# Phase integration bias in a motion grouping task

**Jessica N. Cali**
Department of Psychology, Neuroscience, and Behaviour, McMaster University, Hamilton, Ontario, Canada

**Patrick J. Bennett**
Department of Psychology, Neuroscience, and Behaviour, McMaster University, Hamilton, Ontario, Canada

**Allison B. Sekuler**
Department of Psychology, Neuroscience, and Behaviour, McMaster University, Hamilton, Ontario, Canada
Rotman Research Institute, Baycrest Health Sciences, Toronto, Ontario, Canada
Department of Psychology, University of Toronto, Toronto, Ontario, Canada

**The perception of the direction of global motion depends on our ability to integrate local motion signals over space and time. We examined motion binding using a task requiring integration of relative phase. Observers completed multiple tasks involving clockwise and counter clockwise motion in a stimulus comprising four sets of linearly arranged dots, two moving horizontally and two moving vertically along sinusoidal trajectories differing in phase. Noise jitter was added along the trajectory perpendicular to each dot's motion. The noise acts as a global grouping cue that improves direction discrimination, but surprisingly, the absence of noise causes consistent below-chance performance (Lorenceau, 1996). We explore this phenomenon and subsequently test the hypothesis that observers perceive reverse motion because their representation of the relative phase of the motion components is systematically biased. We employ a number of different objective and subjective measures of motion integration and measure the phenomenon in both younger and older adults. Taken together, the results presented in the current article demonstrate that noise can promote global grouping in the stimulus and that confident, incorrect responses can be observed in the absence of correct global grouping. Generally, the current result raises the possibility that an integration bias could exist in other motion tasks.**

## Introduction

The integration of information across space and time is a fundamental aspect of motion perception. Such integration occurs over multiple spatial scales: For example, motion perception is thought to depend on the integration of information across multiple features on a single object, as well as the integration of information across multiple moving objects (e.g., Adelson & Movshon, 1982; Williams & Sekuler, 1984; Alais et al., 1998; Braddick, 1993; Snowden & Verstraten, 1999). The current article focuses on the integration of spatially separated sets of elements that have different motion trajectories.

We used a stimulus created by Lorenceau (1996) to study issues related to the integration of sinusoidal motion trajectories that differ in phase. First consider a stimulus consisting of four contours that are arranged to form a square and that occluders are placed over the square's four corners. When the square moves in a circular orbit around a central point, the presence of the occluders means that each horizontal and vertical contour undergoes sinusoidal motion in a direction that is perpendicular to its orientation. In other words, the contours move only along horizontal or vertical trajectories, and the direction of orbital motion (i.e., clockwise vs. counter clockwise) is conveyed by the relative phase of the horizontal and vertical components. Nevertheless, observers typically report seeing the square orbit the central point and are capable of discriminating clockwise and counter clockwise motion, which implies that observers somehow integrate the horizontal and vertical motions and are sensitive to their relative phase. To investigate this integration process, Lorenceau modified the moving square stimulus by replacing the four contours with four colinear sets of five evenly spaced, high-contrast dots, and each set of dots underwent sinusoidal motion in the direction perpendicular to its orientation. Hence,

Citation: Cali, J. N., Bennett, P. J., & Sekuler, A. B. (2020). Phase integration bias in a motion grouping task. *Journal of Vision*, *20*(7):31, 1–31, https://doi.org/10.1167/jov.20.7.31.

https://doi.org/10.1167/jov.20.7.31 Received November 5, 2019; published July 30, 2020 ISSN 1534-7362

like the original square stimulus, the modified stimulus contained motion only in the horizontal and vertical directions, and the relative phase of the horizontal and vertical components could be adjusted such that the motion in the modified stimulus was identical to that produced by clockwise or counter clockwise orbital motion in the original stimulus. Gestalt principles such as proximity and common fate would favor grouping the dots into four colinear sets of dots, whereas the perception of global motion requires grouping across at least two perpendicularly arranged sets of colinear dots. Lorenceau measured response accuracy in a task that required observers to discriminate clockwise and counter clockwise global motion. Across conditions, he varied the amount of dynamic, independent motion noise that was added to the trajectory of each individual dot. Surprisingly, he found that accuracy *improved* as the level of dynamic noise increased. However, he also found that accuracy was significantly *below* chance at long stimulus durations when the stimulus did not contain dynamic noise. That is to say, observers consistently misperceived the direction of motion in some conditions.

Previous studies examining this issue focused primarily on the question of whether the moving sets of elements were or were not integrated into a pre determined, correct motion percept (Lorenceau & Alais, 2001; Lorenceau & Shiffrar, 1992; McDermott & Adelson, 2004; Murray et al., 2001; Shiffrar & Pavel, 1991; Lorenceau & Zago, 1999). This forced-choice aspect of the experimental method may have missed some important characteristics of motion integration, especially in the case of illusory, below-chance direction discrimination. Therefore, the current experiments examined motion integration using a continuous measure of integration that allowed for the possibility that subjects perceived illusory or anomalous integrated motion.

Several aspects of motion perception change during normal healthy aging (Hutchinson et al., 2012). For example, minimum motion thresholds and motion coherence thresholds increase with age (Snowden & Kavanagh, 2006), meaning that older adults need faster speeds to accurately perceive coherent motion. Furthermore, older adults are less sensitive to global motion in random dot kinematograms and identify its direction less accurately (Bennett et al., 2007; Ball & Sekuler, 1986; Roudaia et al., 2010). In the current study, we test older adult observers to identify how integration of spatially separated moving elements changes with age.

In the current study, we conduct a series of experiments to explore motion integration by measuring the effects of local motion noise on the discrimination of global orbital motion using stimuli similar to those described by Lorenceau (1996). The first experiment replicated Lorenceau's findings, including the observation of below-chance performance with long-duration stimuli that lacked local motion noise, and extended these findings to a sample of older adult observers. Experiments 2 and 3 focus on factors that produce below-chance performance, examining different types of noise and cues that are hypothesized to affect integration. Experiment 4 presents a series of experiments designed to measure biases in phase integration in both younger and older adults. Finally, Experiment 5 extends the findings of the previous experiments by examining integration using a related stimulus that produces the chopstick illusion (Anstis, 2007).

# Experiment 1

## Method

### *Observers*

Twenty-four naïve observers—12 younger adults (19–27 years old, $M = 23$; 3 female) and 12 older adults (64–78 years, $M = 70$; 7 female)—participated in the experiment. All observers tested in the current study possessed normal or corrected-to-normal Snellen visual acuity. Observers received partial course credit or \$10 per hour for their participation. Each participant provided informed consent prior to the start of the experiment, and all experimental protocols used in the current article were approved by the McMaster University Research Ethics Board.

### *Apparatus and stimuli*

Stimuli were generated on an Apple iMac computer using MATLAB and the Psychophysics and Video Toolbox (Brainard, 1997; Pelli, 1997) and presented on an NEC MultiSync FE992 100-Hz monitor with a display size of 36 × 27 cm (35.6 pixels/cm). Average luminance was 30.5 cd/m$^2$. A chin and forehead rest was used to stabilize viewing position. Participants viewed the display binocularly through natural pupils from a viewing distance of 114 cm.

The stimulus consisted of four sets of colinear dots that were positioned along the sides of a virtual square 3.7 deg in diameter (Figure 1; see Supplementary Materials for video). Each set contained five dots (diameter = 0.05 deg, luminance = 95.6 cd/m$^2$) that were evenly spaced and spanned 3.2 deg. The dots were white and presented on a uniform gray background (luminance = 43.9 cd/m$^2$, Weber contrast = 2.14). The sets of dots moved sinusoidally, and the parameters of motion were derived in the following way. Consider a square that is moved in the picture plane so that its center orbits a stationary, central fixation point with a rotation amplitude of 0.40 deg of visual angle and a rotation frequency of 0.83 Hz. Such orbital motion can be decomposed into horizontal and vertical sinusoidal components: The speed of rotation is determined by

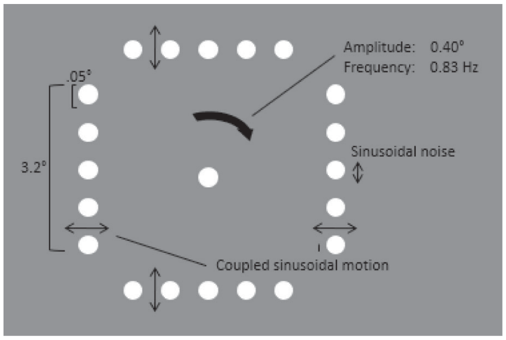



Figure 1. Square stimulus used in Experiments 1, 2, and 4. Each white dot was 0.05 deg of visual angle. The stimulus consisted of 20 dots arranged into four sets of five dots. Each set of dots spanned 3.2 deg. The large arrows depict the main motion components: The two horizontal lines of dots moved vertically along a sinusoid, while the two vertical lines moved horizontally on a sinusoid 90 deg off phase. The small arrows depict the local motion noise: The path of each dot was displaced sinusoidally in the direction perpendicular to the main motion component. The central dot was a fixation point. On each trial, the intial position of the set of dots was randomly chosen to center around a point on the central orbital trajectory. As such, the stimulus was not centered on the fixation point.

the frequency of the sinusoidal components, and the direction of the square's orbit (i.e., clockwise vs. counter clockwise) is determined by their relative phase. In our stimuli, the horizontal and vertical motion components were applied to the vertical and horizontal sets of dots, respectively (Figure 1). Because the local motion of the dots was horizontal or vertical, the direction of global, orbital motion (clockwise vs. counter clockwise) that was used to construct the components could be determined only by noting the relative phase of the two component motions. In addition to the main sinusoidal motion components, independent motion noise was added to each dot. The noise varied the location of a dot sinusoidally in the direction that was perpendicular to the main motion. The frequency of the sinusoidal noise was 3 Hz and the starting phase was selected randomly for each dot on each trial. Each dot's motion therefore would be composed of the main component motion—vertical motion for horizontal dots and horizontal motion for vertical dots—and random phase sinusoidal motion in the orthogonal direction. Noise amplitude, which varied across trials, was 0, 0.027, or 0.081 deg.

### *Procedure and experimental design*

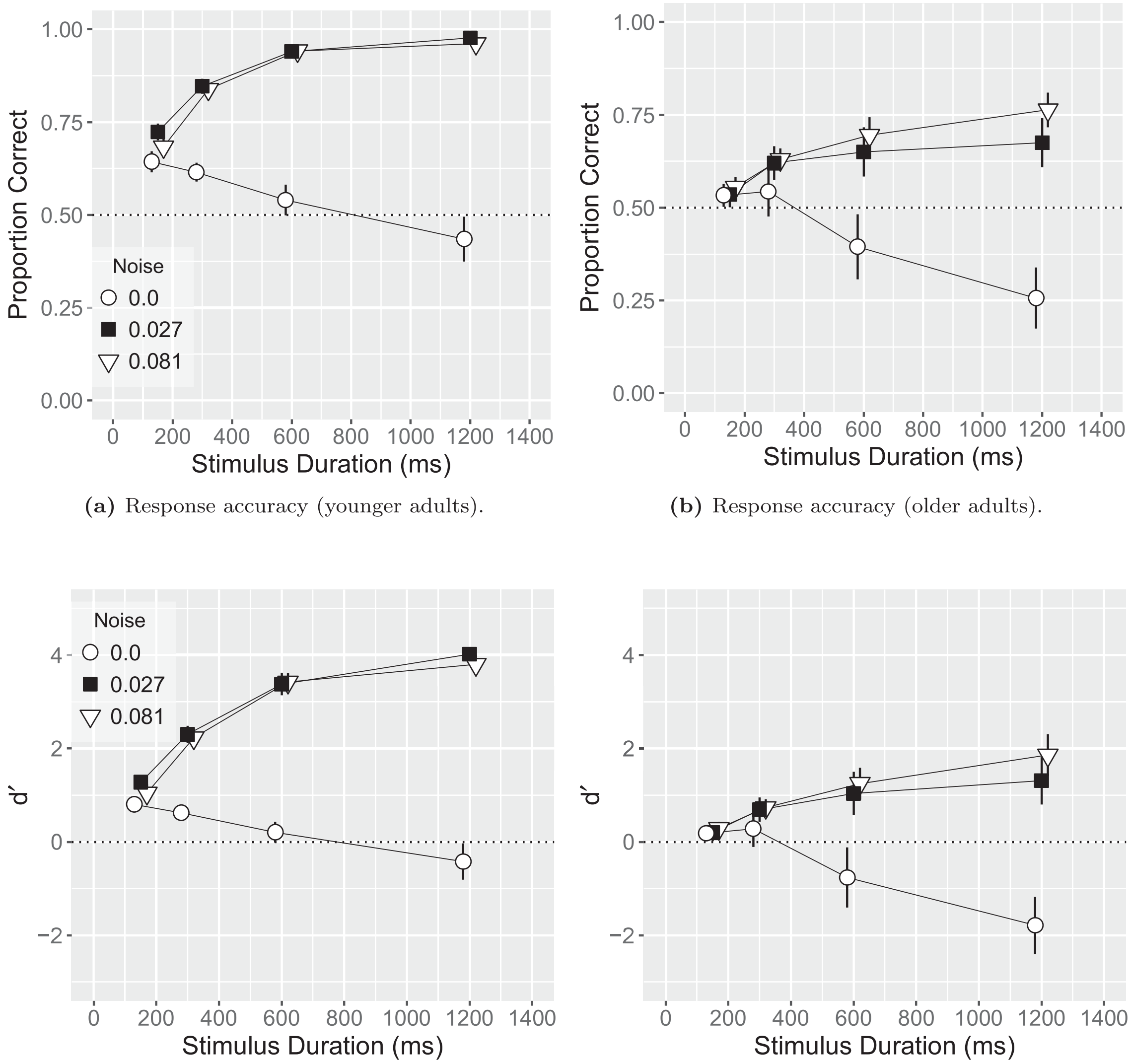
Each trial consisted of the presentation of a central fixation point and a single moving stimulus, followed by a blank, uniform screen. The subject's task was to report whether the direction of global, orbital motion was counter clockwise or clockwise by pressing one of two labeled keys on a standard computer keyboard. An experimental session began with 10 practice trials in which the stimulus was a square consisting of four high-contrast lines that orbited the fixation point for 600 ms. The 10 practice trials were repeated until a subject responded correctly on at least 8 trials. The practice trials ensured that subjects could discriminate clockwise and counter clockwise motion, as well as understood the task and how to use the keyboard to respond. The experimental trials, which immediately followed the practice trials, used the dot stimulus shown in Figure 1. There were four stimulus durations (150, 300, 600, and 1,200 ms) and three motion noise amplitudes (0, 0.027, and 0.081 deg of visual angle) for a total of 12 conditions. With our motion parameters, it took 1,205 ms for the stimulus to go completely around the fixation point; therefore, the stimulus did not complete an orbit of the fixation point at all stimulus durations. There were 50 trials per condition, or 600 trials in total. Experimental trials were blocked by stimulus duration; motion noise levels were randomly intermixed within each block. Each block was self-paced, and an experimental session lasted approximately 50 minutes. No feedback was given during the experiment. Each trial began 1,500 ms following the response to the previous trial.

## Results

All statistical analyses in the current article were performed in R (R Core Team, 2017). In all experiments, the Huynh-Feldt estimate of sphericity ($\tilde{\epsilon}$) was used to adjust $p$ values of $F$ tests conducted on within-subject variables with more than 1 degree of freedom, and either adjusted $R$-squared ($R^2_{adj}$) or Cohen's $d$ is reported as a measure of association strength and effect size. Proportion correct data were arcsine transformed to correct for their nonnormal distribution.

Response accuracy and sensitivity ($d'$) are plotted as a function of stimulus duration and noise amplitude for both age groups in Figure 2. On average, response accuracy and sensitivity were higher in younger compared to older subjects, but the effects of noise and duration on both measures were qualitatively similar in both age groups. For example, when the noise amplitude was greater than zero, response accuracy and sensitivity in both age groups increased with increasing stimulus duration; however, in the zero-noise condition, response accuracy and sensitivity *decreased* with increasing stimulus duration in both age groups.

We analyzed $d'$ and arcsine-transformed proportion correct with separate 2 (age) × 4 (duration) × 3 (noise) analysis of variance (ANOVA). The two ANOVAs yielded very similar results, and therefore we report only the results of the ANOVA on proportion correct. The main effects of age, noise, and duration were significant, as were all of the two-way interactions between age, noise, and duration (Table 1).




