PERGAMON

Vision Research 43 (2003) 531–548

Vision Research

www.elsevier.com/locate/visres

The dynamics of bi-stable alternation in ambiguous motion displays: a fresh look at plaids

Jean-Michel Hupé *, Nava Rubin

Center for Neural Science, New York University, 4 Washington Place, New York, NY 10003, USA

Received 10 January 2002; received in revised form 21 August 2002

Abstract

Prolonged observations of moving plaids lead to bi-stable alternations between coherency and transparency. However, most studies of plaids used brief presentations and a 2AFC between the two interpretations, thus overlooking the dynamical aspect of plaid perception. In other domains, most notably binocular rivalry, it was shown that the dynamics of the bi-stable alternations reveal important insights about the underlying mechanisms. Here we develop methods to study the dynamics of plaid perception. Observers continually indicated their percept (coherency or transparency) during presentations that lasted 1–5 min. Two measures of the relative strength of the coherency percept were derived from those data: $C/[C+T]$, the relative time spent seeing coherency, and RTtransp, the response time to report transparency. Those measures are independent of each other yet tightly correlated, and both show systematic relations to manipulations of plaid parameters. Furthermore, the two measures are sensitive to manipulations in wide parametric regimes, including ranges where brief-presentation methods suffer from "ceiling" and "floor" effects. We conclude that studying the dynamics of bi-stability in plaids can provide new and unsuspected findings about motion integration and segmentation.

Keywords: Dynamics; Bi-stability; Motion segmentation; Integration

1. Introduction

A central problem to vision processing is how the brain computes a global percept from many isolated local cues. In motion processing, a popular illustration of this problem is the aperture problem: when a moving straight line is viewed through an aperture so that its endpoints are not visible, only the component of the motion perpendicular to the line's orientation can be observed (Wallach, 1935; English translation in Wuerger, Shapley, & Rubin, 1996). Marr and Ullman (1981) noted that the brain is constantly faced with the aperture problem, because of the small receptive field sizes of neurons in early visual cortex. The resolution of the ambiguity inherent in local motion measurements requires a global process. Global motion computation involves two fundamental processes: integration and segmentation. In real world scenes, the visual system is faced with multiple, often overlapping objects that can move in different directions, leading to a complex array of local motion measurements. Thus, on the one hand there is a need to combine, or *integrate* local motion signals that arise from the same object, while on the other hand it is necessary to *segment* motion cues that arise from different objects (Braddick, 1993). A classic stimulus that illustrates those conflicting demands is the plaid (Adelson & Movshon, 1982; Wallach, 1935, 1976). A moving plaid can be seen either as a single object moving rigidly ("coherent motion") or as two gratings sliding over each other ("transparent motion"). In the first interpretation, the integration process is dominant, while in the second interpretation the motion segmentation process is stronger and the grating components of the plaid are segmented from each other. Plaids have been a particularly useful tool to study the mechanisms of motion integration and segmentation, since observers' tendency to perceive coherency versus transparency can be manipulated systematically through many parameters, such as the angle between the gratings, the

* Corresponding author. Centre de Recherche Cerveau et Cognition, CNRS-UPS UMR 5549, Université Paul Sabatier, 133, route de Narbonne, 31062 Toulouse Cedex, France. Tel.: +33-5-62-17-28-06; fax: +33-5-62-17-28-09.

E-mail addresses: jean-michel.hupe@cerco.ups-tlse.fr (J.-M. Hupé), nava@cns.nyu.edu (N. Rubin).

0042-6989/02/$ - see front matter
PII: S0042-6989(02)00593-X

spatial frequency or the speed (Adelson & Movshon, 1982; Movshon, Adelson, Gizzi, & Newsome, 1985).

Under prolonged observation, the perception of plaid stimuli switches back and forth between the coherent and the transparent interpretation—it is bi-stable (Wallach, 1935). Such bi-stability is observed even when one of the two percepts is strongly dominant in short observations. [1] Perhaps because a forced-choice judgment of "coherent" or "transparent" is difficult for long presentation times, most studies of plaids used brief presentations (typically 1 or 2 s; but see von Grunau & Dubé, 1993). However, reliable methods to study bi-stable percepts in prolonged presentations have been developed in other domains, such as the Necker cube, figure/ground ambiguous stimuli (e.g. Rubin, 1921; English translation in Rubin, 1958), and, most notably, binocular rivalry (see Blake, 1989, 2001; Blake & Logothetis, 2002; Lehky, 1988; Leopold & Logothetis, 1999; Levelt, 1968 for reviews). In those domains, researchers studied the dynamics of perceptual alternations by asking observers to continually report which of the two (or more) possible interpretations they are perceiving at every moment; each trial lasted between dozens of seconds and a few minutes. This method has been developed most extensively in studies of binocular rivalry, where various measures based on the continual-report data were shown to have reliable relations to parametric manipulations of the stimulus.

In the study reported here, we develop methods similar to those used in binocular rivalry to study the dynamics of perceptual alternations in plaid stimuli. We assess the methodological validity of this approach, and use it to study motion integration and segmentation in plaids. A possible concern about the dynamics approach is that motion perception is very sensitive to adaptation processes. It has been reported that the perception of a plaid as coherent or transparent can depend on previous exposure to motion stimuli (Movshon et al., 1985). This might lead one to suspect that experimental methods using long presentation times could be more susceptible to adaptation processes than brief-duration 2AFC methods. We therefore decided to address this issue first. In a preliminary experiment, we examined the durations spent perceiving coherency and transparency over very long observation times (5 min), and found that there were no grounds for concerns about adaptation (see Section 2). Based on these encouraging results, we moved on to derive from the dynamics data two measures of the strength of coherency versus transparency percepts in plaids. Further experiments showed that these measures are reliably related to parametric manipulations. Furthermore, our results indicate that dynamics-based measures can be more sensitive than brief-presentation 2AFC measures, and reveal effects which were not known until now.

[1] See http://cns.nyu.edu/home/hupe/plaid_demo.

2. Preliminary experiment

The purpose of our preliminary experiment was to test whether the probability of perceiving the coherent and transparent interpretations is stable or whether it changes over time (e.g., due to adaptation). Observers (the two authors) watched a moving plaid for 5 min and reported their percept ("coherent" or "transparent") continually by pressing down one of two mouse buttons. (If the observer was unsure of the percept no button was pressed; this option was used less than 2% of the time.) The stimulus is as described in Section 3, with the following specific parameters: global (plaid) direction of motion: upwards; angle between the grating directions of motion ('α'): 115°; grating speed: 1°/s; duty cycle: 30%. The experiment was repeated 10 times with the same stimulus, but with very long breaks between consecutive trials: there were at most two trials per day (one in the morning and one in the evening).

Fig. 1 shows the durations of the coherent (a) and transparent (b) percepts for the 10 trials. Three observations stand out from the data.

(i) *The distributions of the durations of the two percepts are quite stable over time.* To quantify this, the data were fit by a linear regression, separately for the transparent and coherent percepts. (The first coherent percept was excluded, see below.) There was a modest but significant negative slope for the coherent percepts for observer JMH (log data: $F(1,118) = 11.8$, $p = 0.0008$), and a borderline-significant positive slope for observer NR ($F(1,98) = 4.66$, $p = 0.033$). For the transparent percept, neither observer showed a significant trend (JMH: $F(1,128) = 0.61$, $p = 0.44$; NR: $F(1,108) = 0.96$, $p = 0.33$).

(ii) *The first percept was always the coherent one*. This result could be specific to the particular set of parameters used, of course, but informal observations indicated that the coherent percept was typically the first one for a very wide range of plaid parameters. This is also consistent with the observations of Wallach (1935) and von Grunau and Dubé (1993).

(iii) *The first coherent percept was considerably longer than the subsequent coherent percepts*. The plaid was perceived as coherent for the first 20–30 s (bold symbols in Fig. 1a). Nevertheless, the transparent percept always occurred eventually, and the coherency periods that followed it were shorter, on average (open symbols in Fig. 1a). Such uniqueness was not observed for the first transparent percept (percept number 2 in Fig. 1b).

Having established that the average perceptual durations are stable over time, we next calculated the relative time spent perceiving coherency (i.e., the probability of

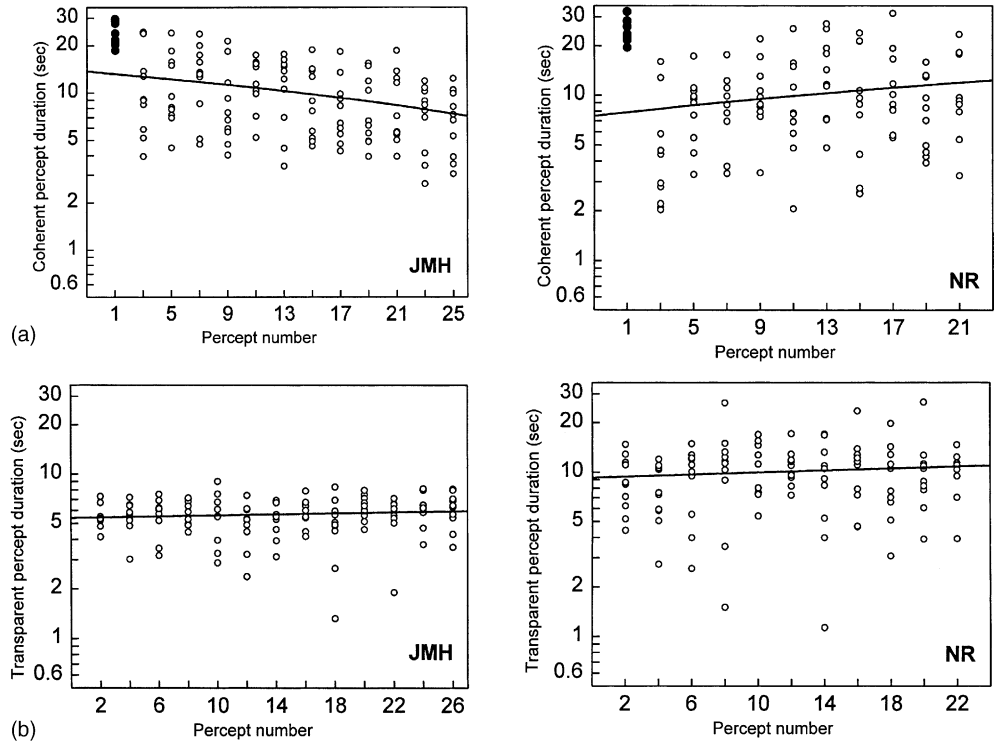

Fig. 1. Results of preliminary Experiment I. The scatterplots show the durations of the coherent (a) and transparent (b) percepts for 10 trials which lasted 5 min each. The durations are plotted as a function of their ordinal position within each trial, for two different observers (JMH, left, and NR, right). The first percept was always coherency (filled circles on top panels), and its mean duration was significantly longer than successive coherent epochs. The distributions of the durations of the two percepts are quite stable over time. (*Note*: only the durations of the first 26 for JMH and 22 for NR perceptual epochs are shown, since later trials had less values; see Fig. 4 and discussion refer footnote 3.)

seeing the coherent percept). It is given by $C/[C+T]$, where C and T denote the cumulative time spent reporting the coherent and the transparent percept over a given observation time. Importantly, the first coherency percept was excluded from C, and will be treated separately. We computed $C/[C+T]$ for successive 40 s durations within each trial, starting with the first report of the transparent percept, for the two observers. The results are shown in Fig. 2. There was no significant change of $C/[C+T]$ over time (i.e., the small change in the average coherency periods over time did not significantly change $C/[C+T]$). Thus, $C/[C+T]$ can be used as a measure of the steady-state probability to perceive coherency in a plaid. (This measure is analogous to that used in binocular rivalry studies; cf. Levelt, 1968).

