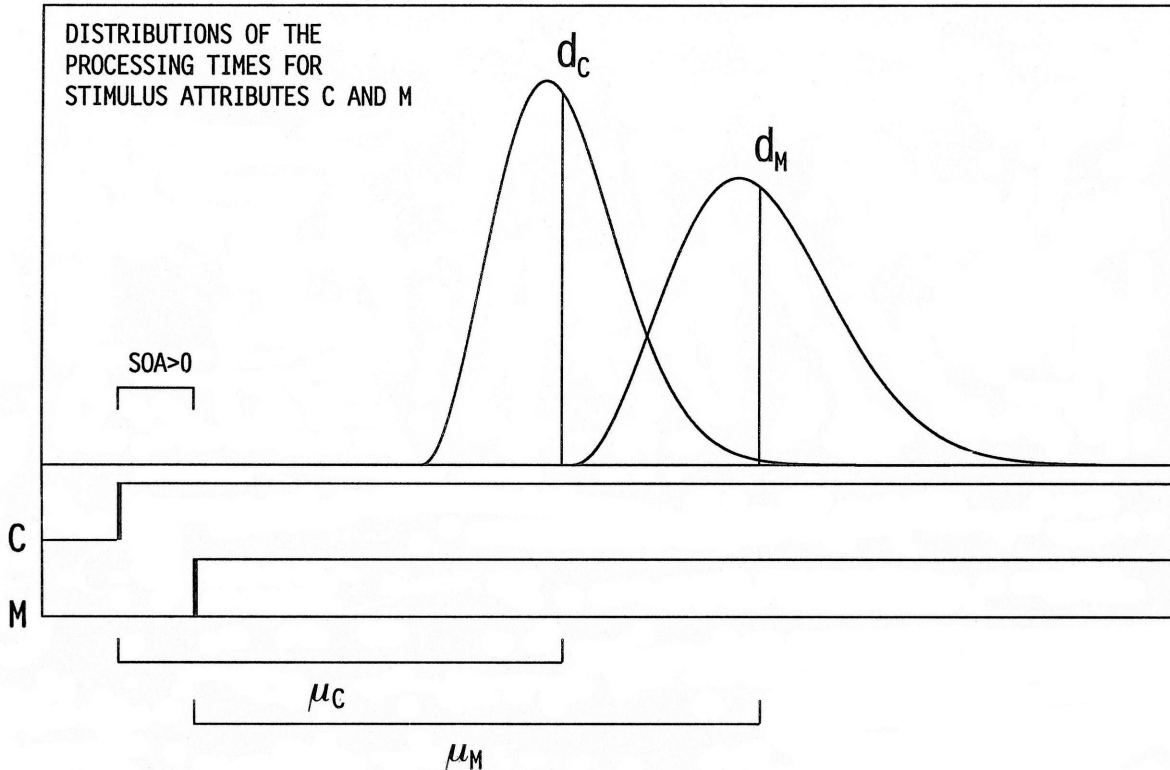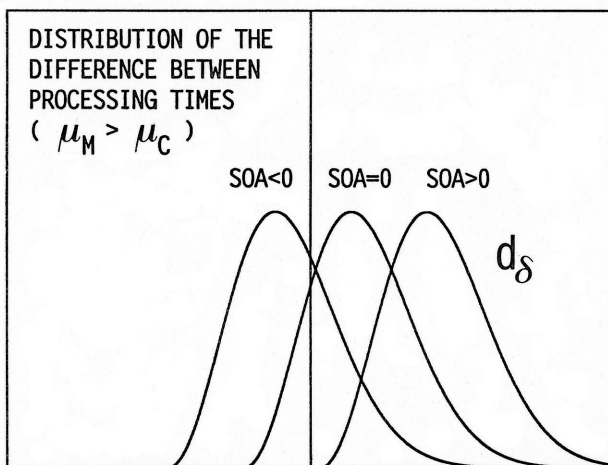
$[x_2]$