**(a)** Response accuracy (younger adults). **(b)** Response accuracy (older adults).

**(c)** Sensitivity (younger adults). **(d)** Sensitivity (older adults).

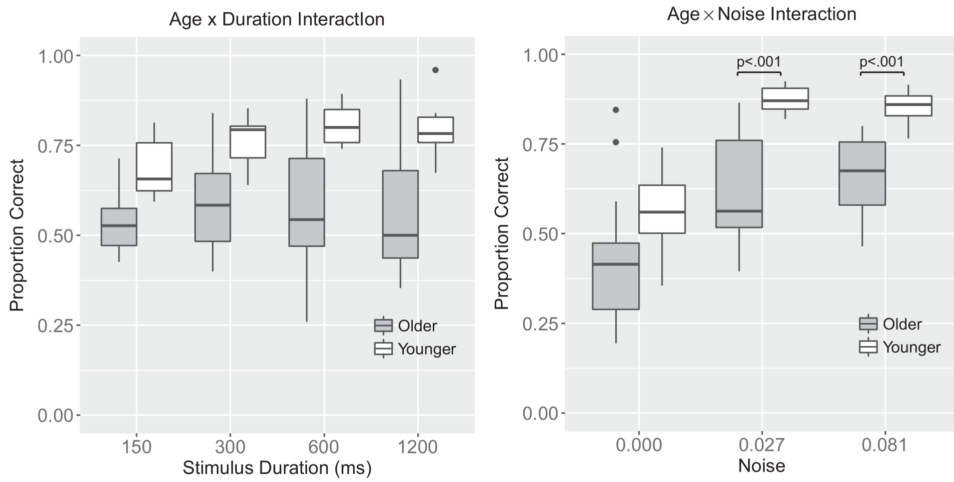
Figure 2. Response accuracy and sensitivity ($d'$) measured in Experiment 1 plotted as a function of stimulus duration and motion noise amplitude for younger (a, c) and older (b, d) participants. Error bars represent ±1 *SEM*.

| **Effect** | ***df*** | $\tilde{\epsilon}$ | ***MSE*** | ***F*** | $\eta_p^2$ | $p_{adj}$ |
|---|---|---|---|---|---|---|
| Age | 1, 22 | – | 0.30 | 21.05 | .30 | 0.0001 |
| Noise | 2, 44 | 0.754 | 0.06 | 106.84 | .30 | <0.0001 |
| Duration | 3, 66 | 0.624 | 0.06 | 15.87 | .08 | <0.0001 |
| Age × Noise | 2, 44 | 0.754 | 0.06 | 11.98 | .05 | 0.0005 |
| Age × Duration | 3, 66 | 0.624 | 0.06 | 4.60 | .02 | 0.02 |
| Noise × Duration | 6, 132 | 0.666 | 0.03 | 46.34 | .18 | <0.0001 |
| Age × Noise × Duration | 6, 132 | 0.666 | 0.03 | 2.22 | .01 | 0.09 |

Table 1. Experiment 1 ANOVA table.

The Noise × Duration interaction reflects the fact that accuracy increased with increasing stimulus duration in the non-zero-noise conditions but decreased with increasing duration in the zero-noise condition (Figure 2). The linear trend of response accuracy across stimulus durations was positive and significantly different from zero in the 0.027-deg noise ($t(23) = 5.58$, $p < 0.001$, $d = 1.61$) and 0.081-deg noise ($t(23) = 8.66$, $p < 0.001$, $d = 2.50$) conditions, whereas the linear trend was negative and significantly different from zero in the zero-noise condition ($t(23) = 5.36$, $p < 0.001$, $d = 1.54$).

The Age × Duration and Age × Noise interactions are illustrated in Figure 3. The Age × Duration interaction reflects the fact that the effect of stimulus duration on response accuracy (after averaging across noise levels) was significant in younger participants ($F(3, 33) = 10.56$, $\tilde{\epsilon} = 0.743$, $p_{adj} = 0.0003$) but not older participants ($F(3, 33) = 1.219$, $\tilde{\epsilon} = 0.555$, $p_{adj} = 0.318$). The Age × Noise interaction was analyzed by comparing accuracy, after averaging across stimulus durations, between age groups: Welch two-sample *t*




(a) Experiment 1 Age × Duration interaction.

(b) Experiment 1 Age × Noise interaction.

Figure 3. Illustration of the Age × Duration and Age × Noise interactions obtained in Experiment 1. (a) Boxplots show accuracy, after averaging across noise levels, for each stimulus duration and age group. The effect of duration was significant in younger, but not older, participants. (b) Boxplots illustrate response accuracy, after averaging across stimulus duration, at each noise level and age group. The effect of noise was significant in both age groups, though it was larger in younger participants. The between-group difference was significant in the two non-zero-noise conditions but not in the zero-noise condition. In both age groups, accuracy in the 0.027 and 0.081 conditions differed from the zero-noise condition but not from each other.

tests found that the age difference was significant in the non-zero-noise conditions (*noise* = 0.027: $t(13.72) = 5.576$, $p < 0.0001$, $d = 2.28$; *noise* = 0.081: $t(17.07) = 6.116$, $p < 0.0001$, $d = 2.50$) but not in the zero-noise condition ($t(16.90) = 1.623$, $p = 0.122$, $d = 0.66$).[1] We also examined this interaction with paired-sample *t* tests comparing accuracy across noise levels within each age group. For younger participants, accuracy was 0.30 higher in the 0.027 condition than the zero-noise condition, a difference that was statistically significant ($t(11) = 9.51$, $p < 0.001$, $d = 2.75$). For older adult participants, the accuracy difference between noise conditions was smaller (i.e., 0.19), though still statistically significant ($t(11) = 5.34$, $p = 0.002$, $d = 1.54$). In both age groups, accuracy in the two non-zero-noise conditions did not differ significantly. Hence, increasing noise from 0 to 0.027 improved response accuracy in both age groups, but the effect was larger in younger adults than older adults.

Interestingly, mean response accuracy was below chance in some of the zero-noise conditions. That is to say, observers were, on average, reporting that the stimulus was moving in the direction opposite to the veridical direction of global motion. We used *t* tests to determine if mean accuracy in three zero-noise conditions—stimulus durations of 600 and 1,200 ms for older adults, and 1,200 ms for younger adults—was significantly lower than chance (i.e., 0.5). Mean accuracy only for older adults in the 1,200-ms condition was significantly lower than 0.5 ($t(11) = 2.943$, one-tailed $p = 0.006$, $d = .850$); however, close inspection of the data suggested that several subjects in both age groups had accuracies that were significantly below chance. To test this idea further, we used Bonferroni-corrected ($\alpha_{FW} = .01$) binomial tests to determine which of the 24 subjects had a response accuracy that was significantly below 0.5 in the zero-noise, 1,200-ms condition. This method identified 13 participants (3 younger and 10 older) with response accuracies that were significantly below 0.5, and the mean proportion correct for these subjects was 0.146. Thus, the data suggest that a few younger participants, and nearly all older participants, responded at below-chance levels in that condition.

Accuracy at the 150-ms stimulus duration was near chance for younger (mean = .64) and older (mean = .53) adults. Proportion correct was significantly different from chance for younger observers ($t(11) = 5.014$, $p < .001$, $d = 2.047$) but not older observers ($t(11) = 1.086$, $p = .301$, $d = .443$). These analyses suggest that younger adults were able to acurately determine direction when viewing the stimulus for very short durations, while older adults were not.

## Discussion

We found that the effect of stimulus duration on response accuracy in a global motion discrimination task depends on the presence of local motion noise. When the local motion noise was present, accuracy increased with increasing stimulus duration. However, when local noise was absent, accuracy *decreased* with increasing duration (Figure 2). Indeed, at the longest stimulus duration (1,200 ms), mean accuracy in older adults (and some younger adults) in the zero-noise condition was below chance. These findings replicate results reported by Lorenceau (1996). We also found that older adults were generally less accurate than younger adults and that this age difference was larger for long-stimulus durations and higher levels of noise; however, the interaction between the effects of noise and stimulus duration did not vary significantly between age groups.

Lorenceau (1996) suggested that the addition of stimulus noise reduces the tendency to organize the stimulus into two sets of parallel, collinear arrangements of dots and therefore makes it more likely that the dots are grouped into a single, coherently moving object. According to this logic, when there is more evidence that the stimulus should be grouped globally and less evidence the stimulus should be grouped locally, global motion direction is perceived more accurately. This hypothesis is consistent with the results of several studies (e.g., Lorenceau & Shiffrar, 1992; Lorenceau, 1996; Shiffrar & Lorenceau, 1996; Lorenceau & Shiffrar, 1999), as well as the results of Experiment 1. Performance is similar in the noise-present and noise-absent conditions at short durations. However, because the effect of the noise builds over time, performance in the two conditions becomes increasingly dissimilar at longer stimulus durations.

Like Lorenceau (1996), we found that the response accuracy in some individuals was below chance in zero-noise conditions with relatively long-stimulus durations. Because all participants successfully completed a block of practice trials, and because below-chance performance was found only in two conditions, it is unlikely that the below-chance performance occurred because participants were confused about the mapping of the response keys or misunderstood the instructions. Rather, participants presumably perceived global motion in the direction opposite to the veridical direction. In the following two experiments, we focus on testing this phenomenon.

# Experiment 2

Experiment 2 investigated two issues. The first issue concerned the nature of the position noise used by Lorenceau (1996) and in Experiment 1. Specifically, we examined whether the observed effects of the noise are specific to sinusoidal position jitter or whether they could be elicited by other types of noise. Experiment 2 therefore included a condition that embedded the stimulus in a dynamic white noise background. The second issue concerned the clarity of the global motion percept in conditions that yielded below-chance performance in some observers. In Experiment 2, while reporting the direction of global motion, observers also reported whether they felt the motion was "definitely," "probably," or "maybe" moving in the reported direction.

## Method

### *Observers*

Twelve naïve observers (17–23 years, $M = 19$ years, 10 female) participated in the experiment.

### *Stimuli*

The experimental apparatus was the same as in Experiment 1. The stimuli also were the same as those used in Experiment 1, except for the following changes. Experiment 2 used only three levels of stimulus motion noise: 0, 0.0135, and 0.027. We did not include a noise amplitude of 0.081 because Experiment 1 found that performance was unaffected by increasing noise beyond 0.027. In addition, we added a condition where the uniform $5.22 \times 5.22$-deg background was replaced with a dynamic Gaussian white noise consisting of 0.05 $\times$ 0.05-deg square pixels. On each video frame, the contrast of each noise pixel was selected randomly and independently from a zero-mean Gaussian distribution with a standard deviation of 0.11. Finally, stimulus duration was fixed at 600 ms. We chose 600 ms because mean accuracy at this duration in Experiment 1 was near chance with zero-noise stimuli, and therefore there was a reasonable chance of seeing performance increase or decrease in conditions that used noise.

### *Procedure*

The procedure was similar to the one used in Experiment 1, except that it was modified to allow participants to register the confidence of their direction judgment on each trial. Specifically, the response screen that immediately followed the stimulus offset prompted observers to respond whether the stimulus "definitely," "probably," or "maybe" moved clockwise or counter clockwise. We refer to these three response alternatives as representing high, moderate, and low confidence, respectively.

The two types of background (uniform vs. dynamic noise) were crossed factorially with three levels of local motion noise (0, 0.0135, and 0.027) to yield a total of six experimental conditions. There were 74 trials per condition, yielding a total of 370 trials. All trial types were intermixed. Prior to the experimental trials,

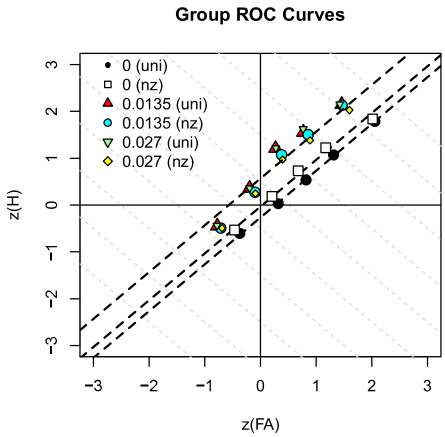



Figure 4. Group ROC data measured with uniform (uni) and noise (nz) backgrounds and local motion noise amplitudes of 0, 0.0135, and 0.027. Each point represents the mean *z*-transformed false alarms and hits for 12 subjects. The dotted lines represent the best-fitting equal-variance Gaussian ROC (i.e., the slope was fixed at 1.0). The ROC cures for the non-zero-motion noise conditions were very similar, so only a single ROC, estimated from the average across those four conditions, is drawn. The light gray lines with a slope of –1 represent constant criterion lines (i.e., $c = \frac{1}{2} \times (z(H) + z(FA))$).

all participants completed the same practice session that was used in Experiment 1. The experiment took approximately 50 min to complete.

## Results

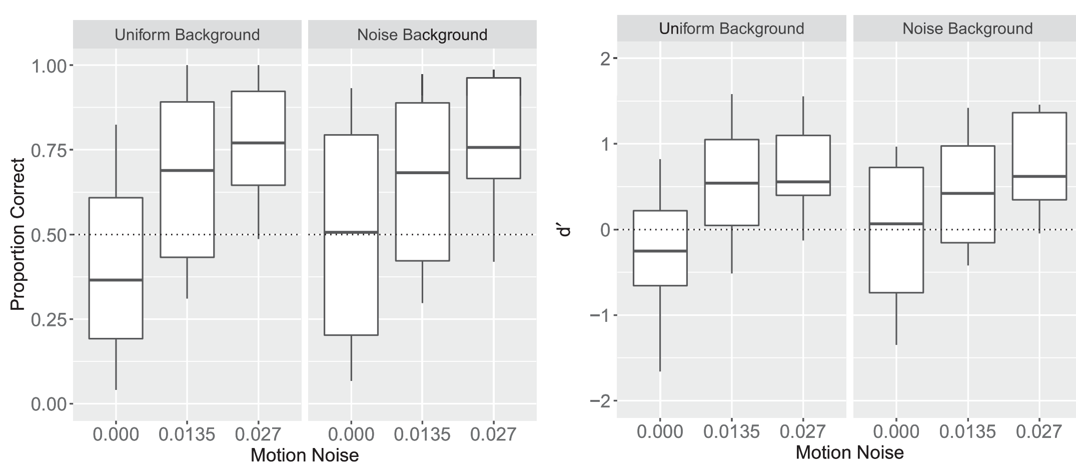
Our analyses address two main questions: (a) How does background noise affect performance in this task compared to motion noise? (b) How confident are observers when they make incorrect and correct decisions?

### *Response accuracy and sensitivity*

Group receiver operating characteristic (ROC) curves were constructed from all participant data and are displayed in Figure 4. The slopes of the ROCs were approximately 1, which demonstrates the validity of using the equal-variance Gaussian model (Wickens, 2002) for calculating $d'$ from these data. The curves also highlight the differences in average performance between the motion noise levels, specifically when the stimulus included motion noise (0.0135 and 0.027) and when it did not (zero motion noise).

Response accuracy, ignoring confidence ratings, and $d'$, calculated from confidence ratings from each observer using the equal-variance Gaussian model (Wickens, 2002), are plotted as a function of motion noise and background type in Figure 5. Accuracy and sensitivity increased with increasing motion noise with both uniform and dynamic noise backgrounds. In addition, performance measured with non-zero-motion noise was similar with uniform and noise backgrounds, but performance in the zero-motion noise condition was poorer with a uniform background than with a dynamic noise background. Finally, in the zero-noise, uniform, background condition, median response accuracy was below chance and $d'$ was less than zero.

We first determined if participants in Experiment 2 performed similarly to younger adult participants in Experiment 1 by comparing performance in the zero-noise, uniform-background condition in Experiment 2 to performance in the zero-noise, 600-ms condition in Experiment 1. Mean response accuracy in the zero-noise, uniform-background condition did not




**(a)** Response Accuracy **(b)** Sensitivity ($d'$)

Figure 5. Proportion correct (a) and sensitivity (b) for all conditions in Experiment 2. Sensitivity was estimated from confidence ratings using the equal-variance Gaussian model (Wickens, 2002).

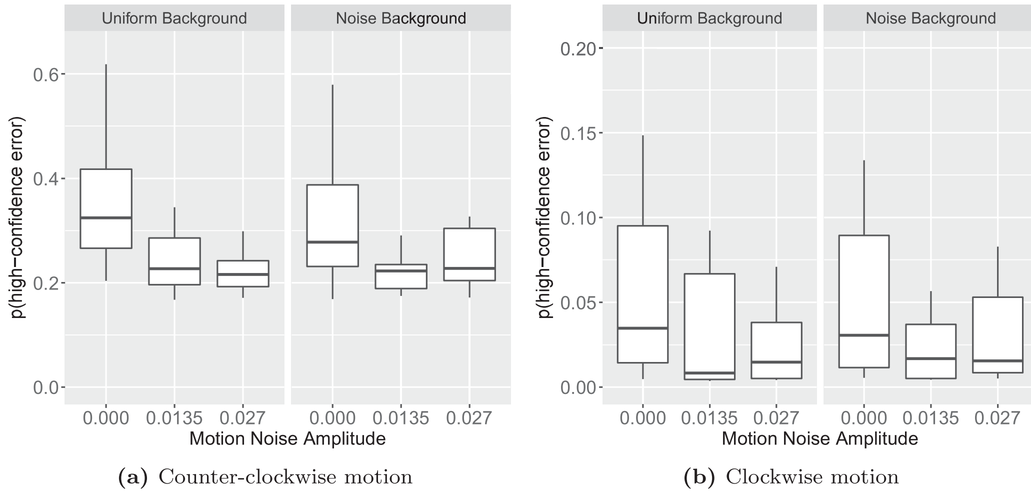



Figure 6. Proportion of high-confidence, incorrect responses obtained with (a) counter clockwise and (b) clockwise global stimulus motion. Note that the scale of the y-axis differs in the two panels.

| **Effect** | ***df*** | $\tilde{\epsilon}$ | ***MSE*** | ***F*** | $\eta_p^2$ | $p_{adj}$ |
|---|---|---|---|---|---|---|
| Noise | 2, 22 | 1 | 0.02 | 57.71 | .21 | <0.0001 |
| Background | 1, 11 | - | 0.01 | 2.14 | .002 | 0.17 |
| Noise × Background | 2, 22 | 1 | 0.01 | 9.05 | .01 | 0.001 |

Table 2. Experiment 2 ANOVA table.

differ significantly from chance in Experiment 2 ($M = 0.406$, $t(11) = -0.699$, $p = 0.498$, $d = -0.51$), nor did it differ from the mean response accuracy measured in the comparable condition in Experiment 1 ($t(17.33) = -1.399$, $p = 0.179$, $d = -0.63$). Similarly, sensitivity ($d'$) in the zero-noise, uniform-background condition did not differ significantly from zero in Experiment 2 ($M = -0.304$, $t(11) = -1.462$, $p = 0.172$, $d = -0.60$), nor did it differ significantly from mean sensitivity measured in the comparable condition in Experiment 1 ($t(21.9) = -1.672$, $p = 0.109$, $d = -0.68$). These results suggest that participants performed similarly in the comparable conditions in Experiments 1 and 2.