Fig. 2 indicates that the coherency and transparency percepts were rather balanced for this stimulus, in terms of their steady-state probabilities ($C[C+T]$ was ~50%).

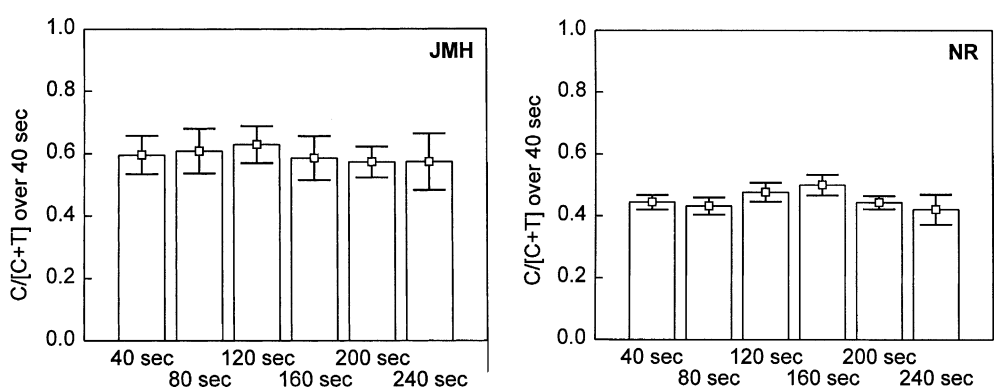

Fig. 2. The probability of the coherent percept is stable over long observation periods. $C/[C+T]$ was calculated for successive 40 s durations in each trial, starting with the first report of the transparent percept (based on data shown in Fig. 1). Bars represent the means of 10 trials, error bars are plus/minus one standard error (here and in all the subsequent graphs). The values of $C/[C+T]$ are not significantly different from each other (JMH: $F(5,54)=1.01$, $p=0.42$; NR: $F(5,54)=0.84$, $p=0.52$).

But this balanced phase came only after a prolonged duration of perceiving a coherent plaid (20–30 s, items (ii) and (iii) above). Informal observations with other plaid parameters suggested that a prolonged initial coherency phase was a common phenomenon. Furthermore, there were indications that the duration of this first coherency phase covaried with $C/[C+T]$. We therefore decided to treat the duration of the first coherency percept as another dependent variable, and termed it RTtransp ('the Response Time to report transparency'). We hypothesized that the variation in $C/[C+T]$ and RTtransp was driven by changes in the relative strength of the underlying coherent and transparent perceptual states (for other findings supporting this hypothesis see Hupé & Rubin, 2000). The experiments described here were therefore designed to explore how $C/[C+T]$ and RTtransp behave as a function of parametric manipulations, as well as how they are related to each other.

3. Methods

3.1. Apparatus

Stimuli were generated on a Silicon Graphics™ Indigo II workstation and displayed on a 19-in. monitor (45 cm viewable screen size) at a frame rate of 76 Hz. The screen resolution was 1280×1024 pixels. The SGI Graphics Library (GL) was used to generate the stimuli.

3.2. Stimuli: rectangular-wave plaids

Plaids composed of rectangular-wave gratings were presented through a circular aperture, 13° in diameter. The luminance of the background outside the aperture was 18 cd/m^2. The gratings comprised dark stripes (24 cd/m^2) on a light background (47 cd/m^2). The dark regions appeared as "figure" because the duty cycle, defined as [(width of dark bar)/(total cycle)], was always less than 50%, i.e., the dark stripes were thinner. The intersections regions' luminance was 19 cd/m^2, putting the plaid in the transparent regime (Stoner, Albright, & Ramachandran, 1990). The two gratings had the same spatial frequency (SF = 0.3 cycle/deg), duty cycle and speed, and the plaids were therefore completely symmetric. The image was refreshed every other frame to allow enough time for drawing the stimuli. (In spite of the reduced effective frame rate the motion appeared smooth; a few of the conditions were rerun with a true 76 Hz rate, by precalculating all frames and displaying them from memory, and the results did not differ at all; data not shown.) A colored fixation point was overlaid on a homogeneous circular patch (2.5° diameter, 18 cd/m^2) that covered the center of the plaid, to minimize OKN eye-movements. Observers were instructed to maintain fixation during the whole duration of stimulus presentation. The stimuli were viewed from a distance of 57 cm in a darkened room.

3.3. Stimuli: sinusoidal wave plaids

Sinusoidal plaids were generated by filling a circular region (7.7° diameter; viewing distance 100 cm) with the following space-time pattern $L(x,y,t) = L_m(1 + A \times [\sin(2\pi f_1(\cos(\theta_1)x + \sin(\theta_1)y - v_1 t)) + \sin(2\pi f_2(\cos(\theta_2)x + \sin(\theta_2)y - v_2 t))])$, where mean luminance $L_m = 15$ cd/m^2, contrast $A = 0.25$ and f_i, θ_i and v_i denote the spatial frequency, direction and speed of each grating; $v_1 = v_2 = 3$°/s. The pattern was precalculated for each frame of a full temporal cycle and displayed from memory at the 76 Hz frame rate. A colored fixation point was overlaid on a homogeneous circular patch (1.5° diameter, 15 cd/m^2) that covered the center of the plaid. The luminance of the background outside the aperture was 15 cd/m^2.

3.4. Observers

Observers were the two authors, five colleagues and five undergraduate students from New York University. The colleagues and students were naïve about the purpose of the experiments. The students were paid 10 dollars an hour for their participation. All participants had normal or corrected-to-normal vision. The two authors participated in the Preliminary experiment and in Experiment III. Four naïve observers participated in Experiment I. Two of them had no previous exposure to plaids. The two authors and seven naïve observers participated in Experiment II. Four of these observers had been previously exposed to plaid stimuli (designated O1, O2, O5 and O9). One observer (designated O8) participated also in Experiment I (but participated in Experiment II first).

3.5. Procedure

Each naïve observer was first shown examples of plaids and asked to describe what she/he saw. Observers typically described first the coherent percept (a pattern moving in a constant direction). Several examples of plaids (with randomly chosen parameters) were displayed until the observer spontaneously reported that the pattern separated into two independently moving gratings, i.e., described the transparent percept. The observers were then given an explanation that the stimulus was in fact ambiguous, and that it was just their perception of it which was changing. The instructions then depended on the experiment the observer participated in. In Experiment I, observers were asked to continually indicate when they perceived coherency by holding down a mouse button and when they perceived transparency by releasing the button. In Experiment II (measuring RTtransp), observers were asked to press a button as soon as they saw the plaid separate into two transparent gratings. In Experiment III and Preliminary experiment, the observers held down one button for coherency and another for transparency, and were allowed to not press any button when they were unsure of their percept.

In Experiment I, the stimulus remained on the screen for 1 min after the first report of transparency (i.e., the duration was RTtransp + 1 min), unless transparency was not reported within 2 min in which case the trial was terminated. In Experiment II, the trial ended after the observer pressed the button to indicate that she/he perceived the pattern as transparent. In Experiment III, the stimulus remained on the screen for 1 min after the first report of a switching percept, or for 2 min if no switch was reported. In all experiments, observers were told that in some trials it may happen that they would not experience the transparent percept at all, and that in such a case the trial would end after 2 min. They were further told that there was nothing wrong with this (not seeing transparency), and asked not to "try" to see more of one or the other percept ("passive" viewing instructions). Observers initiated each trial by pressing a mouse button. Naïve observers received a few practice trials before collection of the data shown.

3.6. Design

The experiments were set up as full factorial designs: all combinations of the different values of the independent variables were used. There were one (Preliminary experiment, Experiments I and III) or two (Experiment II) repetitions of the complete set of parameters in a randomized order.

3.7. Data analysis

The cumulative times spent reporting the coherent and transparent percepts, C and T, respectively, were measured from after the first perceptual switch to the end of each trial. The relative time seeing coherency in the steady-state phase is therefore given by $C/[C+T]$.

RTtransp was defined as the time from stimulus onset to the first report of transparency.

If a perceptual switch was not reported within the 2 min limit, $C/[C+T]$ was set to 0 or 1 (depending on the reported percept); if $C/[C+T]$ was 1, RTtransp was set to 120 s. Note that since the first epoch was excluded from the computation of the cumulative times, $C/[C+T]$ and RTtransp are methodologically independent.

The independent variables were categories, such as observer identity, and continuous predictors (or covariates), such as the angle "α" between the gratings' directions of motion (see Section 4 for specific variables and values). Data were run through an analysis of covariance (ANCOVA; Statistica, StatSoft™), with either RTtransp (Experiments I and II) or $C/[C+T]$ (Experiments I and III) as the dependent variable. A condition of validity of this analysis is that the noise in the data is normally distributed. This condition was satisfied for $C/[C+T]$ (e.g., Kolmogorov–Smirnov test for Experiment I: $d = 0.04$, $N = 191$, not significant). But the distribution of the RTtransp values was highly skewed and the distribution of the residuals was significantly different from normal in both Experiments I and II. Transforming RTtransp values to their natural logarithm provided the best correction [2] (e.g., see Fig. 3). Another condition of validity of an ANOVA is that the variances be homogeneously distributed. To test this, the standardized residual values were plotted as a function of the ANCOVA-predicted values, and these scatterplots were visually inspected for each analysis. The variances of $C/[C+T]$ were judged to be homogeneously distributed (Experiments I and III). For ln(RTtransp), the variances were homogeneously distributed in Experiment I but not II. This issue will be addressed in Section 4. The analysis of residuals was also used to remove outlier values (when z-score were too low or too high: 25 outlier values in Experiment II, 3 in the sinusoidal plaids in Experiment III, none in Experiment I).