$[x_1/x_2]$

$[x_1]$ $[x_1/x_2]$

DISTRIBUTIONS OF THE
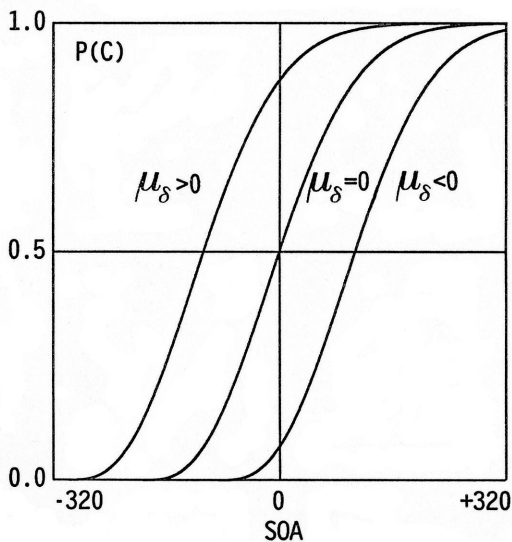PROCESSING TIMES FOR
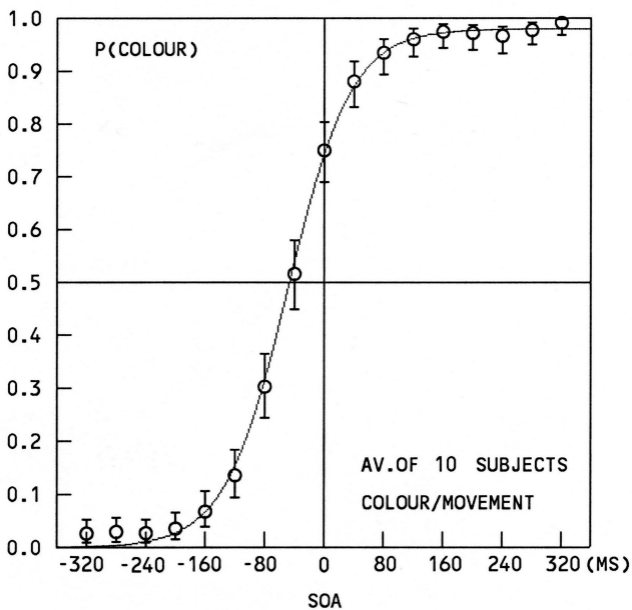STIMULUS ATTRIBUTES C AND M

d_C
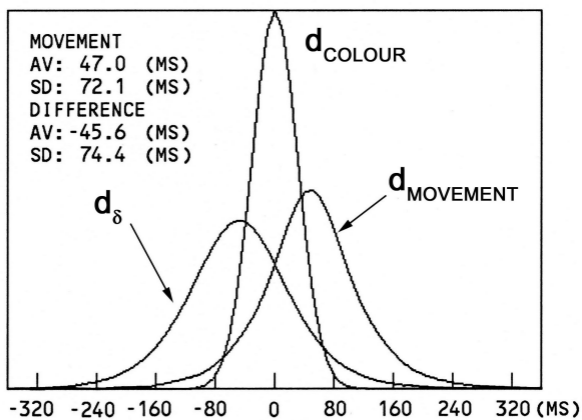
d_M

SOA>0

C

M

μ_C

μ_M

DISTRIBUTION OF THE
DIFFERENCE BETWEEN
PROCESSING TIMES
($\mu_M > \mu_C$)

SOA<0   SOA=0   SOA>0

d_δ

ANSWER: M <-- 0 --> ANSWER: C

1.0   P(C)

$\mu_\delta > 0$ $\mu_\delta = 0$ $\mu_\delta < 0$

0.5

0.0

-320   0   +320
SOA

Top panel:

MOVEMENT
AV: 47.0 (MS)
SD: 72.1 (MS)
DIFFERENCE
AV: -45.6 (MS)
SD: 74.4 (MS)

d_{COLOUR}

$d_{MOVEMENT}$

d_δ

-320 -240 -160 -80 0 80 160 240 320 (MS)

Middle panel:

P(COLOUR)

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0

AV.OF 10 SUBJECTS
COLOUR/MOVEMENT

-320 -240 -160 -80 0 80 160 240 320 (MS)

SOA

Bottom panel:

Z(COLOUR)

4.0
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
-4.0

CORRELATION 0.968
MEDIAN   -0.034
JND   0.124

COLOUR/MOVEMENT

-320 -240 -160 -80 0 80 160 240 320 (MS)

SOA