Arcsine-transformed proportion correct data, as well as $d'$, were analyzed with a 3 (motion noise) × 2 (background) within-subject ANOVA. Essentially identical results were obtained with both ANOVAs, so only the results of the ANOVA performed on accuracy are described here. There was a significant main effect of motion noise that was modulated by a significant Motion Noise × Background interaction (Table 2). The interaction reflected the fact that the linear trend of accuracy across levels of motion noise was larger with a uniform background ($t(11) = 9.23$, $p < 0.0001$, $d = 3.77$) than with a dynamic noise background ($t(11) = 4.50$, $p = 0.0009$, $d = 1.84$).

We also analyzed the Motion Noise × Background interaction by comparing the effect of background at each level of motion noise. We found that the difference between accuracy with uniform and dynamic noise backgrounds was significant in the zero-motion noise conditions ($t(11) = 3.36$, $p = 0.006$, $d = .97$), but did not differ with motion noise of 0.0135 ($t(11) = -0.816$, $p = 0.432$, $d = .24$) or 0.027 ($t(11) = -0.069$, $p = 0.946$, $d = .02$). These analyses suggest that the effect of a dynamic background noise on discrimination accuracy was greatest when the stimulus did not include local motion noise.

#### *High-confidence errors*

Figure 6 shows the proportion of high-confidence ratings given an incorrect response plotted as a function of local motion noise, background (uniform vs. noise), and the direction of global motion. Overall, more high-confidence errors were made for counter clockwise than clockwise global motion. However, for both counter clockwise and clockwise stimulus motion, the median proportion of high-confidence errors was higher in the absence of local motion noise. In addition, the effect of motion noise was, to a first approximation, similar in the uniform- and noise-background conditions.

To examine these trends quantitatively, we analyzed the arcsine-transformed data in Figure 6a,b with separate 3 (motion noise) × 2 (background) within-subject ANOVAs. With counter clockwise motion, the main effects of local motion noise ($F(2, 22) = 14.568$, $\eta_p^2 = 0.29$, $\tilde{\epsilon} = 0.508$, $p_{adj} = 0.0027$) and background ($F(1, 11) = 5.089$, $\eta_p^2 = 0.003$, $p = 0.0454$) were significant, as was the Noise × Background interaction ($F(2, 22) = 7.380$, $\eta_p^2 = 0.02$, $\tilde{\epsilon} = 0.925$, $p_{adj} = 0.0045$). The interaction was significant because the proportion of errors was higher in the uniform-background condition compared to the noise-background condition when the local motion noise amplitude was zero ($t(11) = 2.723$, $p = 0.0193$, $d = 0.79$) or 0.0135 ($t(11) = 2.076$, $p = 0.062$, $d = 0.60$), but it was lower in the uniform-background condition when the motion noise amplitude was 0.027 ($t(11) = -3.426$, $p = 0.006$, $d = 0.99$). However, in both background conditions, more high-confidence errors were made in the zero-motion noise condition than in the 0.0135 motion noise ($t(11) = 3.854$, $p = 0.008$, $d = 1.11$) or 0.027 motion noise ($t(11) = 3.791$, $p = 0.008$, $d = 1.09$) conditions.

With clockwise stimulus motion, the ANOVA yielded a significant main effect of local motion noise amplitude ($F(2, 22) = 9.137$, $\eta_p^2 = 0.11$, $\tilde{\epsilon} = 0.527$, $p_{adj} = 0.010$). The main effect of background ($F(1, 11) = 1.903$, $\eta_p^2 = 0.003$, $p = 0.195$) and the Noise × Background interaction ($F(2, 22) = 2.288$, $\eta_p^2 = 0.01$, $\tilde{\epsilon} = 0.773$, $p_{adj} = 0.140$) were not significant. Pairwise comparisons across the three levels of local motion noise, after averaging across the two background conditions, indicated that the proportion of high-confidence errors was higher in the zero-noise amplitude condition compared to the 0.0135 ($t(11) = 3.369$, $p = 0.0063$, $d = 0.97$) and 0.027 ($t(11) = 2.824$, $p = 0.0166$, $d = 0.82$) conditions, but the 0.0135 and 0.027 conditions did not differ from each other ($t(11) = -0.269$, $p = 0.793$, $d = -0.08$).

These analyses suggest that high-confidence errors occurred significantly more frequently when there was no local motion noise. In addition, the effects of motion noise were similar, though not identical, in the uniform- and noise-background conditions, which means that high-confidence errors in our experiment were more affected by local motion noise than by the nature of the background.

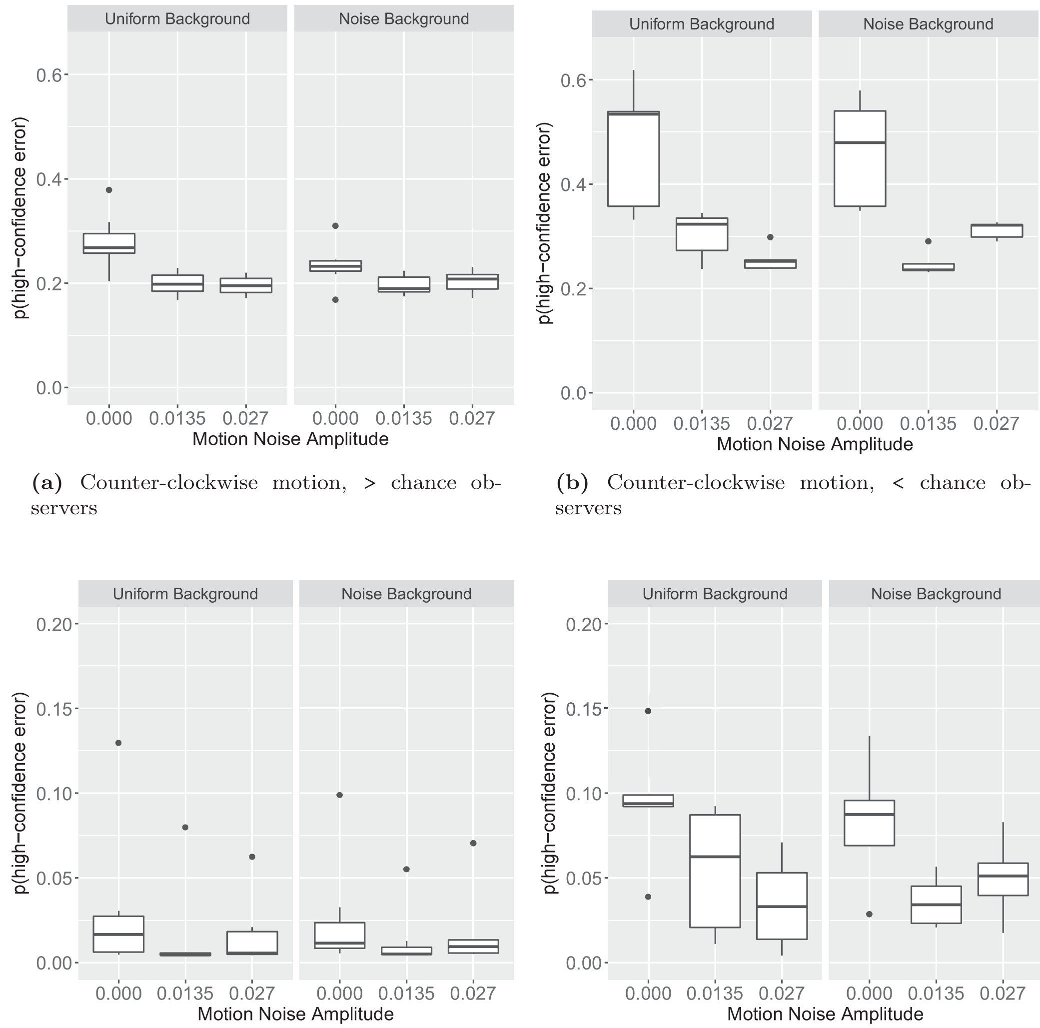
Experiment 1 suggested that some, but not all, younger observers consistently report seeing global orbital motion in the direction opposite to the veridical motion. We wondered whether the effects of local and/or background noise shown in Figure 6 differed between observers who were or were not prone to seeing global motion in the incorrect direction. Therefore, we divided participants into two groups depending on whether their overall accuracy across all conditions was above or below chance performance (0.50): Five participants produced overall accuracy scores lower than 0.50, while the remaining seven participants produced overall accuracy above 0.50. Figure 7 shows the proportion of high-confidence, incorrect responses as a function of local motion noise, background, the direction of global motion, and observer performance. Overall, the low-performing observers made more high-confidence errors compared to the high-performing observers, especially when the stimulus contained no local motion noise.

We analyzed the data in Figure 7 using a linear contrast that tested whether the arcsine-transformed proportion of high-confidence errors in the zero-noise condition differed from the mean proportion of high-confidence errors in the 0.0135 and 0.027 noise conditions. With counter clockwise global motion, the proportion of errors in the zero-noise condition was higher than in the other two noise conditions ($F(1, 10) = 21.99$, $p = 0.001$, $\eta_p^2 = 0.69$), and this effect of motion noise differed significantly between the low- and high-performance groups ($F(1, 10) = 7.86$, $p = 0.018$, $\eta_p^2 = 0.44$). Similar results were obtained for clockwise global motion: The proportion of errors was greater in the zero-noise stimulus condition than in the 0.0135 and 0.0270 noise conditions ($F(1, 10) = 12.686$, $p = 0.005$, $\eta_p^2 = 0.559$); however, although the difference in the linear contrast between the low- and high-performance groups was in the correct direction, it was not statistically significant ($F(1, 10) = 4.689$, $p = 0.055$, $\eta_p^2 = .319$). Taken together, the analyses provide evidence that the effects illustrated in Figure 6 differed for low- and high-performance observers.

## Discussion

The data from Experiment 2 clarify and extend the findings of Experiment 1. We found that dynamic background noise affected global direction discrimination only when the stimulus did not contain local motion noise. Recall that we manipulated the stimulus background to determine if the addition of white noise acted similarly to local motion noise and therefore increased discrimination accuracy. Our results were consistent with our predictions: Adding background white noise to stimuli that had zero local motion noise significantly increased response accuracy from approximately 40% correct (i.e., below chance) to approximately 50% correct (i.e., chance) in the global motion discrimination task (see Figure 5). This result raises the possibility that the local motion noise used here is not particularly special but instead is just one of several types of noise that leads to higher performance in this task.

The second major finding of Experiment 2 concerns the confidence of observers when they make errors. Several observers in Experiment 1 consistently performed below chance in some conditions, which suggests that they perceived the direction of the global orbital motion incorrectly. By analyzing the confidence ratings to incorrect discrimination, Experiment 2 showed that observers often are *confident* in their incorrect responses, and this effect was larger in observers who were more likely to perceive incorrect global motion. These analyses suggest that the below-chance performance found in some conditions




(a) Counter-clockwise motion, > chance observers

(b) Counter-clockwise motion, < chance observers

(c) Clockwise motion, > chance observers

(d) Clockwise motion, < chance observers

Figure 7. Proportion of high-confidence, incorrect responses obtained with (a, b) counter clockwise and (c, d) clockwise global stimulus motion in Experiment 2. Results for seven participants with above-chance accuracy are shown in a and c; results for five observers with below-chance accuracy are shown in b and d. Note that the scale on the y-axis differs between figures in the top and bottom rows.

represents a genuine mis perception of global motion direction rather than an effect of response bias and/or guessing on trials in which no clear direction was perceived.

## Experiment 3

Experiment 3 examined whether the illusion of opposite motion is due to differences in attention in noise-absent and noise-present stimuli. According to this hypothesis, when the stimulus is grouped (i.e., when noise is present), attention is allocated to the entire stimulus as one object, whereas when it is not grouped (i.e., when noise is absent), attention is allocated to only one motion component. Importantly, attending to only one motion component could cause the motion of the second component to be computed in reference to the attended component, as is found in the Duncker illusion (Zivotofsky, 2004) and the reference repulsion effect (Rauber & Treue, 1998). To test this idea, we changed the color of a single dot in each of two sets of collinear dots from white to red. The red dots were placed either within the same motion component (i.e., on parallel sets of dots) or different motion components (i.e., on perpendicular sets of dots). According to saliency models of attention (Itti & Koch, 2000), the red dots should draw visual attention to those locations and therefore bias observers to attending to the corresponding motion components. If attending to a single motion component is causing the illusion of opposite motion in the zero-noise condition, then increasing the saliency of a single motion component should decrease response accuracy in the noise-present condition. On the other hand, increasing the saliency of dots positioned on different motion components should make it more likely that observers will attend to both motion components and therefore increase response accuracy in the noise-absent condition.

## Method

### Observers

Five observers participated in the experiment (20–34 years, $M = 24$ years, five female). To increase the likelihood that the fixation instructions would be followed reliably, Experiment 3 used only experienced psychophysical observers. One observer was one of the authors and the other four were experienced psychophysical observers who were naïve to the purpose of the experiment.

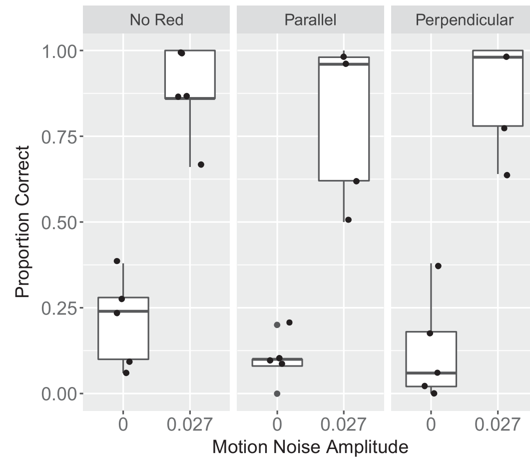



Figure 8. The effects of cues on response accuracy in Experiment 3. The bars indicate the mean cueing effect for each noise level and each cue type. The symbols show accuracy for individual participants. Error bars represent ±1 *SEM*.

### Stimuli

The experimental apparatus was the same as in Experiments 1 and 2. The stimuli were similar to those used in Experiments 1 and 2 with the following exceptions. Experiment 3 used only two levels of local motion noise: 0 and 0.027 dva, and the stimuli were presented only on a uniform gray background for a duration of 1,200 ms. There were three stimulus conditions. In the no-cue condition, the stimulus consisted entirely of white dots, as in the prior experiments. In the perpendicular-cue condition, the stimulus contained two red dots, one replacing the second dot from the left on the top set of dots, and the other replacing the second dot from the top on the right set of dots. Each red dot was easily discriminated from the neighboring white dots, and therefore this manipulation potentially provides an attentional cue that could promote integration of the two different motion components associated with the two sets of dots. Stimuli in the parallel-cue condition contained two red dots that replaced the second dot from the left in the top and bottom sets of dots. If attention to particular motion components is important in this task, this stimulus would encourage local grouping, because the red dots provide an attentional cue to group the two parallel lines that are part of the same motion component. A video of the parallel and perpendicular cue conditions is provided in the Supplementary Materials.

### Procedure

Observers completed three blocks of 100 trials. As in previous experiments, each trial began with the presentation of a fixation point followed by the stimulus. Participants judged whether the global orbital motion was clockwise or counter clockwise, and the next trial began 1,500 ms after the response. No feedback was given. Observers were instructed to maintain central fixation throughout the entire experiment. Half of the trials contained stimuli with zero-motion noise and half with 0.027-dva motion noise. Also, half of the trials contained clockwise motion and half counter clockwise motion. Observers completed the no-cue condition first, followed by the perpendicular-cue condition and then the parallel condition. Unlike Experiments 1 and 2, observers did not complete a practice session. The experiment lasted approximately 15 min.

## Results

Response accuracy is displayed for all conditions in Figure 8. Response accuracy was significantly greater than chance in all noise-present conditions (all $p < 0.041$, $d > 1.33$) and was significantly less than chance in all zero-noise conditions (all $p < 0.006$, $d > 2.18$). A 3 (cue) × 2 (noise) within-subjects ANOVA on arcsine-transformed accuracy revealed a significant main effect of noise ($F(1, 4) = 20.5$, $p = 0.01$, $\eta^2_G = 0.77$). The main effect of cue ($F(2, 8) = 2.67$, $p_{adj} = 0.16$, $\eta^2_P = 0.03$) and the Cue × Noise interaction ($F(2, 8) = 1.06$, $p_{adj} = 0.38$, $\eta^2_G = 0.01$) were not significant. The effect of cue also was not significant when the noise-present ($F(2, 8) = 1.56$, $p_{adj} = 0.28$, $\eta^2_P = 0.28$) and noise-absent ($F(2, 8) = 2.99$, $p_{adj} = 0.13$, $\eta^2_P = 0.43$) conditions were analyzed separately.