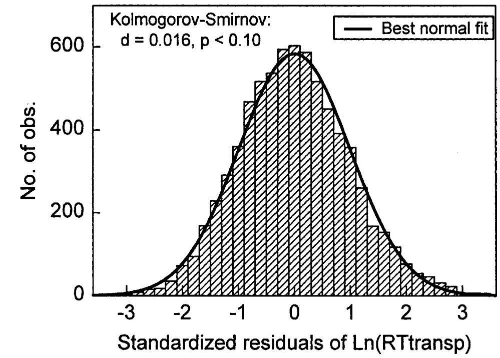

Fig. 3. The histogram of the standardized residuals of ln(RTtransp) for Experiment II is well fit by a Gaussian ($N = 7279$, 25 outliers excluded).

4. Results

4.1. Experiment I

This experiment tested the effect of three independent variables on $C/[C+T]$ and RTtransp. The variables used were: α, the angle between the gratings' directions

[2] The distribution of the subsequent percept durations was also well fit by a log-normal function. In other domains of bi-stability, like binocular rivalry, Gamma functions have typically been used, but log-normal functions are in fact as good or even better (Lehky, 1995).

of motion (90°, 105° and 120°), the gratings' speed (1°/s, 2°/s, 3°/s, and 4°/s), and the global direction of motion of the plaid (four oblique directions, ±45° and ±135°). α and speed were treated as continuous predictors and direction and observer identity as categories. α was chosen since it was previously shown to have a powerful effect on the tendency to perceive coherency versus transparency (Adelson & Movshon, 1982; Kim & Wilson, 1993). Speed was also suggested as a central factor in plaid perception (Farid & Simoncelli, 1994; Smith, 1992; von Grunau & Dubé, 1993). Finally, we chose to vary the global direction in order to avoid between-trial adaptation (previous studies showed that coherency is affected only by adapting stimuli which move in the same direction as the 'test' plaid; Movshon et al., 1985; von Grunau & Dubé, 1993).

4.1.1. Results

4.1.1.1. Dynamics of the perceptual alternations. Fig. 4 presents the durations of successive coherency and transparency epochs, averaged across all observers and parametric configurations. (Only the 149 trials for which the number of perceptual alternations was six or more were included; [3] the seventh bar shows data from 133 trials, since the remaining trials terminated within that period.) The data confirm the results obtained in the Preliminary experiment. First, the average duration of the first, coherent percept (RTtransp) is much longer than the duration of successive coherent percepts. Second, the average duration of the subsequent coherent and transparent percepts are stable over time ($F(2,444) = 2.07$, $p = 0.13$ and $F(2,428) = 2.15$, $p = 0.12$, respectively). This validates $C/[C+T]$ as a reliable measure of the steady-state probability to perceive coherency. More generally, these results support the validity of the dynamics approach for studying plaids. They alleviate two potential concerns about prolonged exposure to plaids in multiple successive trials: the long initial coherency phase is a general phenomenon (i.e., occurs not only for a temporally isolated trial), and there is no indication of between-trial adaptation.

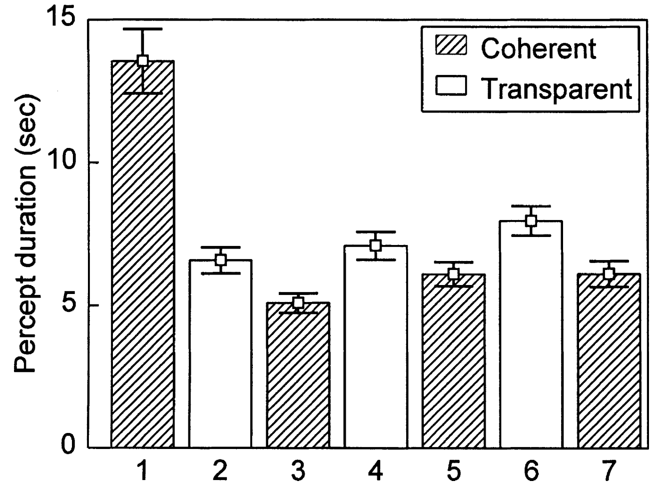

Fig. 4. Average durations of successive coherent and transparent percepts in Experiment I. Observation time was limited to (RTtransp + 1 min). Trials which produced less than six alternations within this limit were discarded, so each average is computed over the same number of trials (149 trials—except the seventh percept: 133 trials). The first bar shows the average of RTtransp values. See text for more explanations.

4.1.1.2. The relation between RTtransp and $C/[C+T]$. Fig. 5 shows a scatterplot of $C/[C+T]$ as a function of ln(RTtransp) for the 48 stimuli and four observers (191 data points; one trial was aborted by one observer). The ln(RTtransp) scatterplot shows a linear relationship with $C/[C+T]$. Interestingly, although RTtransp was transformed to log values merely to obtain a normal distribution of residuals (see Section 3), Fig. 5 now in-

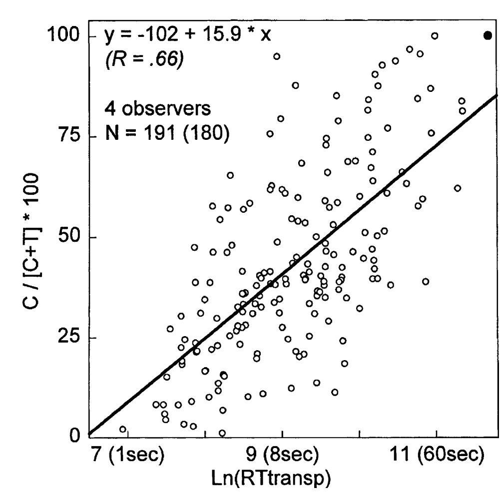

Fig. 5. Correlation between ln(RTtransp) and $C/[C+T]$ in Experiment I (same data as in Fig. 4). Each point in the scatterplot represents one trial. The filled circle (top right) indicates 11 trials where sliding was not reported within the allowed 2 min (see Section 3). These values were excluded from the regression analysis.

[3] When computing the average durations from continual-report data, it is important to avoid artifacts associated with the last part of the observation time of trials. For a fixed viewing time, the total number of epochs within a trial depends on what durations happened to occur on that trial. Trials that happened to contain many short-period alternations would also have more periods, on average. This means that if one tallied *all* the epochs, less and less trials would contribute to the later epochs, and the distribution of durations in those epochs would in turn be skewed towards low values. Averaging all values would thus lead to an apparent decrease in average epoch duration over time. However, this would be an artifact of the calculation method, not a true decrease. (We believe that this is what accounts for the apparent decrease in percept duration in plaids reported by von Grunau & Dubé, 1993.) Such a "boundary artifact" can be avoided by identifying the minimal number of alternations reached in all trials, and basing the analysis only on data up to that point.

dicates that ln(RTtransp) is in fact the appropriate dependent variable to consider, since it is proportional to the steady-state probability of seeing the coherent percept.

The correlation between $C/[C+T]$ and ln(RTtransp) is clear but not very strong ($R = 0.66$). There were indications that this was due to differences between individual observers' slopes. The best-fit slopes for the four observers were not the same (data not shown). To obtain enough data to examine the tightness of the correlation in an individual observer, one of the authors (JMH) performed a similar experiment with more parameters and trials (216 stimuli; total observation time: about 10 h). Fig. 6 shows that the correlation between RTtransp and $C/[C+T]$ is indeed tight, and the linear relation with ln(RTtransp) is very strong ($R = 0.90$).

The specific relationship revealed between $C/[C+T]$ and RTtransp has an important implication for brief-presentation methods. Let us define *RTtransp[C50]* as the value of ln(RTtransp) for which $C/[C+T]$ is 0.5, i.e., when the coherent and the transparent percepts have equal probability in the steady-state. The RTtransp[C50] value for the four naïve observers (Fig. 5) is 14.6 s (individual values: 15, 9, 14 and 28 s) and for JMH it is 19 s (Fig. 6). This means that if we used a brief-presentation 2AFC method, stimuli which have a steady-state transparency probability near 50% would yield "coherency" responses almost 100% of the time, since they take between 10 and 30 s to start sliding. This, in turn, would preclude the possibility of observing any effect of parametric manipulations. Our data indicate that this methodological problem of a 'ceiling effect' is a primary concern about brief-presentation methods (see also Section 4.3).

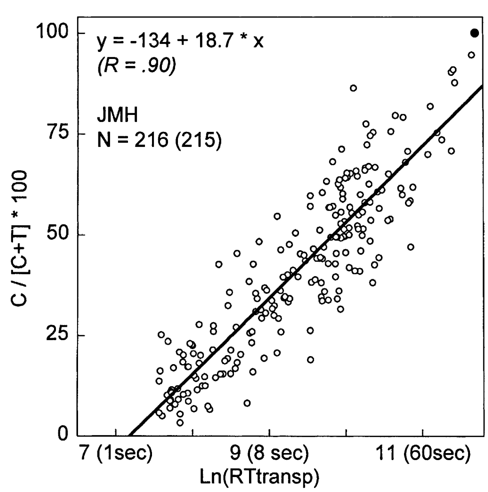

Fig. 6. Correlation between ln(RTtransp) and $C/[C+T]$ for observer JMH. Data were gathered in an experiment similar to Experiment I, but with more parameters and trials. Same conventions as in Fig. 5 (sliding was not reported within 2 min for only one trial).

4.1.1.3. The effect of parametric manipulations. Next, we examined the effect of the parametric manipulations on RTtransp and $C/[C+T]$. Table 1 summarizes the results of ANCOVAs performed on these two dependent variables. Both α and the speed had significant effects, on both $C/[C+T]$ and RTtransp. The global direction of motion did not have a significant effect (but see below, Section 4.2). Observer identity was a significant factor only for $C/[C+T]$. F values were comparable for the analysis on both dependent variables (except for the observer effect), and most of the variance in the data could be accounted for by α. Fig. 7a and b illustrate the effects of α and speed, respectively, on $C/[C+T]$ and ln(RTtransp). The data are collapsed across the different values of speed (Fig. 7a) and α (Fig. 7b), as well as across observers and the four values of global direction of motion. Although the tight correlation between $C/[C+T]$ and RTtransp already indicated that the two curves should behave similarly, their quantitative agreement is impressive.

The effect of increasing α on the tendency for coherency (reducing it) is in agreement with previous studies (Adelson & Movshon, 1982; Kim & Wilson, 1993), underscoring the validity of our dynamics-based measures. The reduction in coherency with increasing speed, while significant, was moderate. This is again consistent with previous studies (Smith, 1992; von Grunau & Dubé, 1993). Note that studies which reported stronger effects of speed used a different manipulation: those studies introduced different speeds to the two gratings (Adelson & Movshon, 1982; Movshon et al., 1985), creating a situation where the two gratings have different attributes, which is known to reduce coherency for other parameters (e.g., contrast or spatial frequency). Finally, the presence of an observer identity effect for $C/[C+T]$ but not RTtransp reflects individual differences in

Table 1
Results of the ANCOVA for Experiment I

$C/[C+T]$	Degrees of freedom	F	p	ln(RTtransp)	Degrees of freedom	F	p
α	1	367.8	<10–16	α	1	301.5	<10–16
Speed	1	15.6	<10–3	Speed	1	22.0	<10–5
Observer	3	12.5	<10–6	Observer	3	1.2	0.30
Direction	3	1.5	0.21	Direction	3	0.2	0.87
Obs. *Direction	9	1.2	0.30	Obs. *Direction	9	0.8	0.63
Error	173			Error	173		

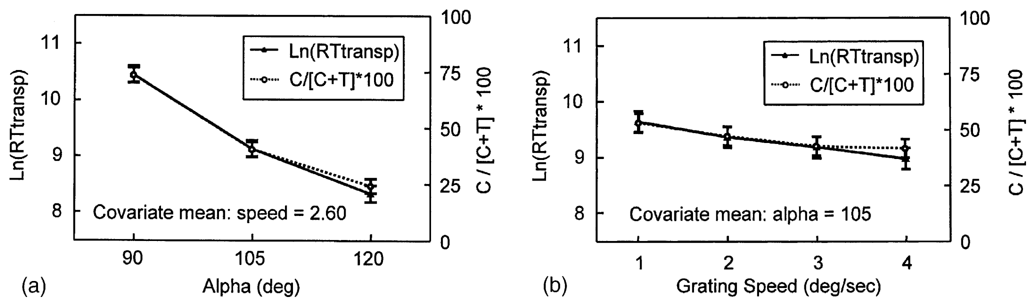

Fig. 7. The effects of α and speed on the two dynamics measures. *Dashed curve*, the probability of the coherent percept, measured by $C/[C+T]$, as a function of the angle α between the grating directions. *Solid curve*, RTtransp as a function of α. The scale of the two vertical axes was set so that the first two points coincide. (a) The effect of α. Each data point represents the least square mean estimated by a linear model with the speed as a covariate and observer as a category (64 trials per data point). (b) The effect of the stimulus speed on $C/[C+T]$ and RTtransp (α covariate, 48 trials per data point).

observers' regression slopes between these two variables, and is incidental to the particular set of parametric values used here (in Experiment II observer identity will be shown to have a strong effect also on RTtransp).

4.2. Experiment II

The tight correlation between RTtransp and $C/[C+T]$ found in Experiment I, and the close agreement of the dependence of the two measures on parametric manipulations, suggested that it may be possible to study plaids by measuring only RTtransp. This would have great methodological advantages, since RTtransp data can be collected much more efficiently. Instead of asking observers to report their percepts continually for prolonged amounts of time, we just have to ask them to indicate the moment when the plaid separated into two transparently moving gratings. Experiment II tested the feasibility of this method. Four independent variables were used: α, the gratings' speed, the duty cycle and the global direction of motion of the plaid (α, speed and duty cycle were treated as continuous predictors, the global direction and observers' identity as categories). Table 2 shows the sets of parametric values used for the three groups of two, four and three observers who participated in the experiment.

The RTtransp protocol is very different from the one used so far, because observers spend a disproportional part of the time experiencing coherency (since the stimulus disappears as soon as they report the transparent percept). This raises again potential concerns about adaptation. It has been shown that viewing a strongly coherent plaid for a while can influence how a subsequent plaid is perceived, e.g., whether coherency or transparency is perceived first (von Grunau & Dubé, 1993). However, this may not necessarily be a problem in cases where many of the stimuli are *not* strongly coherent (in the steady-state), and they are simply viewed only for the initial, coherent phase. Also, the results from Experiment I suggested that randomizing the global direction of motion helps to avert adaptation effects (in this experiment, we increased the number of directions from four to eight). As we shall see, the data obtained indeed indicate that the values of RTtransp obtained in this protocol are comparable to those obtained in the long-presentation protocol used before.

Table 2
Values of the parameters used for the three groups in Experiment II

	Group 1		Group 2		Group 3	
	# Values	Values	# Values	Values	# Values	Values
Observers	2	JMH–NR	4	O3–O6	3	O7–O9
α (deg)	4	130–140–150–160	4	120–135–150–165	3	120–135–150
Speed (°/s)	4	1.3–1.9–2.6–3.2	3	0.65–1.6–2.6	3	1.0–2.1–3.1
Duty cycle (%)	4	10–20–30–40	3	10–25–40	3	15–25–35
Duty cycle (%)			3	17.5–32.5–47.5		
Directions (deg)	8	0–45–· · ·–315	8	0–45–· · ·–315	8	0–45–· · ·–315
Repetitions	2		2		1	
# Sessions	2		4	(3 duty cycle val./session)	1	
# Trials/obs.	1024		1152		216	

Note that the observers of group 2 were tested with a total of six different values of duty cycle. However, not all six values were presented in each session. In each of the four sessions, observers were presented only three values of duty cycle.

Table 3
Results of the ANCOVA for Experiment II

ln(RTtransp)	Degrees of freedom	F	p
α	1	8596.6	<10–16
Speed	1	701.6	<10–16
Duty cycle	1	120.9	<10–16
Observer	8	526.9	<10–16
Direction	7	154.8	<10–16
Obs. *Direction	56	7.7	<10–16
Error	7204		

4.2.1. Results

Table 3 shows the results of the ANCOVA. Most of the variance in the data is accounted for by α ($F[1, 7204] = 8596.6$). The other independent variables all had highly significant effects, too. The effect of the global direction of motion was predominantly due to a greater tendency for sliding when the plaids moved in oblique directions. (This is why the global direction did not show an effect in Experiment I, where only oblique directions were used; a more detailed analysis will be given elsewhere.) There was also a significant effect of observer identity.

4.2.1.1. The effect of α. Fig. 8 shows the effect of α (the difference between the gratings' direction of motion) on RTtransp for each of the nine observers. The effect of α was strong and monotonic for each of the observers. The dependence on α was nearly linear for some observers (O2, O6, O8, solid lines), but asymptoted at larger values of α for others (O1, O3, O4, O5, O7, O9, dashed lines). Noting that the deviation from linearity tended to be present for the observers whose average RTtransp values were lower (solid curves), we suspected it may arise from a floor effect. (The dashed graphs show a greater tendency to asymptote as RTtransp goes below 1 s; fitting a quadratic curve for each observer led to significantly greater second-order coefficients for the 'fast' observers; Mann–Whitney, $p < 0.05$.) To test this conjecture, we re-examined the results of each observer by separating the data to subsets which had low versus high RTtransp values for other parameters (low-RTtransp set: higher speeds and oblique global directions; high-RTtransp: low speeds and horizontal direction). For each observer, the data from the high-RTtransp set was indeed much better fit by a linear curve than the low-RTtransp set (Wilcoxon, $p < 0.008$). Dividing the parameter space to high-RTtransp and low-RTtransp sets also allowed us to address the issue mentioned in Section 3, that the residuals were not homogeneously distributed in this Experiment. The separate analysis shows that the deviation from homogeneity is found only for the low-RTtransp set, relating it to a floor effect. Further details of the analyses mentioned above can be found in http://cns.nyu.edu/home/hupe/plaid_demo/suppl.htm. Thus, we conclude that increasing α, which reduces the tendency for coherency, leads to a linear decrease in RTtransp, except near the minimum possible response time, where RTtransp asymptotes towards this value.

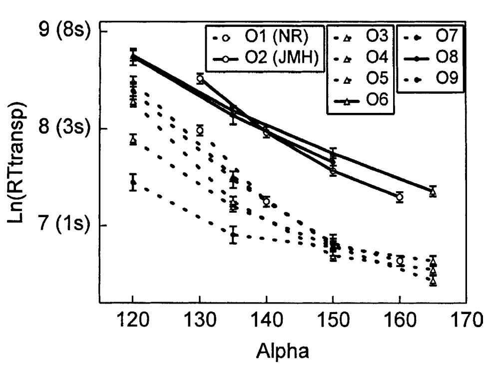

Fig. 8. The effect of α on RTtransp for all observers (Experiment II). Same conventions as in Fig. 7. The covariates were the speed and the duty cycle. The global direction of motion was a category (8 values). Three different sets of parameters were used with different observers (see Table 2). Each data point was computed from 256 (O1–O2), 288 (O3–O6) and 72 (O7–O9) trials.

4.2.1.2. The effect of speed. Fig. 9 shows RTtransp as a function of the gratings' speed for the nine observers. The effect of speed was non-linear, with a significant effect on RTtransp only at slow speeds (<1.5°/s). The non-linearity of the curves is not the result of a floor effect: the curves asymptote at high speeds for observers with high average RTtransp values (solid curves) just as much as for observers with low average RTtransp values (dashed curves). Moreover, this non-linearity was present independently of the range of other parameters. This is shown in Fig. 10 for α: the effect of speed is similar for values of α that give high as well as low average RTtransp. The interaction between α and speed, though significant, is small compared to the main effects. Our conclusion if therefore that speed by itself has little effect on the mechanisms of integration and segmentation for grating speed above 1.5°/s. Other questions are whether in the range where speed does have an effect

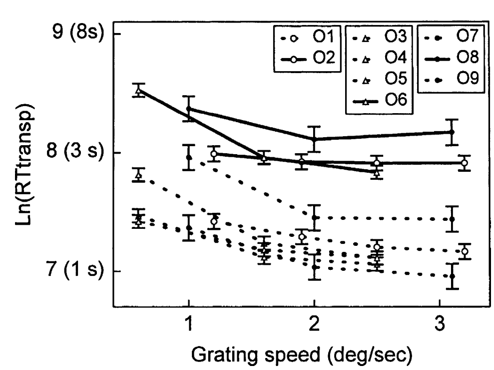

Fig. 9. The effect of the gratings' speed on RTtransp. For each value of speed, RTtransp values were averaged over the different values of α, global direction and duty cycle. Three groups of observers were tested with different sets of speeds (see Table 2).

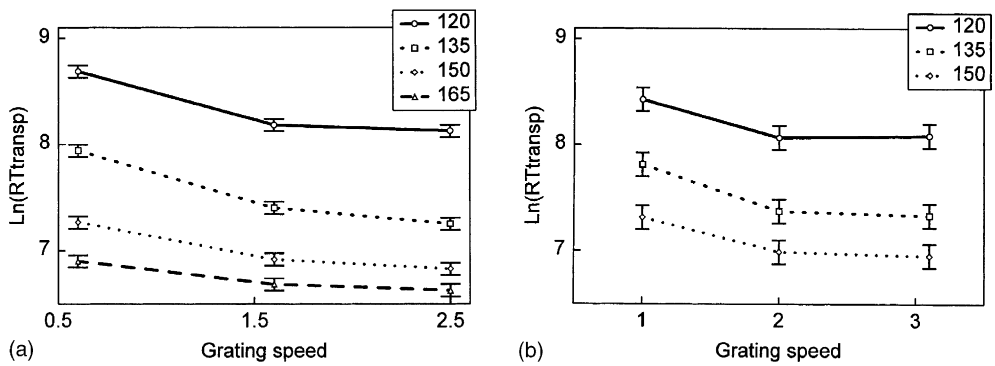

Fig. 10. Interactions between α (inset) and grating speed: (a) results for observers 03–06 ($n = 384$ for each data point); (b) average for observers O7–O9 ($n = 72$).