## Discussion

Experiment 3 was designed to test whether manipulating the saliency of one or both motion components using visual cues would change response accuracy relative to a no-cue condition. We found no evidence that the cues influenced accuracy: In all cue conditions, accuracy was well above chance in the noise-present conditions and well below chance in the noise-absent conditions. The attention cue was presumed to orient attention to particular motion components (Itti & Koch, 2000); however, we did not attempt to measure the degree of attention change between conditions. Assuming the cue was successful, these findings demonstrate that the perceived global motion was not affected significantly by cues that draw attention to a one or both motion components and reduces the likelihood that the illusory motion seen in zero-noise stimuli was caused by observers covertly attending to one of the two motion components. Motion velocity has been shown to be computed in reference to figures rather than backgound (Johansson, 1950). In the present stimulus, we sugest that observers might be using the stationary fixation point as a reference to compute the motion of the dots, rather than one of the two motion components.

# Experiment 4

Experiment 4 focuses on the factors that produce below-chance performance. We assume that the perception of the global, orbital motion in the Lorenceau stimulus requires observers to encode the relative phase of the horizontal and vertical motion components and that a misperception of the direction of motion occurs when the phase is encoded incorrectly. The fact that Experiments 1, 2, and 3 found that some observers *consistently* misperceived the direction of motion in some conditions suggests that there was a consistent error, or bias, in the perceived relative phase of the two components. To test our hypothesis, we measured the magnitude of each observer's phase bias on every trial and measured the association between bias and their direction discrimination responses made to the same stimuli. Specifically, we asked users to adjust the relative phase of the horizontal and vertical motion components until the stimulus appeared to be rotating clockwise or counter clockwise (specified at the beginning of each trial). We calculated the difference between the observer's setting and the actual relative phase that was required to produce the desired motion. This method is presented in three separate groups of observers. In Experiment 4a, we test younger adults in a high (0.027 deg) and zero-noise stimuli. In Experiment 4b, we again test younger observers, but using medium noise (0.0135 deg) and zero-noise stimuli. In Experiment 4c, we build on our findings in Experiment 1, testing older adult observers in high- and zero-noise stimuli.

## Method

### *Observers*

Thirty observers between the ages of 18 and 29 ($M$ = 20 years; 22 female) participated in Experiment 4a. Fifteen observers between the ages of 18 and 21 ($M$ = 19 years; 11 female) participated in Experiment 4b. Fifteen observers between the ages of 62 and 85 ($M$ = 73 years; 9 female) took part in Experiment 4c. All participants were naïve to the purposes of the experiment.

### *Apparatus and stimuli*

Apparatus was identical to Experiment 1. The stimuli were similar to Experiment 1, but the relative phase of motion was not fixed to ±90 deg in all conditions. When the four sets of dots orbit the central fixation point, clockwise orbital motion produces horizontal motion (of the vertical sets of dots) that leads the vertical motion (of the horizontal dots) by 90 deg, whereas counter clockwise motion produces horizontal motion that lags the vertical motion by 90 deg. Relative phases between 0 and ±90 deg produce clockwise/counter clockwise motion that is elliptical rather than circular: Reducing relative phase from ±90 deg to 0 deg results in progressively narrower elliptical motion, and a relative phase of 0 deg corresponds to diagonal motion. In this experiment, we measured the direction of global motion that was perceived when the relative phase of the horizontal and vertical motions was set to different values in a phase adjustment task. Because there was no actual orbital motion in the stimulus, the horizontal and vertical motions were, individually, uninformative, and therefore the direction judgment could be based only on the relative phase of the two motion components. As in Experiment 1, the amplitude and frequency of the sinusoidal motions were set to values that corresponded to a global orbital rotation amplitude of 0.4 deg and a frequency of 0.83 Hz, and the trajectory of each dot was perturbed by sinusoidal jitter, or noise, in the direction that was orthogonal to the main motion. The amplitude of this sinusoidal motion jitter was either 0 or 0.027 deg in Experiments 4a and 4c, and 0 or 0.0135 deg in Experiment 4b. The frequency was 3 Hz, and the starting phase was selected randomly for each dot on each trial.

### *Procedure and experimental design*

Sensitivity to relative phase was measured with an adjustment task and a discrimination task. In the adjustment task, each trial began with the instruction "please adjust the stimulus until it is clockwise/counter clockwise," which was presented in the center of the display for 2 s. The stimulus was presented immediately after the instruction was removed. On each trial, observers adjusted relative phase by turning a knob until they perceived global motion in the pre-specified direction (i.e., clockwise or counter clockwise). Observers were informed that rotating the knob varied relative phase over a limited range (i.e., ±720 deg, or ±2 cycles) around the starting value. Specifically, observers were told that if they felt they had hit a "wall" in stimulus space, they should continue exploring the stimulus by turning the dial in the other direction. Observers were encouraged to adjust the stimulus to achieve the most compelling percept of the target global motion possible and, when they were satisfied, to press the space bar on a computer keyboard to end the trial. On 50% of the trials, the initial setting of relative phase was correct: It was set to the value (−90 or 90 deg) that corresponded to the target direction of global, orbital motion. On the remaining 50% of the trials, the starting phase was set to uninformative values of 0 or 180 deg. When the starting relative phase was 0 deg, the four sets of dots formed a square that moved coherently along a diagonal path from the lower left to the upper right of the stimulus display. When the starting relative phase was 180 deg, the four sets of dots moved coherently along a diagonal path from the lower right to the upper left of the display. See Supplementary Materials for a video demonstration of phase adjustment. For each initial value of relative phase, on half of the trials, the adjustments made by the observer altered the phase of the horizontal motion (of the vertical dots), and on the other half, the adjustments varied the vertical motion (of the horizontal dots). There were a total of 96 trials. On 48 trials, the trajectories of the individual dots were perturbed by sinusoidal noise (amplitude = 0.027 deg or 0.0135 deg), and on the remaining 48 trials, the stimulus did not contain noise. Noise and no-noise stimuli, as well as all initial values of relative phase, were presented in a random order. In Experiment 4c, the target direction and trial number were written at the top of the response screen in the adjustment task for Participants 8 to 15 based on feedback given by Participants 1 to 7. This change was made to reduce memory demands during the trial and to provide feedback about the time remaining in the experiment.

In the discrimination task, on each trial, the relative phase was set to produce either clockwise or counter clockwise circular motion (i.e., ±90 deg) and observers reported whether they perceived global motion in the clockwise or counter clockwise direction. Observers completed 100 trials, 50 with local, sinusoidal noise and 50 with no noise. For each type of noise, the global motion was clockwise on half of the trials and counter clockwise on the other half. The type of noise and direction of global motion were randomly intermixed. Stimulus duration was 600 ms and was followed by a response screen that contained six buttons that indicated three levels of confidence—"maybe," "probably," and "definitely"—that the global motion was in the clockwise or counter clockwise direction. Participants indicated their response by selecting one button with a computer mouse.

The inter trial interval in both tasks was 1 s. All participants completed the adjustment task first, followed by the discrimination task.

## Experiment 4a

### *Results*

*Adjustment task.* Data are reported as phase adjustment error from the value that produces the desired target motion. An error of 0 means the stimulus was adjusted to the correct value. Errors *between* $\pm\pi/2$ mean that phase was set to values that produce elliptical (rather than circular) global motion in the correct direction. An error of exactly $\pi/2$ or $-\pi/2$ means that relative phase was set to a value that produces diagonal global motion. Finally, errors between $\pi/2$ and $\pi$ or between $-\pi/2$ and $-\pi$ mean that phase was set to a value that produces elliptical global motion in the incorrect direction. The sign of the phase error was not informative. Consequently, the absolute value was taken for all responses for the purpose of analysis. Therefore, the ability for an observer to produce either perfectly clockwise or perfectly counter clockwise motion can be taken by the distance between their phase error on a particlar trial and either 0 or $\pi$, whereas their response is generally accurate if the phase error is less than $\pi/2$. The 20% trimmed mean was calculated to estimate the average phase error in each condition for each observer.[2]

Figure 9 displays data from several representative observers. In each figure, the adjustment errors from individual trials are divided into four sets depending on whether the stimulus did or did not contain local noise and whether the initial global motion was in the correct direction or was ambiguous. The results from most observers resembled the results from Observers 5 and 11, who responded correctly in the noise condition but incorrectly in the no-noise condition. Indeed, most of the errors on no-noise trials fell between $\pi/2$ and $\pi$, or between $-\pi/2$ and $-\pi$, which is consistent with the hypothesis that observers perceived the global motion in the wrong direction (as in Experiment 1). Observers 7 and 25 did not perform differently on the noise and no-noise trials: The adjustment errors suggest that Observer 7 nearly always perceived global motion in the direction opposite to the true motion, whereas

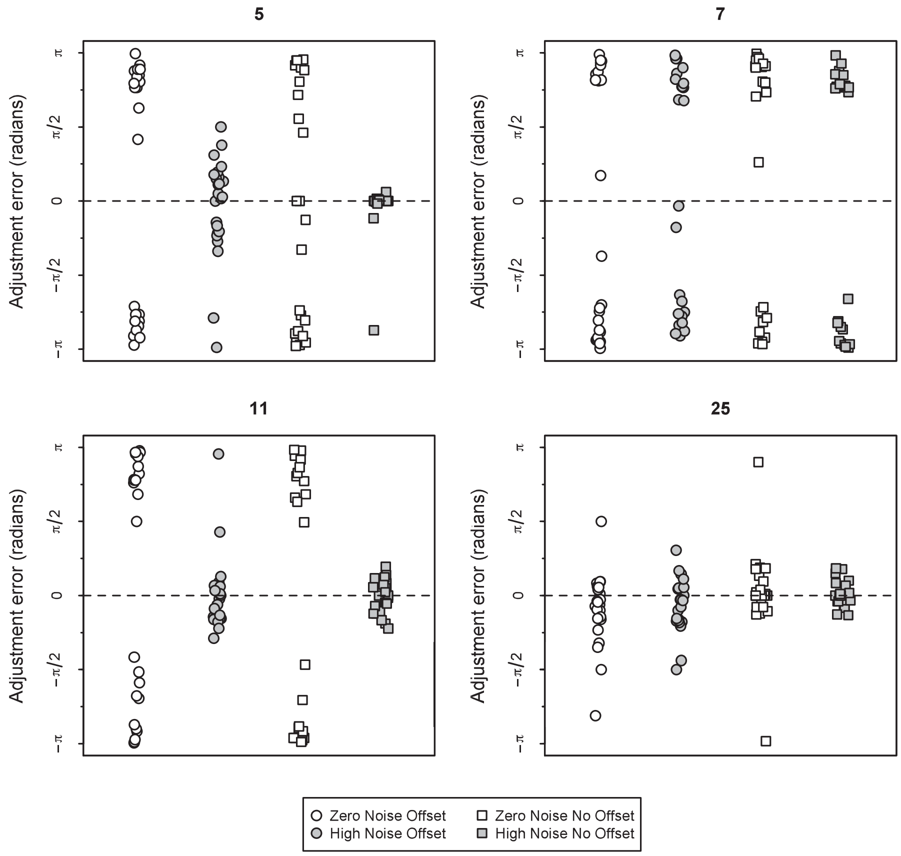



Figure 9. Phase adjustment data from four representative observers in Experiment 4a. Each symbol corresponds to a phase setting from a single trial. Results from high-noise and zero-noise trials are represented by filled and unfilled symbols, respectively. The square and circle symbols indicate results from, respectively, trials in which the initial global motion was in the correct direction (i.e., no offset; relative phase was −90 or 90 deg) or was in an ambiguous direction (i.e., offset; relative phase was 0 or 180 deg).

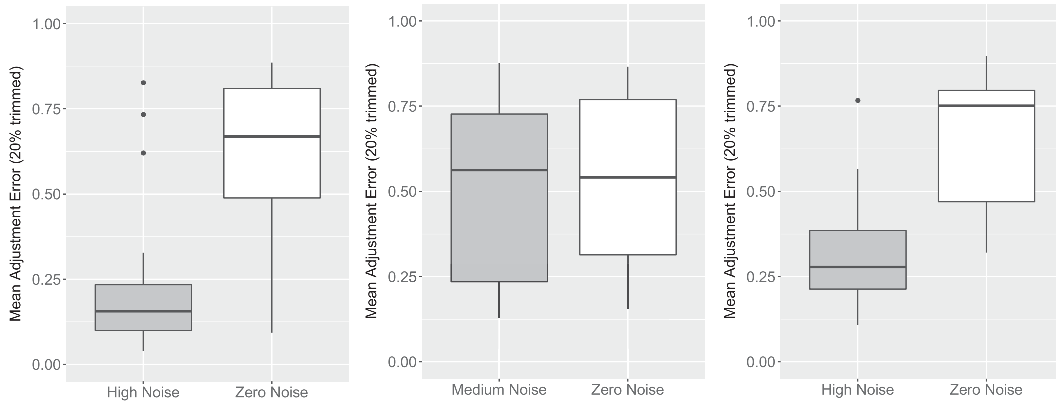
Observer 25 nearly always correctly perceived the direction of global motion.

The boxplots in Figure 10a illustrate the distributions of trimmed mean phase error in the zero-noise and high-noise conditions in $\pi$ radians. Average phase error was greater in no-noise trials than high-noise trials, ($t(29) = 9.31$, $p < 0.0001$, $d = 1.70$). The possible range of phase errors spans from 0 to 1 $\pi$ radians. If an observer was responding randomly, a uniform distribution of responses would be expected, leading to an average phase error of 0.50 $\pi$ radians. Therefore, to determine if observers perceived illusory motion, $t$ tests compared average phase error in the zero-noise and high-noise conditions to 0.50 $\pi$ radians. One-tailed $t$ tests indicated that phase error was significantly greater than 0.50 $\pi$ in the zero-noise condition ($t(29) = 3.28$, $p < 0.003$, $d = 0.60$) and significantly less than 0.50 $\pi$ in the high-noise condition ($t(29) = 8.18$, $p < 0.0001$, $d = 1.10$).

*Discrimination task.* The proportion of correct responses in the discrimination task, collapsing across all confidence ratings, is shown in Figure 11a. The difference between proportion correct on high-noise ($M = 0.838$) and no-noise ($M = 0.239$) trials was significant ($t(29) = 12.37$, $p < 0.0001$, $d = 2.25$). Furthermore, accuracy on zero-noise trials was significantly below 0.50, which represents chance performance ($t(29) = 6.55$, $p < 0.0001$, $d = 1.20$), whereas accuracy on high-noise trials was significantly above 0.5 ($t(29) = 7.46$, $p < 0.0001$, $d = 3.38$).

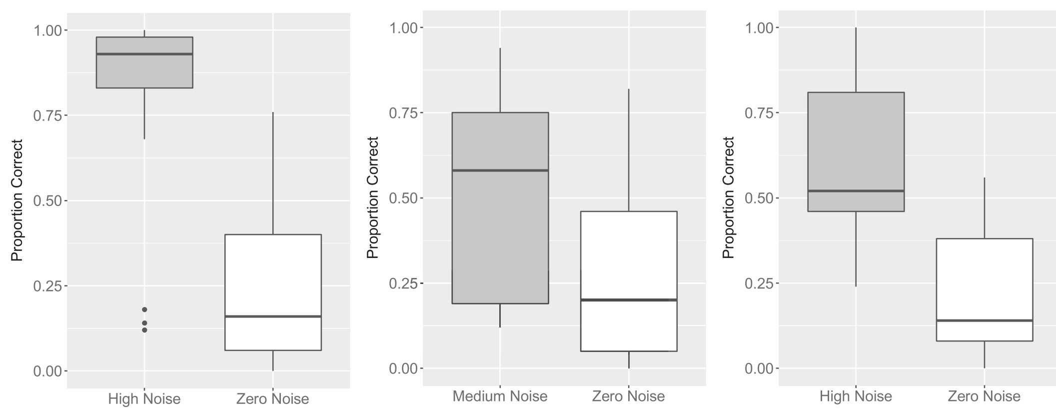
Figure 12a displays the average proportion of high-confidence responses (i.e., observers selected "definitely" on the response screen) that were *incorrect*. Three participants were removed from this analysis because they did not make at least one high-confidence response in both noise levels. The mean proportions in the zero-noise ($M = 0.15$) and high-noise ($M = 0.76$) conditions differed significantly ($t(26) = 7.42$, $p < 0.0001$, $d = 11.42$).

*Relating adjustment responses and discrimination accuracy.* The association between performance in the adjustment and discrimination tasks is depicted in Figure 13a. The linear association between response accuracy in the discrimination task and mean phase error in adjustment task was significant in both the high-noise ($b = -1.19$, $t(28) = -13.21$, $p < 0.001$,




(a) Experiment 4a (b) Experiment 4b (c) Experiment 4c

Figure 10. Boxplots illustrating the distributions of average (i.e., 20% trimmed mean) phase error across observers in Experiment 4. In each plot, phase error is plotted in units of $\pi$ (i.e., $1 = \pi$ radians). In the noise-present condition, the stimulus contained local, sinusoidal jitter with an amplitude of 0.027 (Experiments 4a and 4c) or 0.0135 (Experiment 4b). Young adults were tested in Experiment 4a ($M_{age} = 20$; $n = 30$) and Experiment 4b ($M_{age} = 19$; $n = 15$); older adults were tested in Experiment 4c ($M_{age} = 73$; $n = 15$). Perfect performance corresponds to 0 phase error. Phase error of 1 (i.e., $\pi$ radians) corresponds to the opposite direction of global motion.