(below 1.5°/s), it is the grating speed or plaid speed that matters (or perhaps both), and whether temporal frequency or speed matters more. More experiments are needed to draw definite conclusions about this. Our results indicate that the dynamics approach has the potential to resolve this issue more comprehensively than was possible before.

4.2.1.3. The interaction between α and speed. The absence of notable interaction between α and speed bears significance on the interpretation of the effect of α. The parameter we manipulated directly was the speed of the gratings (V_g), but since we also manipulated α, identical grating speeds corresponded to different plaid speeds ($V_p = V_g/\cos[\alpha/2]$). Previously, it has been suggested that the effect of α could be at least partly explained by the change in plaid speed (Farid & Simoncelli, 1994; Farid, Simoncelli, Bravo, & Schrater, 1995). However, if that were the case, we should have observed a strong interaction between speed and α (comparable to the effect of α itself). The very weak interaction between speed and α therefore rules out this hypothesis. Another related issue is the linearity of the effect of α. The linear dependence of coherency was obtained for constant grating speeds (see also Section 4.3). Would this behavior change had we instead held the plaid speed constant? The weak (or even the absence of) effect of grating speed above 1.5°/s predicts that the answer is no, at least as long as very slow speeds are not used. In Experiment III, we confirmed this directly on one observer (JMH): varying α between 45° and 135° while holding plaid speed at 4°/s (resulting in grating speeds in range 1.5°/s–3.7°/s, average = 2.7°/s) yielded results virtually identical to those obtained when gratings' speed was held constant at 2.5°/s (resulting in plaid speeds in the range 2.7°/s–6.5°/s, average = 4.1°/s): $F(3,48) = 0.48$, $p = 0.70$ ($N = 64$). Moreover, the relationship between α and $C/[C+T]$ was perfectly linear whatever speed was manipulated. (The combined data are presented in Fig. 16, dashed line.) To summarize, our results indicate that the effect of α is linear and that it is not mediated by either grating or plaid speed.

4.2.1.4. The effect of duty cycle. Fig. 11 shows RTtransp as a function of the gratings' duty cycle for the three groups of observers. Gratings comprised of "thinner" bars had a greater tendency to slide than gratings with "thicker" bars. This effect, while significant (cf. Table 3), was fairly small. The effect of duty cycle on the tendency of plaids to slide had been reported before (Stoner & Albright, 1992; Stoner & Albright, 1996), but only when it affected the figure/ground interpretation of the stimulus, so it was suggested that it was related to changes in segmentation cues. In our experiment, the dark bars were always thinner that the light bars, and were systematically perceived as the figure. Therefore, it is unlikely that the segmentation cues explanation could account for our small effect of duty cycle. Another possibility is that the effect is mediated by changes in contrast, since the contrast of the stimulus is affected by manipulations of the duty cycle. (The Michelson contrast is defined only when the duty cycle is 50%. To compute the contrast when the duty cycle was different

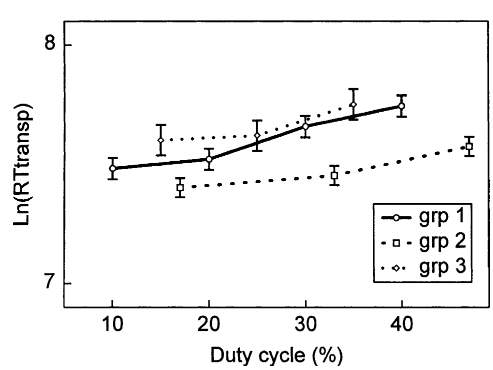

Fig. 11. The effect of duty cycle on RTtransp for the three groups of observers. RTtransp values were averaged over the different values of α, speed and global direction of motion. The data shown here for group 2 were computed from the data of two sessions using the same values of duty cycle (see Table 2).

from 50%, we used the formula: [Maximum_luminance − Minimum_Luminance]/[2× average_luminance].) It has been reported that the probability of coherency increases as contrast is increased (Smith, 1992). In our stimuli, when the duty cycle was 10%, the contrast was 26%; and when the duty cycle was 47.5%, the contrast was 32%. This variation is very small, but so is the effect of duty cycle. More experiments would therefore be needed to decide whether this effect of duty cycle is mediated by contrast.

4.3. Experiment III

In Experiments I and II, the effect of α on the relative strength of coherency was not only monotonic but also near-linear (Figs. 7a and 8). This is different from what was reported previously with short presentation 2AFC methods. Kim and Wilson (1993) obtained a sigmoid-shaped function with a transition between almost 100% coherency to 100% transparency around α of 90° (cf. Fig. 14b). The rapid transition in Kim and Wilson's data occurred over an α range of 36.8°, out of the wide range of 106.4° they used. We therefore asked whether the near-linear behavior of $C/[C+T]$ and RTtransp in Experiments I and II may happen to be valid only in the more restricted range of αs used there (30° and 45°, respectively). To test this, we measured $C/[C+T]$ and RTtransp over a wide range of α values: 15°–165°. Fig. 12 shows that the variation of both measures is well-fit by a linear curve over this entire range. Saturation in coherency ($C/[C+T]=1$) and transparency ($C/[C+T]=0$) was reached only when α approached its lowest and highest possible values of 0° and 180°, respectively.

There was another major difference between our Experiments I and II and Kim and Wilson's (1993) experiment, besides the different methods: their stimuli occupied a very different place in parameter space. Most notably, the plaids were composed of sinusoidal gratings with unequal spatial frequencies. To test whether these parametric differences might have led to the qualitatively different results they obtained, we set out to replicate their stimuli as closely as possible (see Section 3) and apply the dynamics approach to them. In order to collect reliable continual-report data, the experimental design was changed from that used by Kim and Wilson (1993) in two ways. The global direction of the stimulus was varied between the four cardinal directions from trial to trial, and a fixation point was superimposed on a small homogeneous circular patch in the center of the stimulus (as in our previous experiments). Finally, while the spatial frequency (SF) ratios we used were identical to those used by Kim and Wilson (1993), the absolute SF values were decreased: 2, 0.67, 0.33 cycle/deg in our experiments instead of 6, 3 and 1 cycle/deg used by Kim and Wilson (see stimulus picture in Fig. 16). This was done because at their speed of 3°/s, some of the original gratings resulted in very high temporal frequencies (e.g., 18 Hz for the 6 cycle/deg grating); under fixation conditions this led to severe flicker or, alternatively, phenomenal disappearance (the grating turned into a homogenous surface). With these changes in place, the perceptual alternations of the sinusoidal plaids had a dynamical behavior similar to that found for the rectangular plaids tested in Experiments I and II, with a log-normal distribution, mean durations stable over time, and a stable $C/[C+T]$ (provided the first percept was excluded), validating it as a measure of coherency (data not shown).

Fig. 13 shows the results for two observers (the authors). Over the entire range of α values tested (45–135 in 30° steps), the variation of $C/[C+T]$ was gradual and well-fit by a linear function. This was true for both SF ratios tested, 3 and 6. These results indicate that the true underlying relationship between α and coherency strength is linear also for the plaid parameters used by Kim and Wilson (1993). Re-analysis of the continual-report data to simulate what a brief-presentation 2AFC design would yield provides direct support that the

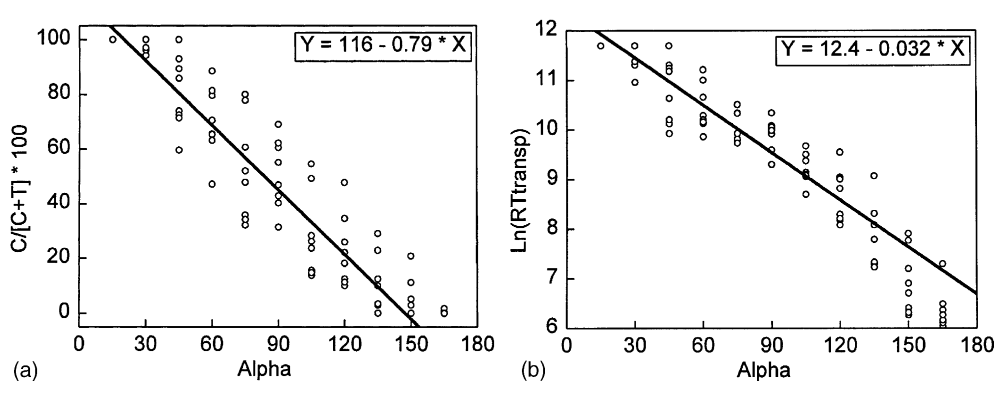

Fig. 12. The dependence of $C/[C+T]$ (a) and RTtransp (b) on α is well-fit by a linear curve for a wide range of α values (fitting for α between 40° and 140°). The stimuli were rectangular plaids similar to those of Experiment II (duty cycle: 25%; speed: 3°/s), presented for [RTtransp + 40 s]. Each data point corresponds to one measure (one of eight possible global directions). The observer was one of the authors (JMH).

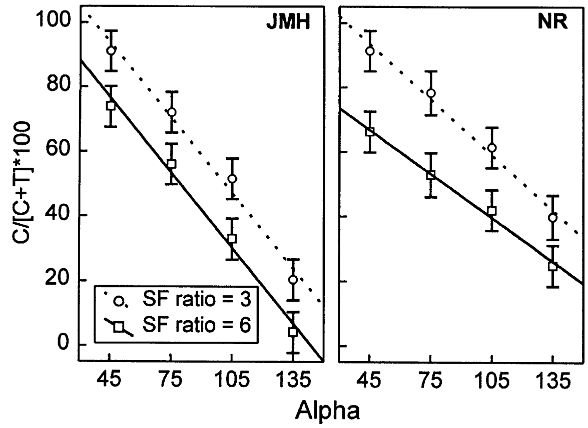

Fig. 13. The linear relationship between the strength of the coherent percept and α, revealed by the dynamics-based measure $C/[C+T]$, remains valid for plaids composed of additive sinusoidal gratings with SF ratios of 3 and 6, similar to those used by Kim and Wilson (1993).

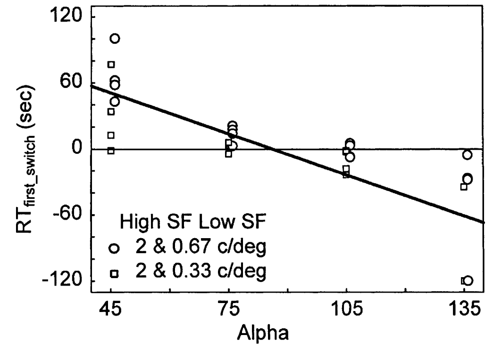

Fig. 15. $RT_{\text{first\_switch}}$ as a function of α for the data presented in Fig. 14a. Positive values indicate that the first percept was the coherent one, negative values that it was the transparent one. $RT_{\text{first\_switch}}$ was computed from stimulus onset. In this set data, the first percept was always reported within 1.1 s after stimulus onset, and the first percept switch never occurred until at least 1.7 s.

sigmoid-shaped function obtained by Kim and Wilson (1993) resulted from ceiling and floor effects inherent to this method. We took the first 1 s of each trial and classified it as "coherent" or "transparent" based on which of the two percepts was reported more within that time window. We then averaged the trials for each value of α to obtain a "% coherency" measure. The resulting curves are shown in Fig. 14a: they saturate at 100% and 0% coherency for small and large α, respectively, closely resembling the functions obtained by Kim and Wilson (1993, Fig. 14b). A more intuitive explanation of this result will be given below.

The sinusoidal plaids behaved similarly to rectangular plaids in terms of the effect of α and of bi-stability, but they showed a prominent difference in another aspect of their dynamical behavior: the first percept was often the transparent one. Recall that for the plaids used previously (Experiments I, II and Fig. 12), the first percept was 'coherency' even when $C/[C+T]$ was much lower than 0.5. But for the sinusoidal plaids of Fig. 13, the first percept behaved in a more symmetrical way: it was often 'transparency' for plaids with $C/[C+T] < 0.5$. To illustrate this, we defined a single measure of the first percept duration, $RT_{\text{first\_switch}}$, by giving it a minus sign for trials that started with transparency. Fig. 15 shows $RT_{\text{first\_switch}}$ as a function of α. For large α values, which yielded $C/[C+T]$ values less than 0.5 (see Fig. 13), the first percept tended to be transparent (compare with Fig. 12b, where the first percept was always the coherent one). This graph also provides better understanding why a brief-presentation 2AFC design would yield a sigmoid curve. Suppose that we used a trial duration of 1 s. $RT_{\text{first\_switch}}$ never went below 40 s for $\alpha = 45°$ (for SF ratio = 3). This means that for these parameters the observer would never get a chance to perceive transparency within 1 s and thus would have responded "coherency" 100% of the trials. Conversely, for $\alpha = 135°$, all trials yielded $RT_{\text{first\_switch}}$ values below −5 s, and therefore the observer would have responded "transparency" 100% of the trials.

What caused the shift towards a symmetric first percept exhibited in Fig. 15? It seemed natural to suspect

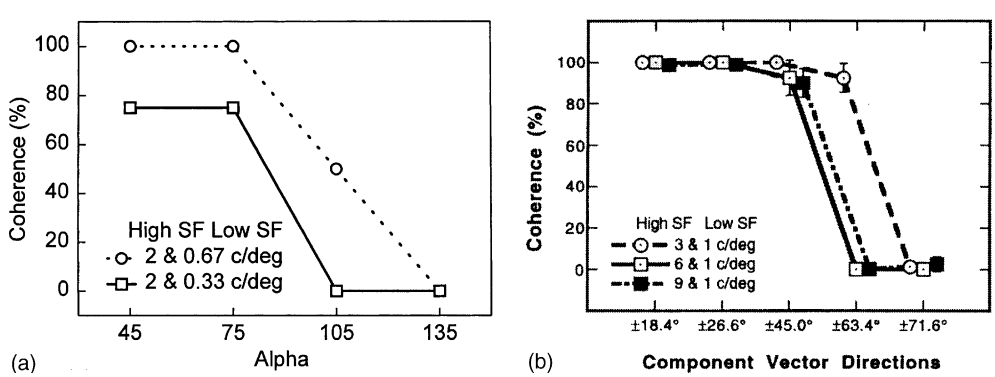

Fig. 14. (a) The continual-report data of Fig. 13 were re-analyzed to simulate what results a 1 s presentation 2AFC method would yield (see text). The frequency of "coherent" responses yields a sigmoid-shape curve as a function of α. Data for JMH, vertical global directions, four values (two repetitions) per datum point; (b) data obtained with 1 s presentation 2AFC method with similar stimuli as those used for Fig. 13 (reproduced, with permission, from Kim & Wilson, 1993, Fig. 3C).

the two most prominent new features of the plaids used here: they were composed of sinusoidal gratings while the plaids in Experiments I and II had rectangular luminance profiles, and the SF ratios of the gratings differed from 1. However, preliminary experiments indicate that neither of these parameters is necessary to induce a change in the nature of the first percept. Rather, the different behavior at stimulus onset seems to occur for plaids that consist of 50% duty cycle gratings, regardless of other aspects of their Fourier content (e.g., sinusoidal versus square-wave) or their SF ratios. This is only a tentative conclusion at this point, supported by results from one observer (JMH). We tested three sets of plaids: (1) square-wave plaids (duty cycle 50%) with SF ratios of 3 and 6; (2) sinusoidal plaids (which have duty cycle 50% by definition) with a SF ratio of 1; (3) square-wave plaids with a SF ratio of 1; the bias towards coherency in the first percept was strongly weakened in all of these configurations. Since 50% duty cycle was the only common feature to these three sets, it suggests it may be a determinant factor. The idea that shifting to 50% duty cycle has such a dramatic effect on the perception of plaids at stimulus onset may seem surprising given the relatively weak effect that varying duty cycle had in Experiment II. But note that in that case, this parameter was always kept in the range below 50%. It is possible that there is a qualitative change when duty cycle reaches 50%, a unique point where the figure/ground interpretation of the gratings becomes ambiguous. More experiments are needed to follow up this intriguing possibility, as well as to reevaluate the possible effect of other parameters on plaids' first percept.

The preliminary explorations described above offered an opportunity to compare the behavior of sinusoidal versus square-wave plaids over a wide range of parameters. The results indicate that the two types of plaids behave very similarly in terms of the relative strength of coherency. This is illustrated in Fig. 16 for the effect of α on $C/[C+T]$. The linear dependence on α was also similar to what was found for the rectangular plaids (dashed line). It is worthwhile to note that this close quantitative agreement occurs even though there is a noticeable difference between sinusoidal and rectangular wave plaids in terms of their phenomenal appearance. For the rectangular (or square) wave plaids, the coherent and transparent percepts were perceptually distinct and the transitions between them very sharp (in time). In contrast, the sinusoidal plaids sometimes gave the impression of non-rigid motion—neither fully coherent, nor clearly transparent. This made the task of reporting coherency/transparency noticeably more difficult for those stimuli. This difficulty is not unique to our continual-report task: several authors who used a brief-presentation 2AFC paradigm have commented on it (e.g. Kooi, De Valois, Switkes, & Grosof, 1992b). Indeed, some authors reported that naive observers needed

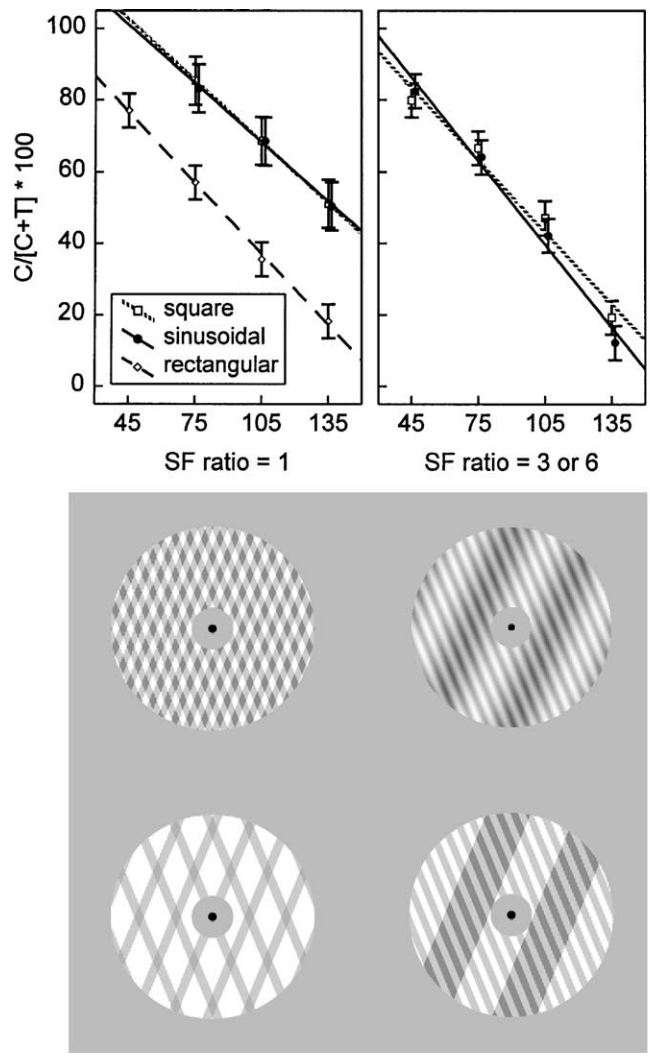

Fig. 16. The linear relationship between the strength of the coherent percept and α, revealed by the dynamics-based measure $C/[C+T]$, remains valid in a wide range of plaid parameter space (all data for observer JMH): (filled dots symbols/solid lines) plaids composed of additive sinusoidal gratings similar to those used by Kim and Wilson (1993). The data for SF ratio = 3 or 6 are the same as those of Fig. 13. For SF ratio = 1, SF was 2 cycle/deg; (square symbols/dotted line), gratings' luminance profile turned to square-wave, all other parameters unchanged; (diamond symbols/dashed line) plaids similar to those used in Experiments I, II and Fig. 12 (duty cycle = 25%, SF = 0.5 cycle/deg, transparent intersections). The pictures illustrate some of the stimuli. Left: Square-wave gratings with SF ratio = 1 and rectangular plaids; Right: sinusoidal and square-wave gratings with SF ratio = 6.

significant practice before they could classify sinusoidal plaids as coherent or transparent consistently (e.g. Movshon et al., 1985). This is very different than what we found for the rectangular wave plaids. There, the coherency/transparency transitions were clear and easy to detect, making the task suitable for unpracticed observers. The finding that sinusoidal and rectangular-wave plaids have similar parametric dependencies is therefore useful methodologically, since it suggests that rectangular-wave plaids may be used in experiments involving "coherent/transparent" judgments without loss of generality, facilitating task performance.