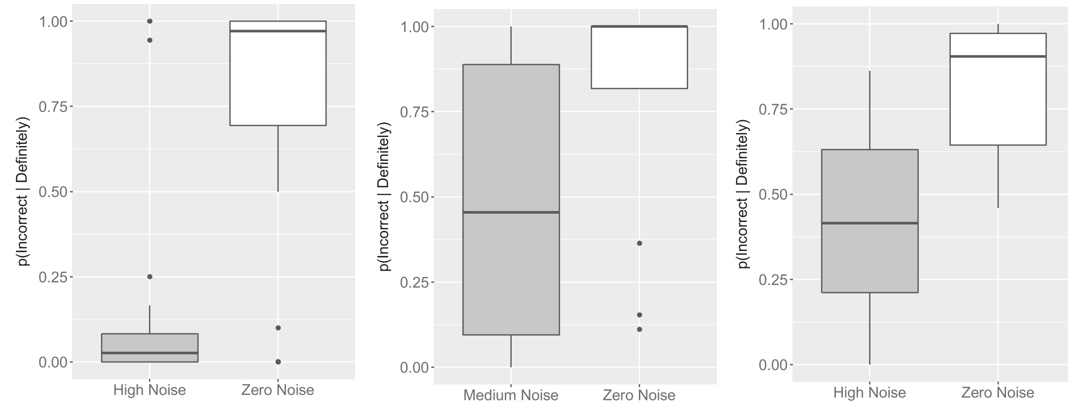

(a) Experiment 4a (b) Experiment 4b (c) Experiment 4c

Figure 11. Proportion correct in the discrimination task, ignoring confidence ratings. The amplitude of local, sinusoidal jitter was 0.0135 and 0.027 deg in the medium- and high-noise conditions, respectively. Young adults were tested in Experiment 4a ($M_{age} = 20$; $n = 30$) and Experiment 4b ($M_{age} = 19$; $n = 15$); older adults were tested in Experiment 4c ($M_{age} = 73$; $n = 15$). Chance performance corresponds to an accuracy of 0.5.

$R^2_{adj} = 0.85$) and zero-noise ($b = -0.76$, $t(28) = -5.57$, $p < 0.001$, $R^2_{adj} = 0.51$) conditions.

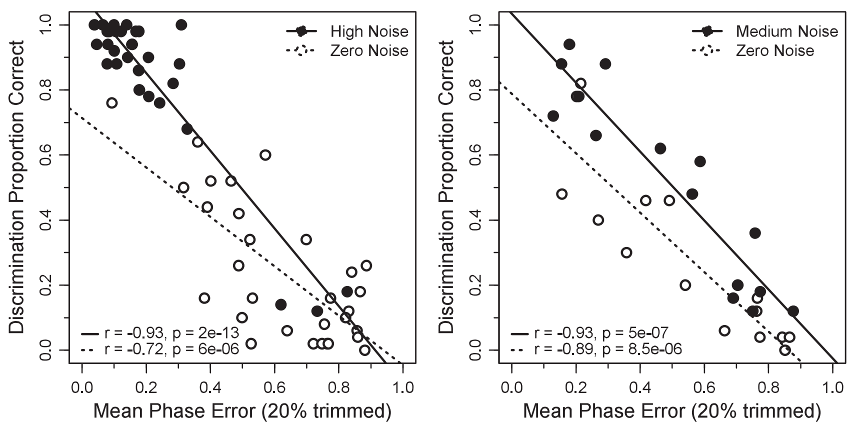
### *Discussion*

Experiment 4a examined the relationship between the relative phase of two motion trajectories and the perceived direction of global orbital motion. Using an adjustment task, we found that some participants set the relative phase of motion components to a value that was more than $\pi/2$ radians away from the correct value for the target motion. Importantly, these same participants performed below chance in the direction discrimination task when the stimulus did not contain local motion noise. Furthermore, observers were more likely to make errors in the discrimination task with high confidence when the stimulus did not contain noise. These results can be taken as strong evidence that most observers perceived motion in the direction opposite to the veridical direction when the stimulus did not contain local dynamic noise, and some observers misperceived the direction of motion even when the stimulus did contain noise. The regression in Figure 13a shows that there was a strong inverse relationship between phase adjustment error and discrimination

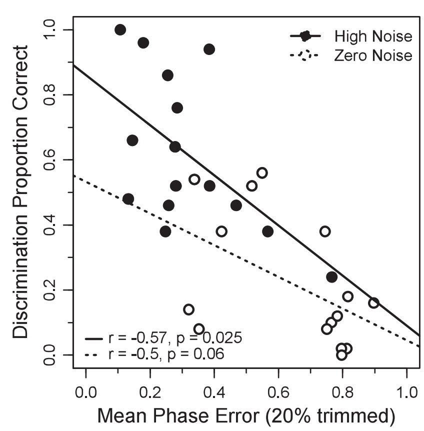



(a) Experiment 4a (b) Experiment 4b (c) Experiment 4c

Figure 12. The proportion high-confidence responses that were incorrect (i.e., $p$(Incorrect | Definitely)) observed in the discrimination tasks of Experiment 4. The amplitude of local, sinusoidal jitter was 0.0135 and 0.027 deg in the medium- and high-noise conditions, respectively. Young adults were tested in Experiment 4a ($M_{age} = 20$; $n = 30$) and Experiment 4b ($M_{age} = 19$; $n = 15$). Older adults were tested in Experiment 4c ($M_{age} = 73$; $n = 15$).




(a) Experiment 4a (b) Experiment 4b




(c) Experiment 4c

Figure 13. Relationship between response accuracy in the discrimination task and the average phase error in the adjustment task. Pearson's $r$ and the $p$ value for each regression are shown in the lower-left corner in each plot.

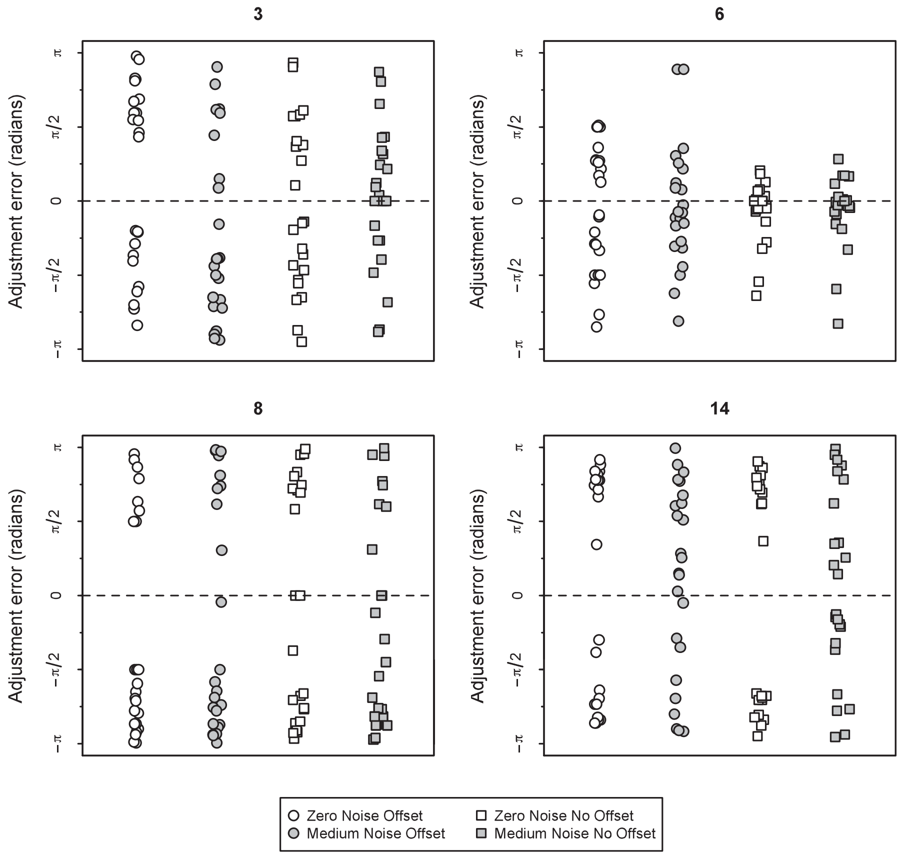



Figure 14. Phase adjustment data from four representative observers in Experiment 4b. Symbol conventions are the same as in Figure 9.

error accuracy both in the presence and absence of local dynamic noise.

Lorenceau (1996) suggested that local stimulus noise improved global motion discrimination by reducing the motion coupling between the two sets of horizontal dots that moved vertically and between the two sets of vertical dots that moved horizontally and therefore made it easier to group all four sets of dots into a single form that moved in one global direction. If this is true, reducing the amplitude of the local motion noise should reduce performance in noise-present stimuli. We tested this hypothesis in Experiment 4b by using a noise amplitude that was one half the value used in Experiment 4a. Experiment 4b also addressed a potential criticism of Experiment 4a, namely, that the high correlation between phase adjustment error and direction discrimination accuracy is due primarily to the presence of three outliers who were observers that had large adjustment errors and poor discrimination accuracy in the high-noise condition (see Figure 13a). By reducing the amplitude of local noise, we hoped to increase the range of phase errors found in the noise-present condition and therefore derive a more robust estimate of the relationship between adjustment error and discrimination accuracy.

## Experiment 4b

### *Results*

*Adjustment task.* Results from the adjustment task from four representative observers are displayed in Figure 14 in radians. Average phase error in $\pi$ radians for all observers is shown in Figure 10b. Average phase error in the medium-noise trials ($M = 0.54$) was slightly greater than the average error in the zero-noise condition ($M = 0.49$), a difference that was statistically significant ($t(14) = 2.44$, $p = 0.028$, $d = 0.63$). Phase error did not differ from chance (.50 phase error) for zero-noise trials ($t(14) = 0.67$, $p = 0.517$, $d = 0.17$) or medium-noise trials ($t(14) = -0.10$, $p = 0.918$, $d = 1.86$).

An ANOVA comparing performance on noise-present and noise-absent trials in Experiment 4a and Experiment 4b yielded a significant Experiment × Noise interaction ($F(1, 43) = 30.97$, $p > 0.0001$, $\eta_P^2 = 0.42$). The phase adjustment error was significantly greater in medium noise (Experiment 4b) than high noise (Experiment 4a) ($t(21.7) = 3.65$, $p = 0.001$, $d = 1.28$), but the average phase error in zero-noise trials did not differ significantly between experiments ($t(23.2) = 1.03$, $p = 0.31$, $d = 0.35$).

*Discrimination task.* The average proportions of correct responses from the discrimination task, ignoring confidence ratings, are shown in Figure 11b. The difference between proportion correct in the medium-noise ($M = 0.512$) and zero-noise ($M = 0.291$) conditions was significant ($t(14) = 5.99$, $p < 0.0001$, $d = 1.54$). Hence, as we found in the previous experiments, adding local motion noise improved direction discrimination accuracy. Furthermore, response accuracy in the zero-noise condition was significantly below chance performance ($t(14) = 3.03$, $p = 0.009$, $d = 0.78$). Accuracy on medium-noise trials did not differ significantly from chance ($t(14) = 0.15$, $p = 0.879$, $d = 1.70$)

A 2 (noise present vs. noise absent) × 2 (experiment) ANOVA comparing proportion correct in Experiments 4a and 4b found a significant Experiment × Noise interaction ($F(1, 43) = 26.37$, $p < 0.0001$, $\eta^2_P = 0.38$). Two-sample $t$ tests comparing accuracy in Experiments 4a and 4b found that accuracy differed on noise-present trials ($t(23.8) = 3.62$, $p = 0.001$, $d = 1.22$) but not noise-absent trials ($t(23.6) = 0.65$, $p = 0.52$, $d = 0.22$).

Figure 12b displays the average proportion of high-confidence responses (i.e., trials on which observers selected “definitely” on the response screen) that were incorrect. Two participants were excluded from this analysis because they did not make at least one high-confidence response for each noise level. The proportion of high-confidence errors in the zero-noise (0.78) and medium-noise (0.45) conditions differed significantly ($t(12) = 3.44$, $p = 0.005$, $d = 0.95$).

An ANOVA comparing the proportion of high-confidence errors on noise-present and noise-absent trials in Experiments 4a and 4b revealed a significant interaction between noise and experiment ($F(1, 38) = 4.33$, $p = 0.044$, $\eta^2_P = 0.10$). Two-sample $t$ tests comparing the proportion of high-confidence errors in Experiments 4a and 4b found a significant difference on noise-present trials ($t(19.2) = 2.46$, $p = 0.023$, $d = 0.91$) but not noise-absent trials ($t(26.8) = 0.16$, $p = 0.874$, $d = 0.05$).

*Relating adjustment responses & discrimination accuracy.* Figure 13b displays the relationship between performance on the adjustment and discrimination tasks for all participants in Experiment 4b. As predicted, reducing the amplitude of the local dynamic noise in this experiment resulted in broader distributions of phase adjustment error and discrimination accuracy in the high-noise condition than was found in Experiment 4a. Nevertheless, as was found in Experiment 4a, response accuracy in the discrimination task was significantly linearly associated with average phase error in the adjustment task in both the noise-present ($b = -1.06$, $t(13) = -9.14$, $p < 0.0001$, $R^2_{adj} = .855$) and noise-absent conditions ($b = -.91$, $t(13) = -7.06$, $p < 0.0001$, $R^2_{adj} = .777$).

The relation between phase adjustment error and discrimination accuracy in Experiments 4a and 4b was analyzed with a linear model that included phase error as a continuous predictor variable, experiment as a categorical predictor variable, and a Phase Error × Experiment interaction term. Noise-present and noise-absent conditions were analyzed separately. In the noise-present condition, the linear association between phase error and discrimination accuracy in Experiment 4a was significant ($b = -1.19$, $t(41) = -9.36$, $p < 0.001$), and neither the effect of experiment ($b = -0.06$, $t(41) = -0.91$, $p = 0.367$) nor the Phase Error × Experiment interaction ($b = 0.14$, $t(41) = -0.96$, $p = 0.341$) were significant. Similar results were obtained in the noise-absent condition: The association between phase error and discrimination accuracy was significant in Experiment 4a ($b = -0.76$, $t(41) = -5.87$, $p < 0.001$), but neither the effect of experiment ($b = -0.15$, $t(41) = -0.78$, $p = 0.439$) nor the Phase Error × Experiment interaction ($b = -0.15$, $t(41) = -0.78$, $p = 0.439$) were significant. Hence, these analyses suggest that the relation between phase adjustment error and direction discrimination accuracy was similar in the two experiments.

### Discussion

In Experiment 4a, phase adjustment error was very small and direction discrimination accuracy was very high in the noise-present condition. To avoid this ceiling effect, Experiment 4b used a local dynamic noise amplitude that was half the amplitude used in Experiment 4a. As expected, reducing the amplitude of the noise resulted in a wider distribution of phase adjustment error and discrimination accuracy across participants (cf. Figures 13a and 13b), and therefore provided a more reliable estimate of the relation between performance in the adjustment and discrimination tasks. Overall, the results of Experiment 4b were consistent with those obtained in Experiment 4a: In both experiments, response accuracy in the discrimination task was higher in the noise-present condition than in the zero-noise condition, discrimination accuracy was significantly below chance in the zero-noise condition but not in the medium-noise condition, and discrimination accuracy in both conditions was significantly correlated with phase error in the adjustment task. One difference between experiments is that discrimination accuracy in the noise-present condition was significantly above chance when the local dynamic noise amplitude was 0.027 dva (Experiment 4a) but not 0.0135 dva (Experiment 4b).

## Experiment 4c

In Experiment 1, we found that older adults were more likely than younger adults to perform below

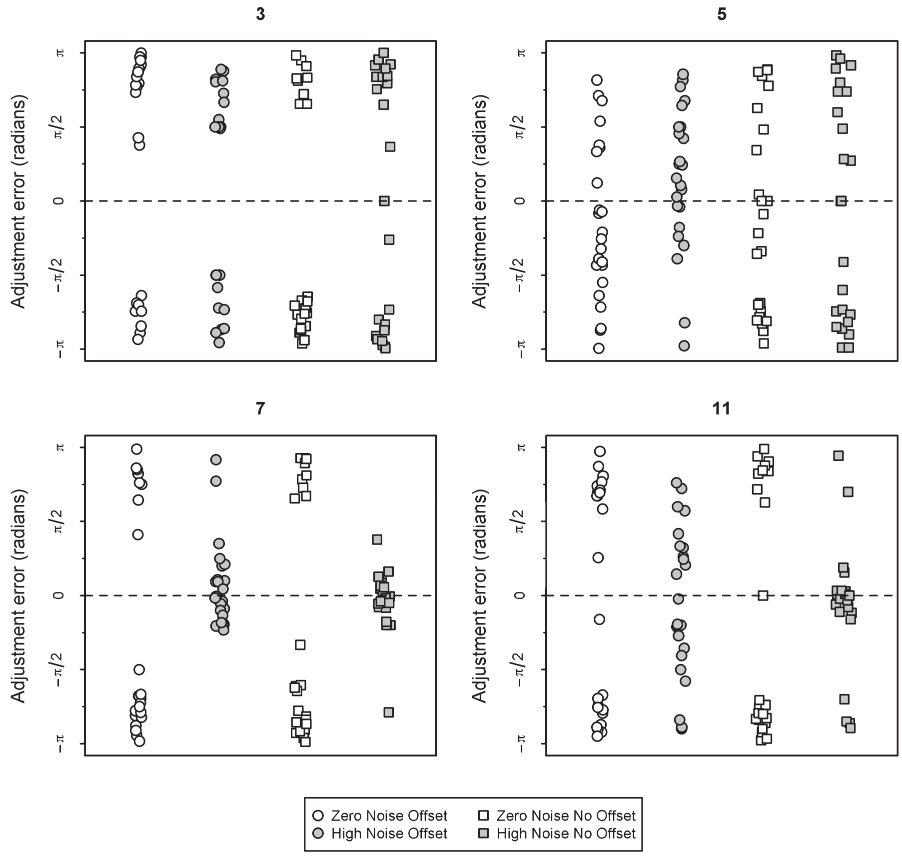



Figure 15. Phase adjustment data from four representative observers in Experiment 4c. Symbol conventions are the same as in Figure 9.

chance at direction discrimination in zero-noise stimuli. We therefore hypothesized that older adults would display larger degrees of phase bias than younger adults.