5. Discussion

We studied the dynamics of perceptual alternations in plaids by asking observers to continually report whether they perceived coherency or transparency. This paradigm has been used extensively in the study of other bi-stable phenomena, most notably binocular rivalry, and we found that it could be naturally adapted to plaids. Furthermore, the systematic and meaningful relationships that the data showed with parametric manipulations confirm the validity of this method for studying plaids.

The observation that moving plaids give rise to bi-stable perceptual alternations goes back to the very first description of such stimuli by Wallach (1935) (English translation in Wuerger et al., 1996), who gave lucid and detailed descriptions of the phenomenology. But from Adelson and Movshon's (1982) paper, which re-introduced plaid stimuli as a tool to study motion processing, and onwards, few authors have commented on the spontaneous transitions between the two interpretations. We found only one study that studied the perceptual transitions directly (von Grunau & Dubé, 1993). In all other studies, researchers used brief presentations (typically less than 3 s) and asked observers to make a 2AFC if the plaid was coherent or sliding. (Some studies addressed possible percept alternations by instructing observers to report their "dominant percept", e.g., Stoner & Albright, 1996 or Lindsey & Todd, 1996, or to report "whether they saw pattern motion", Stoner et al., 1990). We can only speculate about the reason for the overwhelming preference for brief-presentation 2AFC methods, but several reasonable explanations come to mind. First, many studies that used plaids did not focus on integration versus segmentation issues (i.e., were not concerned whether the plaid looked coherent or transparent), but rather used plaids to study the integration process. Those studies focused on the perceived direction and/or speed of the plaid (Alais, Wenderoth, & Burke, 1994; Bowns, 1996; Derrington, Badcock, & Henning, 1993; Derrington, Badcock, & Holroyd, 1992; Ferrera & Wilson, 1991; Stone, Watson, & Mulligan, 1990; Welch, 1989; Wilson & Kim, 1994b; Yo & Wilson, 1992). Consequently, plaid parameters were chosen so that the coherent percept was dominant (but see Kooi, De Valois, Grosof, & De Valois, 1992a for a discussion about possible biases in these studies due to the presence of some sliding motion).

Another reason for the preference of the brief-presentation paradigm might have been a concern that the prolonged observation method would be more susceptible to adaptation, attentional effects, and/or eye movements. However, our results indicate that factors such as adaptation and attention do not, in fact, hinder the possibility of using the continual-report approach. Figs. 1 and 2 showed that there is no consistent within-trial adaptation, and that the probability of the coherent percept, $C/[C+T]$, was stable over durations as long as 5 min. Fig. 4 indicated that the between-observer mean lengths of the coherent and transparent epochs do not change significantly over time (except for the first coherent percept, or RTtransp). The constancy of the mean durations over time refutes the idea that adaptation accumulates over time in prolonged observations. [4,5] As for between-trial adaptation, although we did not test this directly, our results showed no evidence for it. The reliable dependence of $C/[C+T]$ and RTtransp on manipulated parameters is an indicator that the data do not suffer from systematic biases. The likeliest reason for the lack of between-trial adaptation effects in our experiments is the systematic randomization of stimuli over a large number of parametric conditions (in particular the randomization of the global direction). Finally, note that we are not claiming that adaptation effects cannot be observed in plaids: surely, that would not be correct. Our conclusion about adaptation is restricted to its effects (or lack thereof) in the paradigms used here. Since the purpose of this paper was not to study adaptation effects, we did not pursue directly how those might be induced in continual-report experiments.

5.1. Two dynamics-based measures: $C/[C+T]$ and *RTtransp*

Based on the continual-report data, we defined two measures of the relative strength of the coherent percept. The first measure is the probability to perceive coherency in prolonged observations, $C/[C+T]$, where C and T represent the cumulative time spent seeing coherency and transparency, respectively. An analogous measure, the relative cumulative time spent in each bi-stable percept, has been used extensively in binocular rivalry studies, and was shown to be systematically related to manipulations of the strength of the stimuli, e.g., via changes in contrast [6] (for reviews see Blake, 1989; Blake,

[4] von Grunau and Dubé (1993) reported a shortening of the perceptual epochs over time, but their conclusion was most likely due to a methodological problem in how they computed the average durations; see footnote 3, Section 4.

[5] There is some evidence that the dynamical behavior in other bi-stable domains may be different, showing increase (Brown, 1955; Long, Toppino, & Kostenbauder, 1983) or decrease (Lehky, 1995) of the alternation rate over time.

[6] In binocular rivalry, the mean durations of the percept have also been shown to be systematically related to the strength of the stimuli (Blake, 1989; Fox & Rasche, 1969; Lehky, 1988; Levelt, 1968; Logothetis, Leopold, & Sheinberg, 1996). Such a relation can be shown also for plaid stimuli, but it goes beyond the scope of the present paper.

2001; Blake & Logothetis, 2002; Lehky, 1988; Leopold & Logothetis, 1999; Levelt, 1968). For plaids, a similar measure has been used previously by von Grunau and Dubé (1993). However, our definition differs from theirs in an important way, because we systematically excluded the duration of the first percept from the calculation of $C/[C+T]$. Our data indicate that if this precaution is taken, $C/[C+T]$ is stable over time (cf. Fig. 2), justifying referring to it as the "steady-state" probability to perceive coherency. The exclusion of the first perceptual period was called for because of its different dynamical behavior: it was consistently longer than the subsequent periods.

The duration of the first coherency percept was used as a second measure of the relative strength of coherency. This measure, termed RTtransp ("the Response Time to see transparency"), is methodologically independent of $C/[C+T]$ yet showed a tight correlation with it and a similar dependency on parametric manipulations. RTtransp was used only for rectangular-wave plaids in Experiments I and II, where the first percept was always 'coherency' (cf. Experiment III). For those stimuli, RTtransp was not only longer than subsequent coherency epochs but, importantly, it could be very long even in cases where, after the first separation of the gratings, the transparent percept was more frequent. Interestingly, from Wallach's descriptions of bi-stability in his (rectangular-wave) plaids it seems that he was aware of the singularity of the first epoch: "A pattern of crossed lines... will be seen to move *for a long time* in the direction of its objective movement. With *prolonged inspection*, however, this motion will break up... two series of oblique lines are seen to move in opposite directions... This divided motion lasts only briefly and downward motion of the unified pattern returns. Now it is replaced *more quickly* by the horizontal motion phase. The two phases continue to alternate..." (Wallach, 1976, p. 212; our italics).

5.2. New results from the dynamics approach

The dynamics approach uncovered many unsuspected findings about motion integration and segmentation, beyond the observations about the dynamics of alternations summarized above. First, manipulating multiple variables in full-factorial designs provided quantitative estimates of the relative strength of different factors

- The parameter found to have the greatest effect was α, the angle between the gratings' direction of motion: increasing α dramatically decreased $C/[C+T]$ and RTtransp. This underscores the importance of the angular separation between motion signals generated by different objects as a primary cue for segmentation, and ecologically makes sense since independent objects tend to move in different directions.
- The effect of α was almost perfectly linear in most of the range of possible α values. This finding has important implications for models of motion integration and segmentation, since it suggests that there is not a "critical" value of α where the system switches from one interpretation to the other (see, e.g., Wilson & Kim, 1994a).
- The effect of α was independent of speed, indicating that its influence on motion integration and segmentation mechanisms is indeed mediated by a difference of direction, not by secondary effects on speed.
- Increasing the gratings' speed decreased coherency measures only within a small range of speed values (below 1.5°/s). This indicates that the neural mechanisms for motion integration and segmentation are insensitive to speed over a wide range.
- Duty cycle had a statistically significant, but very small effect on coherency (as long as it was below 50% so the that the figure/ground relationship of the gratings was unaltered; Stoner & Albright, 1996; Stoner et al., 1990).

In addition to these findings, which the paper described in details, the dynamics approach led to other observations that merit further examination.

- The global direction of the plaid is an important factor: plaids moving in oblique directions slide more easily than plaids moving in cardinal directions (Hupé & Rubin, 2001).
- Preliminary experiments (cf. Section 4.3) indicated that the asymmetry between coherency and transparency at stimulus onset was significantly reduced for stimuli for which the duty cycle was 50% (sinusoidal as well as square-wave plaids). The elimination of bias for seeing coherency first may be caused by the figure/ground ambiguity of the gratings composing these plaids.
- The alternations between coherency and transparency were perceptually much clearer for square/rectangular-wave plaids than for sinusoidal plaids, although the quantitative measures revealed a similar dependency on parametric manipulations. This finding may again be related to differences in figure/ground segmentation between the two types of plaids.
- The dynamics method can also shed light on the debated issue of the effect of the luminance of the plaid's intersections. Stoner and co-workers (1990,1996) showed that the probability of coherency was decreased when the intersections' luminance caused the gratings to look transparent (statically). Their results were obtained with a fixed value of α (135°), the 2AFC method being able to show the effect only within a small range of parameter space: Kim and Wilson (1993) repeated the experiment with a smaller value of α, 44°, and found that plaids were perceived as

coherent in 100% of the trials, regardless of the luminance of the intersections. Dynamics-based measure indicated however that the effect of intersections' luminance extends over a wide range of α values (unpublished observations; you can visit http://cns.nyu.edu/home/hupe/plaid_demo; see also Plummer & Ramachandran, 1993).

5.3. Comparing the dynamics approach with the brief-presentation 2AFC method

An important conclusion of this study is that the dynamics-based measures of the strength of coherency are more sensitive than those derived from brief presentation methods. This was demonstrated most directly for the effect of α, where $C/[C+T]$ and RTtransp showed a gradual, near-linear dependence on α over most of the range tested, while brief-presentation measures yielded a sigmoid-shape relationship (compare Figs. 7a, 8, 12, 13 and 16 with Fig. 14). The floor and ceiling effects exhibited in Fig. 14 represent a general problem of brief-presentation methods not limited to this case. For example, the observation of interactions between speed and α (Farid & Simoncelli, 1994; Farid et al., 1995) were likely to be biased by floor and ceiling effects. In any brief presentation experiment, an underlying variation in the strength of coherency will reveal itself only in parameter regimes where the duration of the first percept (either coherency or transparency) happens to be of the same order of magnitude as the trial duration. Elsewhere, the perception of the plaid during the brief trial will be dominated by the first percept (which is typically much longer), and therefore parametric manipulations will appear to have no effect. An example is the effect of the luminance of the plaid's intersections: the brief-presentation method revealed its underlying effect only in regimes where RTtransp happened to be close to the presentation duration ($\alpha \sim 135°$, Stoner et al., 1990; Stoner & Albright, 1996), and masked it elsewhere (Kim & Wilson, 1993).