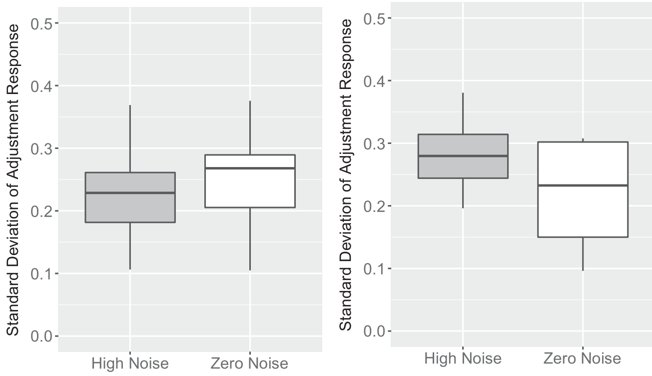
### *Results*

*Adjustment task.* Average adjustment error from Observers 1 to 7 was compared to Observers 8 to 15 to check for differences that may have occurred due to the change in experimental procedure described above. A two-sample $t$ test revealed no difference between the two groups for the adjustment task ($t = 0.59$, $p = 0.56$, $d = 0.23$), and so the data were combined across groups for the following analyses.

Results from the adjustment task from four representative observers are displayed in Figure 15. Average phase errors in the zero- and high-noise conditions are shown in Figure 10c, in $\pi$ radians. Mean error in the two conditions differed significantly ($t(14) = 4.48$, $p < 0.0001$, $d = 1.16$). Average phase error in zero-noise trials was significantly less than 0.5 (i.e., the phase error predicted by random responding) ($t(14) = 2.74$, $p = 0.016$, $d = 0.71$), whereas average phase error in the high-noise trials was significantly greater than 0.5 ($t(14) = 4.03$, $p = 0.001$, $d = 1.78$).

Phase adjustment error measured in older participants was compared to the adjustment error measured in younger participants in Experiment 4a with a 2 (noise present vs. noise absent) × 2 (experiment) ANOVA. The main effect of noise ($F(1, 43) = 83.17$, $p > 0.0001$, $\eta_P^2 = 0.66$) was significant, but the main effect of experiment ($F(1, 43) = 1.68$, $p = 0.202$, $\eta_p^2 = 0.04$) and the Noise × Experiment interaction ($F(1, 43) = 1.08$, $p = 0.305$, $\eta_p^2 = 0.02$) were not.

A comparison of Figures 9 and 15 suggests that the responses from older adults in Experiment 4c were more variable and less precise than responses from younger adults in Experiment 4a. The standard deviations of the absolute value of adjustment error, computed for each participant and condition, are plotted in Figure 16. To quantitatively assess this comparison, a 2 (noise present vs. absent) × 2 (experiment) ANOVA compared the mean standard deviation of adjustment errors across conditions and experiments. Unlike what was found by the analysis of mean adjustment error, the analysis of the standard deviation yielded a significant Noise × Experiment interaction ($F(1, 43) = 10.22$, $p = 0.003$, $\eta_P^2 = 0.19$). A two-sample $t$ test comparing the mean standard deviation in the high-noise condition in Experiments 4a and 4c was significant ($t(30.8) = 2.97$, $p = 0.006$, $d = 0.91$), but the difference between




(a) Experiment 4a (b) Experiment 4c

Figure 16. Standard deviation of response in adjustment task, calculated within each participant for Experiment 4a ($M_{age}$ = 20; $n = 30$) and Experiment 4c ($M_{age} = 73$; $n = 15$). The amplitude of local, sinusoidal jitter was 0.0135 in the high-noise condition and absent in the zero-noise condition.

experiments in the zero-noise condtion was not significant ($t(22.5) = 1.12$, $p = 0.274$, $d = 0.39$).

*Discrimination task.* The average proportions of correct responses in the discrimination task, ignoring confidence ratings, are shown in Figure 11c. The difference between response accuracy in the high-noise ($M = 0.617$) and zero-noise ($M = 0.219$) conditions was significant ($t(14) = 4.31$, $p < 0.001$, $d = 1.11$). Furthermore, response accuracy in the zero-noise condition was significantly below 0.50, which represents chance performance ($t(14) = 5.43$, $p < 0.0001$, $d = 1.40$).

An ANOVA comparing performance on high-noise and zero-noise trials in Experiments 4a and 4c produced a significant Noise × Experiment interaction ($F(1, 43) = 4.50$, $p = 0.040$, $\eta_P^2 = 0.09$), suggesting that the effect of noise on accuracy differed in younger and older adults. A two-sample $t$ test found that the difference between accuracy in Experiments 4a and 4c was significant on high-noise trials ($t(29.2) = 2.89$, $p = 0.007$, $d = 0.90$) but not zero-noise trials ($t(30.4) = 0.31$, $p = 0.762$, $d = 0.09$).

Figure 12c displays the average proportion of high-confidence responses that were incorrect. One participant did not make at least one high-confidence response in each noise level and therefore was not included in this analysis. The proportion of high-confidence errors in the zero-noise ($M = 0.82$) and high-noise ($M = 0.41$) conditions differed significantly ($t(13) = 4.14$, $p = 0.001$, $d = 1.11$).

An ANOVA comparing the proportion of high-confidence responses that were incorrect in Experiments 4a and 4c found significant main effects of noise ($F(1, 39) = 4.31$, $p = 0.044$, $\eta_P^2 = 0.10$) and experiment ($F(1, 39) = 60.11$, $p < 0.0001$, $\eta_P^2 = 0.61$), but the Noise × Experiment interaction was not significant ($F(1, 39) = 2.23$, $p = 0.144$, $\eta_P^2 = 0.05$). This analysis suggests that high-confidence errors were more frequent in high-noise conditions and more frequent in older than younger subjects, but the effect of noise was similar in both age groups.

*Relating adjustment responses & discrimination accuracy.* Figure 13c displays the data and regression line comparing the results of each observer in both tasks in Experiment 4c. Regressing direction discrimination accuracy onto average phase error in the adjustment task yielded a significant relationship on high-noise trials, ($b = .6112$, $t(13) = 2.249$, $p = 0.0425$, $R^2_{adj} = .2247$) but not zero-noise trials ($b = .3764$, $t(13) = 2.098$, $p = 0.056$, $R^2_{adj} = .1956$). The relationship between discrimination accuracy and adjustment error in Experiments 4a and 4c was evaluated with linear models that included adjustment error as a continuous predictor, experiment (4a or 4c) as a categorical predictor, and an Adjustment × Experiment interaction. The high-noise and zero-noise conditions were analyzed separately. In the high-noise condition, the linear model fit the data reasonably well ($R^2_{adj} = 0.727$). As noted earlier, the association between discrimination accuracy and adjustment error was significant in Experiment 4a ($b = -1.19$, $t(41) = -9.01$, $p < 0.001$). The difference between the regression slopes in Experiments 4a and 4c was not significant ($b = 0.42$, $t(41) = 1.72$, $p = 0.093$); however, the intercept of the regression line was lower in Experiment 4c than Experiment 4a (0.86 vs. 1.09, $b = -0.23$, $t(41) = -2.76$, $p = 0.009$). In the zero-noise condition, the overall fit of the linear model was poorer ($R^2_{adj} = 0.405$). As noted earlier, there was a significant relationship between discrimination accuracy and adjustment error in Experiment 4a ($b = -0.76$, $t(41) = -5.25$, $p < 0.001$), and the difference between the regression slopes ($b = 0.27$, $t(41) = -1.05$, $p = 0.297$) and intercepts ($b = 0.27$, $t(41) = -1.05$, $p = 0.297$) in the two experiments was not significant. These analyses suggest that the relation between discrimination accuracy and phase adjustment error was similar in younger (Experiment 4a) and older (Experiment 4c) observers, although (a) overall accuracy in the high-noise condition was lower in older observers, and (b) the association between discrimination accuracy and phase error was weaker in older compared to younger subjects (cf. Figures 13a,c).

### Discussion

Experiment 4c measured the relationship between the perceived relative phase of two motion trajectories and perceived global motion in older adults. Specifically, Experiment 4c tested the hypothesis that the illusion of opposite motion is stronger in older than younger adults.

As expected, we found that phase adjustment error was smaller in the high-noise condition than in the

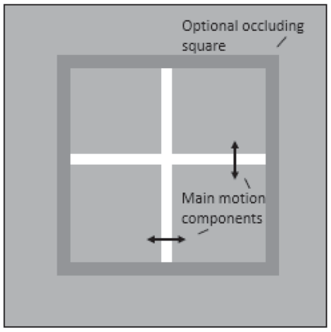



Figure 17. Depiction of the stimulus used in Experiment 5. Each of the white bars displayed sinusoidal motion, which varied in phase depending on condition. The gray occluding square was present on half the trials, representing the occluder present/absent manipulation.

zero-noise condition (see Figure 10c). Contrary to our hypothesis, there was no evidence that the magnitude of the effect of noise on phase adjustment error differed between older observers and the younger adults in Experiment 4a, although analyses conducted on the standard deviations of phase adjustment errors from each participant indicated that adjustments were more variable in older than younger adults. As was found with younger observers in Experiment 4a, discrimination accuracy for older adults was higher in high-noise trials than zero-noise trials; however, this effect of noise was smaller than the effect found in younger observers in Experiment 4a. Finally, as was found in Experiments 4a and 4b, discrimination accuracy was inversely related to phase adjustment error.

## Experiment 5

The prior experiments used stimuli that were similar to the one illustrated in Figure 1. In Experiment 4, we sought to extend the findings of the prior experiments by testing whether similar effects are obtained with a different stimulus described by Anstis (2007), which produces the so-called chopstick illusion. The stimulus consisted of a pair of crossed horizontal and vertical bars, each moving sinusoidally in a direction that is orthogonal to the bar's orientation (Figure 17). When the motions are out of phase by 90 deg, the summed motion of both the bars is circular in either a clockwise or counter clockwise direction around a central fixation point. In both this stimulus and those used in the above experiments, it is theoretically possible to monitor each of the motion components separately and sum the phases to get a correct answer to direction of motion, without perceiving a single coherent percept. The responses of observers who completed the task this way and who perceived the stimulus as multiple, separate components could be similar to the responses of observers who integrated the two moving lines into a single coherent percept. To gain a better understanding of how behavior in this task is related to perceived coherence, the current experiment therefore compared discrimination accuracy with ratings of perceived coherence.

### Method

#### Observers

Twenty naïve observers (11 female, mean age 20.25 years) participated in this experiment. One participant withdrew from the study, leaving 19 observers in the final sample.

#### Apparatus and stimuli

The apparatus used in this experiment was identical to the prior experiments. The stimulus used in this experiment is depicted in Figure 17. The stimulus consisted of a pair of crossed, horizontal and vertical white bars presented against a gray background. Each bar was 1.85 deg long and 0.10 deg wide. The luminance of each bar was 95.6 cd/m$^2$ and background luminance was 43.9 cd/m$^2$. Each bar moved sinusoidally along the direction perpendicular to its principal axis (i.e, the horizontal bar moved vertically and the vertical bar moved horizontally). In the discrimination task, the motion of the bars was offset by ±90 deg so that the added motion of the two bars created either clockwise or counter clockwise motion. In the adjustment task, the relative phase of the two bars was adjustable, as in Experiment 4. On half of the trials, a thin (0.05-deg visual angle) square border surrounded the stimulus.

#### Procedure

Prior to the experiment, participants were shown a demonstration consisting of dynamic plaid patterns that was used to explain the concept of motion coherence. Each plaid was displayed in a circular aperture 3.13 deg in diameter. Each grating component moved at a speed of 0.08 deg/s toward either the upper-left or upper-right corner of the display. The first, coherent plaid consisted of two square-wave gratings of 0.68 cpd oriented ±45 deg from vertical and a total root mean square (RMS) contrast of 0.52. This stimulus evoked a percept of a plaid pattern that drifted coherently upward. The second, incoherent plaid similarly consisted of two square-wave gratings oriented ±45 deg from vertical, but one grating had a fundamental frequency of 1.23 cpd while the other had a fundamental of 0.14 cpd, RMS contrast of 0.56. This plaid did not evoke a percept of coherent upward motion, but rather a percept of each grating sliding over one another (Stoner & Albright, 1996). Participants were shown each type of plaid while the experimenter described the concepts of coherent and incoherent motion, and were asked to describe the motion of each plaid. We defined coherent motion as motion that appeared to unite into a single-motion trajectory and incoherent motion as motion that appeared disjointed. The stimuli remained visible for as long as was necessary for the participants to understand the distinction between coherent and incoherent motion.

The experimental trials began immediately following the demonstration and followed the same procedures used in Experiment 4. Observers completed an adjustment task followed by a discrimination task. In the adjustment task, the target direction of orbital motion was indicated by the word “clockwise” or “counter clockwise” being shown in the center of the display for 2 s at the start of each trial. The moving stimulus then appeared in one of four starting phases: 90 deg, which produced clockwise orbital motion; −90 deg. Which produced counter-clockwise motion; and two neutral phases (0 and 180 deg) that produced diagonal motion. As in Experiment 4, the participant used a dial to adjust the relative phase of the stimulus to produce the target motion and pressed the space bar when satisfied. Unlike prior experiments, participants then used the mouse to click a button on the screen to rate the motion coherence on a 4-point (0–3) scale: 0 was labeled “low coherence” and 3 was labeled “high coherence.” The following trial began 1500 ms following the coherence rating. Participants completed 32 occluder-present trials and 32 occluder-absent trials. The adjustment task lasted approximately 40 min.

The discrimination task was similar to prior experiments, except that observers made motion coherence ratings on each trial. The stimulus duration was 600 ms, either in clockwise or counter clockwise global motion. Observers made their response following stimulus offset by pressing one of two labeled keys on a standard computer keyboard. Following the direction response, four buttons labeled, 0, 1, 2, and 3 appeared on the display. Observers rated the perceived coherence of the immediately preceding trial. Participants completed 72 occluder-present trials and 72 occluder-absent trials. The discrimination task lasted approximately 15 min.

## Results

Average phase error is reported in the adjustment task, and proportion correct is reported in the discrimination task. Coherence ratings in both tasks were transformed from a 4-point scale to scores ranging from 0 to 1. The distributions of all data are shown in Figure 19.

### Adjustment task

Results from the adjustment task from four representative observers are displayed in Figure 18. Average phase errors in the occluder-present and occluder-absent conditions are shown in Figure 19a. Mean phase error in the occluder-present ($M$ = 0.239) and occluder-absent ($M$ = 0.371) conditions differed significantly ($t(18) = 2.51$, $p = 0.0220$, $d = 0.575$).

### Discrimination task

The average proportions of correct responses in the discrimination task are shown in Figure 19c. The difference between response accuracy in the occluder-present ($M$ = 0.789) and occluder-absent ($M$ = 0.561) conditions was significant ($t(18) = 4.079$, $p <$ 0.001, $d = 0.936$).

### Relating adjustment responses & discrimination accuracy

Figure 20 displays the data and regression line comparing the results of each observer in both tasks in Experiment 5. Regressing direction discrimination accuracy onto average phase error in the adjustment task yielded a significant relationship on occluder-present trials ($b = -0.593$, $t(18) = -7.268$, $p = <0.0001$, $R^2_{adj} = 0.7422$) and not occluder-absent trials ($b = -0.178$, $t(18) = -1.227$, $p = 0.237$, $R^2_{adj} = 0.0273$).

### Below-chance performance

One goal of Experiment 5 was to test whether the illusion of opposite orbital motion obtained with the Lorenceau stimulus could be seen with a different stimulus. However, on average, performance in all four conditions in the adjustment and discrimination tasks is above chance levels (see Figure 19). The means of both occluder-present ($t(18) = 5.44$, $p = <0.0001$, $d = 1.25$, $M = 0.24$) and occluder-absent ($t(18) = 2.64$, $p = 0.017$, $d = 0.61$, $M = 0.37$) trials were significantly below 0.5 $\pi$ phase error in the adjustment task. Similarly, the means of both occluder-present ($t(18) = 4.11$, $p < 0.001$ $d =$ 0.94, $M = 0.79$) and occluder-absent ($t(18) = 0.79$, $p$ $= 0.446$, $d = 0.18$, $M = 0.56$) trials were significantly above 0.5 proportion correct in the discrimination task.

Although observers did not display below-chance performance in either task on average, we tested whether performance was significantly below chance in the discrimination task in individual observers, using Bonferroni-corrected ($\alpha_{FW} = .01$) binomial tests. We found that proportion correct in the discrimination task was significantly less than 0.5 in six observers (mean accuracy = 0.11) when occluders were absent,

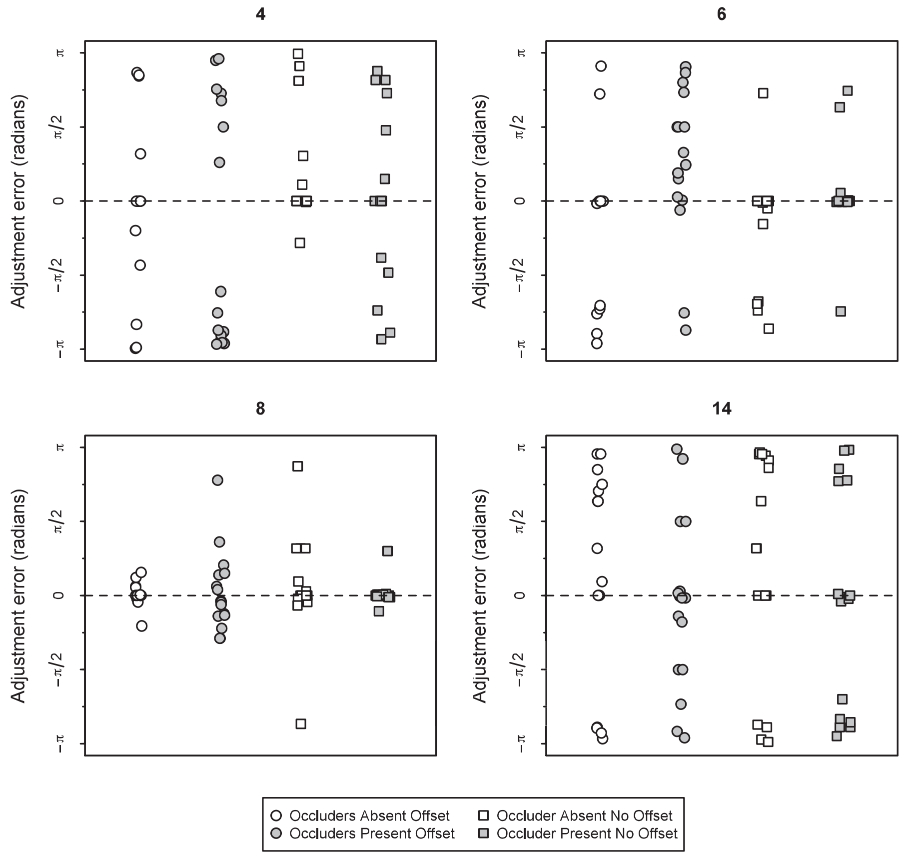



Figure 18. Phase adjustment data from four representative observers in Experiment 5. Each symbol corresponds to a phase setting from a single trial. Results from occluder-present and occluder-absent trials are represented by filled and unfilled symbols, respectively. The square and circle symbols indicate results from a trial in which the initial global motion was in the correct direction (i.e., no offset; relative phase was −90 or 90 deg) or was in an ambiguous direction (i.e., offset; relative phase was 0 or 180 deg).

and remained below chance in three of these observers (mean accuracy = 0.19) when occluders were present.

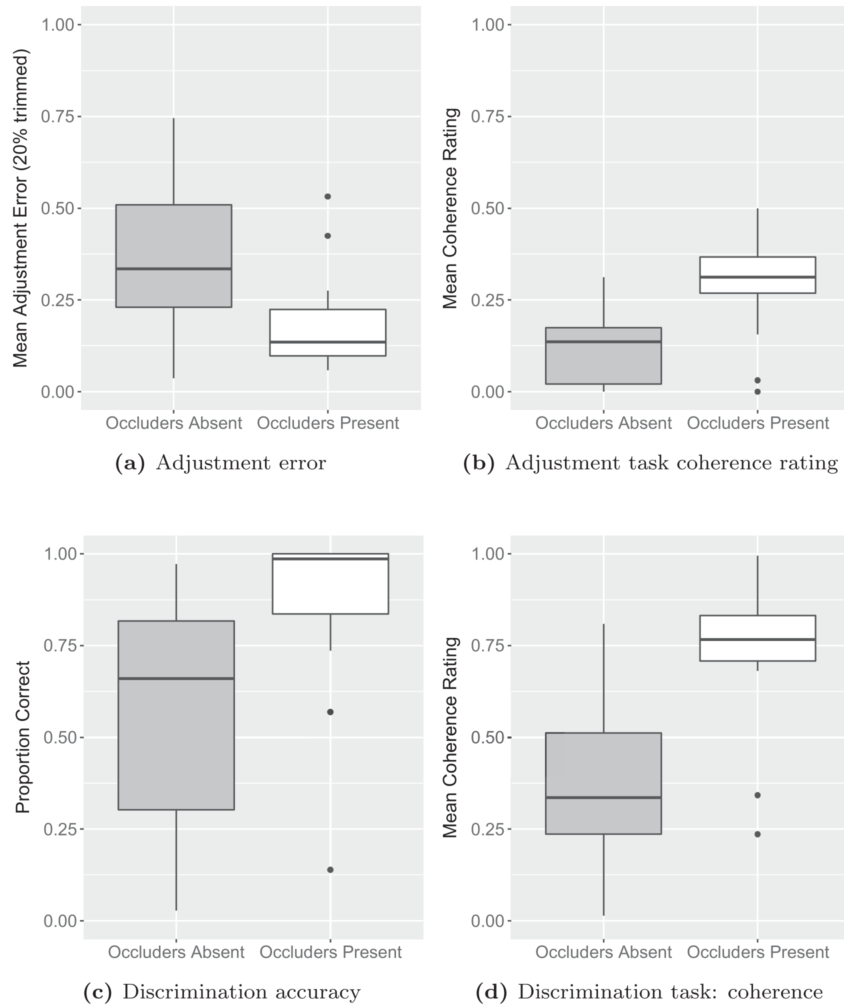
### Coherence ratings

Average motion coherence ratings were analyzed in a 2 (task) × 2 (occluder present vs. absent) ANOVA. Results of this analysis are displayed in Table 3.

There was a significant main effect of occluder, reflecting the fact that coherence ratings were higher in the occluder-present condition. There also was a significant main effect of task, indicating that coherence ratings were higher in the discrimination task than the adjustment task. Finally, there was a significant Occluder × Task interaction. To interpret this interaction, we conducted $t$ tests comparing coherence ratings in the occluder-present and occluder-absent conditions in each task. Although the effect of the occluder was larger in the discrimination task, coherence ratings were significantly higher in the occluder-present condition in the discrimination task ($t(18) = 4.718$, $p = 0.0002$, $d = 1.082$) and the adjustment task ($t(18) = 3.818$, $p = 0.0013$, $d = 0.876$).

### Relating coherence and phase error

The relationship between coherence and phase adjustment error was evaluated with an ordinal logistic mixed-model regression using the Ordinal package in R (Christensen, 2019a; R Core Team, 2017). Statistical significance was assessed with $p$ values estimated using adaptive Gauss-Hermite quadrature approximation with 10 quadrature points, as recommended in Christensen (2019b). The model included adjustment error as a continuous predictor, occluder (absent or present) as a categorical predictor, an Adjustment × Occluder interaction, and subjects as a random effect. The model was reasonably well defined (condition number of Hessian: 259.35). The interaction between adjustment error and occluder was significant, indicating the slopes of the regression lines in the occluder-present and occluder-absent conditions differed significantly ($b = -2.043$, $z = -4.785$, $p < 0.0001$). The intercepts of the regression lines also differed significantly between the occluder-present and occluder-absent conditions ($b = 1.306$, $z = 6.894$, $p < 0.0001$). Therefore, we fit separate models data from




(a) Adjustment error (b) Adjustment task coherence rating (c) Discrimination accuracy (d) Discrimination task: coherence

Figure 19. Boxplots illustrating results of Experiment 5. In the occluders-present condition, the stimulus was surrounded by the outline of a square; in the occluder-absent condition, that square was absent. (a) The distributions of average (i.e., 20% trimmed mean) phase error across observers in the adjustment task. Perfect performance corresponds to 0 phase error. Phase error of 1 corresponds to the wrong direction of global motion. (b) The distribution of mean coherence ratings in the adjustment task. (c) The distribution of proportion correct in the discrimination task. (d) The distribution of mean coherence rating in the discrimination task.

the occluder present and occluder-absent conditions. We found that there was a significant association between rated coherence and adjustment error in the occluder-present condition ($b = -1.727$, $z = -4.787$, $p < 0.0001$) but not in the occluder-absent condition ($b = -0.237$, $z = -0.772$, $p = 0.4403$). These analyses suggest that the relation between rated coherence and phase adjustment error is different depending on whether occluders are present or absent. For occluder-present trials, there is a relationship between coherence and phase adjustment error, but that relationship does not exist in occluder-absent trials.

## Discussion

Experiment 5 compared subjective and objective measures using performance in an adjustment and discrimination task and rated coherence to stimuli in each task. There were three goals in Experiment 5: (a) to explore perception of ambiguous motion in a different stimulus, (b) to examine whether observers experience illusory motion in this stimulus, and (c) to compare the objective measures used in the prior experiments in the current study with subjective experience of perceived coherence.

To address first and second goals, we conducted analyses on the present data mirroring those used in Experiment 4. We found that the main findings of Experiment 4 extend to the stimulus used in Experiment 5. On average, occluder-present stimuli produced more accurate performance than occluder-absent stimuli in the adjustment and discrimination tasks. There was a significant relationship between adjustment and discrimination performance in occluder-present stimuli. Unlike the prior experiments, the average

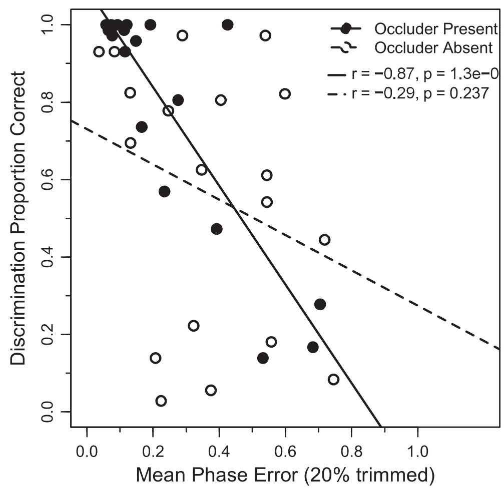



Figure 20. Relationship between response accuracy in the discrimination task and the average phase error in the adjustment task for Experiment 5. Pearson's *r* and *p* values for each regression are shown in the top-right corner of the plot.

| **Effect** | ***df*** | ***MSE*** | ***F*** | $p\eta^2_p$ | $p_{adj}$ |
|---|---|---|---|---|---|
| Occluder | 1, 18 | 0.04 | 25.24 | .58 | <0.0001 |
| Task | 1, 18 | 0.03 | 59.28 | .77 | <0.0001 |
| Occluder × Task | 1, 18 | 0.01 | 7.91 | .31 | 0.01 |

Table 3. Experiment 5 ANOVA of rated coherence.

discrimination and adjustment performance measures for occluder-absent stimuli were above chance. However, several observers exhibited performance significantly below chance in both tasks, suggesting that at least some observers experienced illusory motion.

The final goal of Experiment 5 concerned ratings of perceived coherence. Occluder-present stimuli were rated more coherent on average than occluder-absent stimuli. In the adjustment task, we were able to look at whether coherence rating on each trial could be predicted by phase error on that trial. In occluder-present stimuli, phase error was indeed related to coherence rating, where lower phase error occurred when coherence was rated higher. However, in occluder-absent stimuli, this relationship did not exist. This suggests that although, on average, rated coherence seems to be related to performance measures, on a trial-by-trial basis, this relationship only exists for some stimuli (here: occluder present) and not others (here: occluder absent).

## General discussion

In Experiment 1, we successfully replicated a surprising result reported by Lorenceau (1996), namely, that adding local motion noise significantly improves the performance in a global motion discrimination task, especially for long-stimulus durations. Interestingly, several observers, particularly older adults, showed below-chance performance when the stimulus was displayed for long durations and did not contain local motion noise. The local motion noise added to the stimulus perturbed the direction of each dot but maintained the global spatial distribution of the stimulus. Hence, local noise improved global motion perception presumably by increasing the likelihood of grouping the different motion trajectories into a single percept.

Experiment 2 examined the reliability of the below-chance performance by asking observers to rate their confidence in their responses. We found that observers were often confident about their incorrect decisions in conditions where the group average response accuracy was below chance. Furthermore, the probability of making high-confidence errors was greater in observers whose overall response accuracy was below chance. These results suggest that observers genuinely perceived global motion in the wrong direction. Experiment 2 also showed that the addition of a dynamic white noise to the stimulus background increased discrimination accuracy from below-chance to near-chance levels. This lends support to the idea that the local motion noise used by Lorenceau (1996) and in Experiment 1 is not the only type of noise that can increase response accuracy. We suggest that any stimulus manipulations that reduce the tendency of the individual dots to be grouped into four linear units would make it easier for the dots to be grouped into a single, moving global form.

Experiment 3 was conducted to test the specific hypothesis that the illusion of opposite motion in zero-noise stimuli is due to selecting a single motion component as a motion reference. Salient color cues were placed in the stimulus at locations that ought to bias observers to attend to either a single motion component or to both motion components. The results of Experiment 3 demonstrated that performance was largely unaffected by this manipulation. A motion reference account of illusory motion does not seem to provide an adequate explanation in the current task.

Eye movements have been implicated in various motion illusions, such as the peripheral drift illusion (Faubert & Herbert, 1999; Beer et al., 2008), the Enigma illusion (Troncoso et al., 2008), and the Filehne illusion (Mack & Herman, 1973). In the present experiments, we instructed participants to maintain central fixation throughout the experiment, although we did not record eye movements. Despite the lack of eye-tracking data, there are several reasons for suspecting that eye movements did not contribute significantly to this illusion. First, the illusion depends strongly on the presence or absence of noise, and it is not clear why eye movements would differ significantly in those conditions. Second, although naïve observers were used in all other experiments in the current study, Experiment 3 also found the illusion in experienced psychophysical observers who (at least in some conditions) were better able to maintain fixation (Cherici et al., 2012) and who were confident that they maintained fixation during each trial. Finally, informal observations by the authors have shown that switching between central fixation and deliberately tracking one set of dots, and covertly attending to one set of dots, does not affect the clear percept of illusory motion in the zero-noise stimulus.

Experiment 4a examined how accurately young adults perceived the direction of global motion using adjustment and discrimination tasks. With both tasks, we found that observers accurately perceived global motion when the stimulus contained local dynamic noise but consistently misperceived the direction of global motion when the stimulus did not contain noise. Furthermore, performance in both tasks was strongly correlated, which suggests that the failure to correctly perceive the direction of global motion was related to an inability to properly encode the relative phase of the two motion components. Experiment 4b replicated the main results of Experiment 4a using a lower amplitude local motion noise.

Experiment 4c focused on relating phase adjustment and discrimination performance and found below-chance performance with both tasks for zero-noise stimuli and above-chance performance for high-noise stimuli. However, older adults were more variable in their responses to high-noise stimuli in the adjustment task and performed worse on high-noise stimuli in the discrimination task compared to the younger adults of Experiment 4a. This is unsurprising, given the prediction that older observers would perform more poorly and produce more variable data due to age-related differences in spontaneous noise in the system, which has been reported to increase with age (Pardhan, 2004; Schmolesky et al., 2000), and calculation efficiency, which has been reported to decrease with age (Bennett et al., 1999; Pardhan et al., 1996). Although it is still debated which of these effects is responsible for decrements to perception with aging, this type of change to the visual system can explain the general increase in participant variability. Various measures used to track variability of response are reported in the Appendix.

The age difference that we observed in the high-noise condition in Experiment 4c may be related to age-related deficits in temporal integration (Arena et al., 2012; Snowden & Kavanagh, 2006) and direction discrimination (Bocheva et al., 2013). However, it is interesting to note that we observed significant age differences in mean accuracy in the discrimination task and in Experiment 1 but not in mean phase error in the adjustment task. The lack of an age difference in the adjustment task may be related to aspects of the task that are separate from motion integration per se. For example, the stimulus duration was considerably longer in the adjustment task than the discrimination task. Bennett et al. (2007) demonstrated that older adults can identify the direction of global motion of random dot kinematograms as accurately as younger adults if given longer stimulus durations. The stimulus duration is fixed in the discrimination task, but not the adjustment task, which may explain the discrepancy in the age difference in the discrimination task, but not the adjustment task. Older adults indeed took 24.49 S longer per trial on average (see Appendix for adjustment trial details).

Experiment 5 demonstrated that the effects reported in the current study are not stimulus specific, and instead probably depend predominately on the requirement to integrate sinusoidal motion across separate pieces of the stimulus, which is common to all stimuli used in the current experiments. In Experiment 5, we measured perceived coherence along with phase bias and discrimination ability. In general, larger phase errors were associated with reduced discrimination accuracy and lower perceived coherence. However, this relationship was found on a trial-by-trial basis in the adjustment task only in the occluder-present condition. From this analysis, we can conclude that if the correspondence between performance and coherence exists for many types of stimuli, it is at least stronger in some stimuli than others.