Another important limitation of the brief-presentation method is that its measure of the probability of coherency depends on the arbitrary choice of a presentation duration: by slightly increasing or decreasing the trial duration one may shift the point of transition of the sigmoid-shaped curve which the methods yields. Furthermore, this manipulation would create a misleading impression that trial duration has an effect on coherency. We conjecture that the observation of Kooi et al. (1992a,b) that increasing the viewing time of plaids from 0.5 to 3 s increased the probability of the transparent percept may be caused by this methodological problem (rather than to adaptation mechanisms specific to the coherent percept, as proposed by the authors).

The dynamics-based measures are not immune to floor or ceiling effects, either. We have seen that, for very "slidy" stimuli, RTtransp values may become so short that they can be driven by response times limitations no less than by the balance between coherency and transparency. This led to observable floor effects (see discussion of Fig. 8 and http://cns.nyu.edu/home/hupe/plaid_demo/suppl.htm). Using the $C/[C+T]$ measures may not necessarily help in such cases, since plaids which start sliding so soon after stimulus onset often do not cohere for very long times (or never), making the method impractical. The other extreme, of very "sticky" stimuli, also presents problems for the dynamics approach. Even if such stimuli would slide after very long observation times, in reality it is not practical to expect trials to last more than a small number of minutes. Nevertheless, an advantage of the dynamics approach is that it allows to infer the presence of such floor/ceiling effects easily from the data (when $C/[C+T]$ asymptotes to 0 or 1).

5.4. Perceptual bi-stability in plaids: implications for physiological and modeling studies

The fact that plaid stimuli are bi-stable has important implications, both experimental and theoretical. Physiological studies often present stimuli for long durations (e.g., in electrophysiology, optical imaging or fMRI). This presents special challenges in the case of bi-stable stimuli, because the physical responses must, at some level, undergo alternations similar to those observed perceptually. Averaging cells' (or fMRI) responses over long durations, which may include more than one perceptual state, therefore becomes problematic, and could potentially mask important effects. One way to address this problem has been to use stimuli which are strongly coherent or strongly transparent, perceptually (Movshon et al., 1985; Rodman & Albright, 1989; Stoner & Albright, 1992). A potentially more powerful method is to collect behavioral data about the appearance of the stimulus in a continual-report paradigm, and look for correlation between the time-course of the perceptual alternations and that of physiological responses (Castelo-Branco et al., 1997), similarly to what has been done for binocular rivalry experiments (Leopold & Logothetis, 1996; Logothetis & Schall, 1989; Polonsky, Blake, Braun, & Heeger, 2000; Tong & Engel, 2001; Tong, Nakayama, Vaughan, & Kanwisher, 1998).

With regard to theories of motion integration and segmentation, the notion that for plaids either coherency (integration) or transparency (segmentation) "wins" promoted models that embedded a mechanism to "decide", or choose between the two possible interpretations of the stimulus. But the bi-stability of plaid perception suggests that another modeling approach may be more appropriate. In binocular rivalry, it is widely accepted that the bi-stability arises from active competition between the rivaling stimuli; in models, this competition is typically implemented via some form of reciprocal inhibition between the neural representations of the two percepts (see, e.g., Blake, 1989; Laing & Chow, 2002; Lehky, 1988). Transferring this approach to the domain of motion integration and segmentation would suggest an architecture where the neural representations of the 'coherent' and 'transparent' interpretations of the stimulus continually compete for dominance. This is a significant departure from present-day approaches to motion segmentation and integration, but one that may well advance our understanding of the underlying mechanisms.

Acknowledgements

We thank Dr. Robert Shapley for useful discussions. This work was supported by the Human Frontier Science Program (J.M. Hupé, LT0463/1999-B), the Sloan Foundation and the National Institute of Health (R01-EY14030-01).

References

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature, 300*, 523–525.

Alais, D., Wenderoth, P., & Burke, D. (1994). The contribution of one-dimensional motion mechanisms to the perceived direction of drifting plaids and their after effects. *Vision Research, 34*, 1823–1834.

Blake, R. (1989). A neural theory of binocular rivalry. *Psychological Review, 96*, 145–167.

Blake, R. (2001). A primer on binocular rivalry, including current controversies. *Brain and Mind, 2*, 5–38.

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience, 3*, 13–21.

Bowns, L. (1996). Evidence for a feature tracking explanation of why type II plaids move in the vector sum direction at short durations. *Vision Research, 36*, 3685–3694.

Braddick, O. (1993). Segmentation versus integration in visual motion processing. *Trends in Neurosciences, 16*, 263–268.

Brown, K. T. (1955). Rate of apparent change in a dynamic ambiguous figure as a function of observation-time. *American Journal of Psychology, 68*, 358–371.

Castelo-Branco, M., Goebel, R., Neuenschwander, S., Lanfermann, H., Zanella, F. E., & Singer, W. (1997). MT/MST activation depends on the interpretation of stimuli: a functional MRI study of the perception of plaids. *Society for Neuroscience Abstracts, 23*, 460.

Derrington, A. M., Badcock, D. R., & Henning, G. B. (1993). Discriminating the direction of second-order motion at short stimulus durations. *Vision Research, 33*, 1785–1794.

Derrington, A. M., Badcock, D. R., & Holroyd, S. A. (1992). Analysis of the motion of 2-dimensional patterns: evidence for a second-order process. *Vision Research, 32*, 699–707.

Farid, H., & Simoncelli, E. P. (1994). The perception of transparency in moving square-wave plaids. *Investigative Ophthalmology and Visual Science, 35*(Suppl.), 1271 (Abstract).

Farid, H., Simoncelli, E. P., Bravo, M. J., & Schrater, P. R. (1995). Effect of contrast and period on perceived coherency of moving square-wave plaids. *Investigative Ophthalmology and Visual Science, 36*(Suppl.), 51 (Abstract).

Ferrera, V. P., & Wilson, H. R. (1991). Perceived speed of moving two-dimensional patterns. *Vision Research, 31*, 877–893.

Fox, R., & Rasche, F. (1969). Binocular rivalry and reciprocal inhibition. *Perception and Psychophysics, 5*, 215–217.

Hupé, J. M., & Rubin, N. (2000). Perceived motion transparency can override luminance/color cues which are inconsistent with transparency. *Investigative Ophthalmology and Visual Science, 41*(Suppl.), 721 (Abstract).

Hupé, J. M., & Rubin, N. (2001). Transparent motion is always more likely for plaids moving along oblique directions than for plaids moving along cardinal directions. *Investigative Ophthalmology and Visual Science, 42*(Suppl.), 736 (Abstract).

Kim, J., & Wilson, H. R. (1993). Dependence of plaid motion coherence on component grating directions. *Vision Research, 33*, 2479–2489.

Kooi, F. L., De Valois, K. K., Grosof, D. H., & De Valois, R. L. (1992a). Properties of the recombination of one-dimensional motion signals into a pattern motion signal. *Perception and Psychophysics, 52*, 415–424.

Kooi, F. L., De Valois, K. K., Switkes, E., & Grosof, D. H. (1992b). Higher-order factors influencing the perception of sliding and coherence of a plaid. *Perception, 21*, 583–598.

Laing, C. R., & Chow, C. C. (2002). A spiking neuron model for binocular rivalry. *Journal of Computational Neuroscience, 12*, 39–53.

Lehky, S. R. (1988). An astable multivibrator model of binocular rivalry. *Perception, 17*, 215–228.

Lehky, S. R. (1995). Binocular rivalry is not chaotic. *Proceedings of the Royal Society of London. Series B, Biological Sciences, 259*, 71–76.

Leopold, D. A., & Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature, 379*, 549–553.

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

Levelt, W. J. M. (1968). *On binocular rivalry*. Paris, Mouton: The Hague.

Lindsey, D. T., & Todd, J. T. (1996). On the relative contributions of motion energy and transparency to the perception of moving plaids. *Vision Research, 36*, 207–222.

Logothetis, N. K., Leopold, D. A., & Sheinberg, D. L. (1996). What is rivalling during binocular rivalry? *Nature, 380*, 621–624.

Logothetis, N. K., & Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science, 245*, 761–763.

Long, G. M., Toppino, T. C., & Kostenbauder, J. F. (1983). As the cube turns: evidence for two processes in the perception of a dynamic reversible figure. *Perception and Psychophysics, 34*, 29–38.

Marr, D., & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London. Series B, Biological Sciences, 211*, 151–180.

Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattas, & C. Gross (Eds.), *Pattern recognition mechanisms* (pp. 117–151). Rome: Vatican Press.

Plummer, D. J., & Ramachandran, V. S. (1993). Perception of transparency in stationary and moving images. *Spatial Vision, 7*, 113–123.

Polonsky, A., Blake, R., Braun, J., & Heeger, D. J. (2000). Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nature Neuroscience, 3*, 1153–1159.

Rodman, H. R., & Albright, T. D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research, 75*, 53–64.

Rubin, E. (1921). *Visuell wahrgenommene figuren*. Copenhagen: Gyldendals.

Rubin, E. (1958). Figure and ground. In D. C. Beardslee & M. Wertheimer (Eds.), *Readings in perception* (pp. 194–203). Princeton, NJ: D. Van Nostrand.

Smith, A. T. (1992). Coherence of plaids comprising components of disparate spatial frequencies. *Vision Research, 32*, 393–397.

Stone, L. S., Watson, A. B., & Mulligan, J. B. (1990). Effect of contrast on the perceived direction of a moving plaid. *Vision Research, 30*, 1049–1067.

Stoner, G. R., & Albright, T. D. (1992). Neural correlates of perceptual motion coherence. *Nature, 358*, 412–414.

Stoner, G. R., & Albright, T. D. (1996). The interpretation of visual motion: evidence for surface segmentation mechanisms. *Vision Research, 36*, 1291–1310.

Stoner, G. R., Albright, T. D., & Ramachandran, V. S. (1990). Transparency and coherence in human motion perception. *Nature, 344*, 153–155.

Tong, F., & Engel, S. A. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. *Nature, 411*, 195–199.

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron, 21*, 753–759.

von Grunau, M., & Dubé, S. (1993). Ambiguous plaids: switching between coherence and transparency. *Spatial Vision, 7*, 199–211.

Wallach, H. (1935). Uber visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung, 20*, 325–380.

Wallach, H. (1976). *On perception*. New York: Quadrangle.

Welch, L. (1989). The perception of moving plaids reveals two motion-processing stages. *Nature, 337*, 734–736.

Wilson, H. R., & Kim, J. (1994a). A model for motion coherence and transparency. *Visual Neuroscience, 11*, 1205–1220.

Wilson, H. R., & Kim, J. (1994b). Perceived motion in the vector sum direction. *Vision Research, 34*, 1835–1842.

Wuerger, S., Shapley, R., & Rubin, N. (1996). On the visually perceived direction of motion by Hans Wallach: 60 years later. *Perception, 25*, 1317–1367.

Yo, C., & Wilson, H. R. (1992). Perceived direction of moving two-dimensional patterns depends on duration, contrast and eccentricity. *Vision Research, 32*, 135–147.