Studies of motion integration have used a variety of methods to assess successful, coherent integration. Studies that use stimuli that are structured such that there is a definitive correct or incorrect response can use objective measures of performance in which the participant presumably either answers correctly on each trial if they perceive the coherent, integrated stimulus (Smith et al., 1999; Shiffrar et al., 1995; Tang et al., 2015). Studies that use stimuli that are inherently ambiguous (Adelson & Movshon, 1982; Anstis & Kim, 2011; Welch, 1989) typically use subjective measures of an observer's perception of motion. Objective measures are sometimes preferred to subjective measures because it is thought that they may be more reliable and less biased. However, objective measures may be only indirectly related to the phenomenon of interest, namely, the perceived coherence of motion. In the current experiments, we were interested in whether an observer perceived one or two motion elements but instead asked the observer to report the integrated motion direction. These tasks measured the extent to which observers can combine motion components but do not measure the percept of the observer. To combat this weakness, in Experiment 5 of the current study, we ask observers to report their subjective experiences. In sum, we used several tasks, both objective and subjective, to characterize the perceptual experience of observers with this class of stimulus. The combination of the variety of methods we use in the current study adds richness and depth to the findings of the prior literature and bridges together prior findings that use any one method. We suggest that using several measures is a strategy to overcome the weaknesses of each method and allows for a richer understanding of the experimental results. This added information fortunately comes with a very small added cost to the data collection and analysis process, and so it is an effort that is likely worthwhile for many researchers to undertake.

The illusory motion seen in the present experiments is reminiscent of the effects of motion repulsion. Motion repulsion is a visual illusion where the trajectory of motion of a field of dots is altered by the presence of a moving reference frame or another superimposed moving field of dots (Marshak & Sekuler, 1979). Despite the similarities, the effects of the current study are likely not due to motion repulsion for a few reasons. First, motion repulsion is most commonly seen in dot fields and not seen for certain stimuli, such as plaids (Kim & Blake, 1996). Given that illusory motion is seen in the current article for both dot and line stimuli, it is unlikely that the illusory motion is simply motion repulsion. Second, if a motion repulsion account of the present illusion were true, one might expect one of the two motion components, either horizontal or vertical, to be used as the reference. Accordingly, the adjustments made in Experiment 4 might be different depending on whether the observer was adjusting the vertical or horizontal component. However, in each set of observers in Experiment 4, there was no significant difference in mean phase error between trials where adjustments were made to the vertical and horizontal components. Last, motion repulsion is not seen when motion directions are 90 deg from one another as they are in the current study (Hiris & Blake, 1996).

A large body of work focuses on the specific conditions that lead to grouping or segmentation. Typically, these studies vary the stimulus to determine which stimulus configurations lead to better grouping. The types of stimulus perturbations studied include varying local noise (Lorenceau, 1996), stimulus contrast and eccentricity (Takeuchi, 1998), the presence and absence of occluders and terminators (McDermott & Adelson, 2004; Vallortigara & Bressan, 1991; Lorenceau & Shiffrar, 1992), notches or features that disambiguate motion (Castet & Wuerger, 1997), and the number of sides of the shape (Tang et al., 2015). In the current article, we do not address the subject of ambiguous motion from this perspective. Instead, we are concerned with the perceptual consequences of not grouping the stimulus: that is, what does the observer perceive in this stimulus when there is no noise in the dots (Experiments 1–4) or when the occluder is absent (Experiment 5)?. In the current article, we have shown that misperception occurs in situations that do not lead to global grouping.

Previous research using the orbiting square stimulus in Figure 1 has focused on the stimulus manipulations that enable observers to integrate the motion components to perceive the correct global motion rather than how or why observers consistently perceive global motion in the wrong direction (Lorenceau & Shiffrar, 1992, 1999; Lorenceau, 1996; Shiffrar & Lorenceau, 1996). Lorenceau (1996) discussed several explanations for the effect, including the possible contribution of pursuit eye movements, but did not experimentally explore the phenomenon. Shiffrar & Lorenceau (1996) examined the perceived global orbital motion of stimuli consisting of four contours arranged to form a diamond and found that discrimination accuracy was below chance when the stimulus comprised wide, high-contrast contours, which suggests that the effects reported in this article are not unique to the specific manipulations we used here. Instead, they may be present whenever perceived global direction depends on relative phase of different motion components in stimuli that lack compelling cues for global grouping.

Our results suggest that the misperception of global motion is associated with a failure to accurately encode the relative phase of the horizontal and vertical motion components in our stimulus. Why might observers fail to encode phase accurately? One possibility is that the visual system fails to efficiently sample the motion stimulus. Recent studies suggest that rhythmic electrical responses in the brain are associated with moment-to-moment fluctuations in attention (e.g., Fries et al., 2008; Busch et al., 2009). Although often studied in detection tasks, the idea that the visual system samples from the environment in a rhythmic fashion has been implicated in motion integration tasks as well. In the current study, it is possible that observers are misrepresenting the phase of one or both elements when attempting to combine them, in a process similar to rhythmic visual aliasing. This misrepresentation may lead to a combination that more closely represents the opposite direction percept and therefore produces performance that is below chance. Interestingly, if this explanation were true, it would imply that aliasing can occur under some conditions (i.e., the stimulus is segmented) and not under other conditions (i.e., the stimulus is grouped). It is also worth noting that the frequency of the motion components in the current stimulus is 0.83 Hz, which is quite low compared to the alpha-band frequencies typically studied in the rhythmic perception research, which range from 8 to 11 Hz. The aliasing literature therefore provides an interesting interpretation of the current results but does not explain the phenomenon studied here.

The experiments reported presently explore how relative phase integration in motion can be understood in a variety of tasks in the presence and absence of global grouping cues. In general, we found good performance when cues that encourage global grouping were present and below-chance performance in the absence of cues that encourage global grouping. This below-chance performance is important for research in perceptual grouping in motion because it is suggestive of a bias in information integration that may be present in other types of integration or grouping tasks.

*Keywords: motion, illusion, grouping*

## Acknowledgments

Commercial relationships: none.
Corresponding author: Jessica N. Cali.
Email: calij@mcmaster.ca.
Address: Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada.

## Footnotes

[1]This result is due to the presence of two older adults with unusually high accuracy (Figure 3b): When those two outliers are removed, the age difference in the zero-noise condition is significant ($t(19.3) = 3.9$, $p = 0.0008$).
[2]Similar results were obtained when average error was calculated as the mean rather than the trimmed mean.

## References

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

Alais, D., Blake, R., & Lee, S. H. (1998). Visual features that vary together over time group together over space. *Nature Neuroscience, 1*, 160–164.

Anstis, S. (2007). The ash-lag effect during illusory chopstick rotation. *Perception, 36*, 1043–1048.

Anstis, S., & Kim, J. (2011). Local versus global perception of ambiguous motion displays. *Journal of Vision, 11*, 1–13, https://doi.org/10.1167/11.3.13.

Arena, A., Hutchinson, C. V., & Shimozaki, S. S. (2012). The effects of age on the spatial and temporal integration of global motion. *Vision Research, 58*, 27–32.

Ball, K., & Sekuler, R. (1986). Improving visual perception in older observers. *The Gerontological Society of America, 41*, 176–182.

Beer, A. L., Heckel, A. H., & Greenlee, M. W. (2008). A motion illusion reveals mechanisms of perceptual stabilization. *PLoS ONE, 3*, 1–7.

Bennett, P. J., Sekuler, A. B., & Ozin, L. (1999). Effects of aging on calculation efficiency and equivalent noise. *Journal of the Optical Society of America A*, 654–668.

Bennett, P. J., Sekuler, R., & Sekuler, A. B. (2007). The effects of aging on motion detection and direction identification. *Vision Research, 47*, 799–809.

Bocheva, N., Angelova, D., & Stefanova, M. (2013). Age-related changes in fine motion direction discriminations. *Experimental Brain Research, 228*, 257–278.

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

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Busch, N. A., Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *Journal of Neuroscience, 29*, 7869–7876.

Castet, E., & Wuerger, S. (1997). Perception of moving lines: Interactions between local perpendicular signals and 2D motion signals. *Vision Research, 37*, 705–720.

Cherici, C., Kuang, X., Poletti, M., & Rucci, M. (2012). Precision of sustained fixation in trained and untrained observers. *Journal of Vision, 12(6)*, 31, https://doi.org/10.1167/12.6.31.

Christensen, R. H. B. (2019a). *Ordinal-regression models for ordinal data*. R Foundation for Statistical Computing. R package version 2019.3-9.

Christensen, R. H. B. (2019b). *A tutorial on fitting cumulative link mixed models with clmm2 from the ordinal package*. Retrieved from https://cran.r-project.org/web/packages/ordinal/vignettes/clmm2_tutorial.pdf.

Faubert, J., & Herbert, A. M. (1999). The peripheral drift illusion: A motion illusion in the visual periphery. *Perception, 28*, 617–621.

Fries, P., Womelsdorf, T., Oostenveld, R., & Desimone, R. (2008). The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *Journal of Neuroscience, 28*, 4823–4835.

Hiris, E., & Blake, R. (1996). Direction repulsion in motion transparency. *Visual Neuroscience, 13*, 187–197.

Hutchinson, C. V., Arena, A., Allen, H. A., & Ledgeway, T. (2012). Psychophysical correlates of global motion processing in the aging visual system: A critical review. *Neuroscience and Biobehavioral Reviews, 36*, 1266–1272.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 1489–1506.

Johansson, G. (1950). Configurations in the perception of velocity. *Acta Psychologica, 7*, 25–79.

Kim, J., & Blake, R. (1996). Direction repulsion between components in motion transparency. *Vision Research, 13*, 187–197.

Lorenceau, J. (1996). Motion integration with dot patterns: Effects of motion noise and structural information. *Vision Research, 36*, 3415–3427.

Lorenceau, J., & Alais, D. (2001). Form constraints in motion binding. *Nature Neuroscience, 4*, 745–751.

Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research, 32*, 263–273.

Lorenceau, J., & Shiffar, M. (1999). The linkage of visual motion signals. *Visual Cognition, 6*, 431–460.

Lorenceau, J., & Zago, L. (1999). Cooperative and competitive spatial interactions in motion integration. *Visual Neuroscience, 16*, 755–770.

Mack, A., & Herman, E. (1973). Position constancy during pursuit eye movement: An investigation of the Filehne illusion. *Quarterly Journal of Experimental Psychology, 25*, 71–84.

Marshak, W., & Sekuler, R. (1979). Mutual repulsion between moving visual targets. *Science, 205*, 1399–1401.

McDermott, J., & Adelson, E. H. (2004). The geometry of the occluding contour and its effect on motion interpretation. *Journal of Visions, 4(10)*, 944–954.

Murray, R. F., Sekuler, A. B., & Bennett, P. J. (2001). Time course of amodal completion revealed by a shape discrimination task. *Psychonomic Bulletin & Review, 8*, 713–720.

Pardhan, S. (2004). Contrast sensitivity loss with aging: sampling efficiency and equivalent noise at different spatial frequencies. *Journal of the Optical Society of America A*, 169–175.

Pardhan, S., Gilchrist, J., Elliott, D. B., & Beh, G. K. (1996). A comparison of sampling efficiency and internal noise level in young and old subjects. *Vision Research, 36*, 1641–1648.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*, 437–442.

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rauber, H. J., & Treue, S. (1998). Reference repulsion when judging the direction of visual motion. *Perception, 27*, 393–402.

Roudaia, E., Bennett, P. J., Sekuler, A. B., & Pilz, K. S. (2010). Spatiotemporal properties of apparent motion perception and aging. *Journal of Vision, 10*, 1–15, https://doi.org/10.1167/10.14.5.

Schmolesky, M. T., Wang, Y., Pu, M., & Leventhal, A. G. (2000). Degradation of stimulus selectivity of visual cortical cells in senescent rhesus monkeys. *Nature Neuroscience, 3*, 384–390.

Shiffrar, M., Li, X., & Lorenceau, J. (1995). Motion integration across differing image features. *Vision Research, 35*, 2137–2146.

Shiffrar, M., & Lorenceau, J. (1996). Increased motion linking across edges with decreased luminance contrast, edge width and duration. *Vision Research, 36*, 2061–2067.

Shiffrar, M., & Pavel, M. (1991). Percepts of rigid motion within and across apertures. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 749–761.

Smith, A. T., Curran, W., & Braddick, O. J. (1999). What motion distributions yield global transparency and spatial segmentation? *Vision Research, 39*, 1121–1132.

Snowden, R. J., & Kavanagh, E. (2006). Motion perception in the ageing visual system: Minimum motion, motion coherence, and speed discrimination thresholds. *Perception, 35*, 9–24.

Snowden, R. J., & Verstraten, F. A. (1999). Motion transparency: Making models of motion perception transparent. *Trends in Cognitive Sciences, 3*, 369–377.

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

Takeuchi, T. (1998). Effect of contrast on the perception of moving multiple Gabor patterns. *Vision Research, 38*, 3069–3082.

Tang, M. F., Dickinson, J. E., Visser, T. A. W., Edwards, M., & Badcock, D. R. (2015). Role of form information in motion pooling and segmentation. *Journal of Vision, 15*, 1–18, https://doi.org/10.1167/15.15.19.

Troncoso, X. G., Macknik, S. L., Otero-Millan, J., & Martinez-Conde, S. (2008). Microsaccades drive illusory motion in the Enigma illusion. *Proceedings of the National Academy of Sciences, 105*, 16033–16038.

Vallortigara, G., & Bressan, P. (1991). Occlusion and the perception of coherent motion. *Vision Research, 31*, 1967–1978.

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

Wickens, T. D. (2002). *Elementary signal detection theory*. Oxford, UK: Oxford University Press.

Williams, D. W., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research, 24*, 55–62.

Zivotofsky, A. Z. (2004). The Duncker illusion: Intersubject variability, brief exposure, and the role of eye movements in its generation. *Investigative Ophthalmology and Visual Science, 45*, 2867–2872.

## Appendix

Tables A1, A2, and A3 report multiple characteristics of individual observer performance meant to provide information about how each observer completed the adjustment tasks in Experiment 4. These values are used to make sure that observers are adjusting the dial at least some amount on every trial and exploring all possible arrangements of the stimulus. Four measures are reported in each table. The first of these measures is the total value of adjustments made, in radians. This represents the total amount of radians adjusted by the observers on the average trial. During the experiment, the dial was checked for movement every 700 ms. Column 2 of Table A1 reports the number of 700-ms increments in which an adjustment is made. This information is useful to combine with the total radians

| Observer number | Total radians adjusted | | Number of intervals with adjustment | | Trial duration (seconds) | |
|---|---|---|---|---|---|---|
| | Mean | *SD* | Mean | *SD* | Mean | *SD* |
| 1 | 41.67 | 4.26 | 29.62 | 3.12 | 9.65 | 12.00 |
| 2 | 35.25 | 23.75 | 24.50 | 11.76 | 17.81 | 11.12 |
| 3 | 26.61 | 13.85 | 27.62 | 10.08 | 25.78 | 16.96 |
| 4 | 10.43 | 10.67 | 19.68 | 10.72 | 27.40 | 16.97 |
| 5 | 10.99 | 4.82 | 20.12 | 12.92 | 24.35 | 27.32 |
| 6 | 25.73 | 15.57 | 23.49 | 17.61 | 21.37 | 16.58 |
| 7 | 40.52 | 11.12 | 24.13 | 6.28 | 14.19 | 16.43 |
| 8 | 41.45 | 20.24 | 15.49 | 11.10 | 13.51 | 9.67 |
| 9 | 10.33 | 2.90 | 30.02 | 5.66 | 11.85 | 12.26 |
| 10 | 23.67 | 14.08 | 30.76 | 15.25 | 23.61 | 18.63 |
| 11 | 24.34 | 9.08 | 46.17 | 16.23 | 29.96 | 22.14 |
| 12 | 46.99 | 14.51 | 29.63 | 3.40 | 15.48 | 16.01 |
| 13 | 148.39 | 47.63 | 24.53 | 6.73 | 18.29 | 19.87 |
| 14 | 104.23 | 28.20 | 47.29 | 27.82 | 32.34 | 21.64 |
| 15 | 8.01 | 4.53 | 5.06 | 2.23 | 9.23 | 6.35 |
| 16 | 49.86 | 18.81 | 39.05 | 17.35 | 20.70 | 19.79 |
| 17 | 9.42 | 2.74 | 18.66 | 8.84 | 15.95 | 11.06 |
| 18 | 37.10 | 17.08 | 30.03 | 20.28 | 12.77 | 12.73 |
| 19 | 31.95 | 13.38 | 20.8 | 7.48 | 17.07 | 12.72 |
| 20 | 41.50 | 16.98 | 37.50 | 11.67 | 48.38 | 27.54 |
| 21 | 6.87 | 5.83 | 33.12 | 12.51 | 14.84 | 16.61 |
| 22 | 5.78 | 1.93 | 32.10 | 11.78 | 16.11 | 12.91 |
| 23 | 51.77 | 22.42 | 26.15 | 15.24 | 19.32 | 12.43 |
| 24 | 15.35 | 7.18 | 21.93 | 11.29 | 24.88 | 10.06 |
| 25 | 30.77 | 4.40 | 66.59 | 16.23 | 25.21 | 26.09 |
| 26 | 71.75 | 36.85 | 55.15 | 23.65 | 50.42 | 41.26 |
| 27 | 127.09 | 19.82 | 74.31 | 25.92 | 43.03 | 38.98 |
| 28 | 131.47 | 6.37 | 65.48 | 14.63 | 18.15 | 27.99 |
| 29 | 54.82 | 22.76 | 90.48 | 21.94 | 23.56 | 24.91 |
| 30 | 52.00 | 22.89 | 60.55 | 12.41 | 21.85 | 19.56 |

Table A1. Experiment 4a.

| Observer number | Total radians adjusted | | Number of intervals with adjustment | | Trial duration (seconds) | |
|---|---|---|---|---|---|---|
| | Mean | *SD* | Mean | *SD* | Mean | *SD* |
| 1 | 23.09 | 6.77 | 105.98 | 31.51 | 35.38 | 37.30 |
| 2 | 2.33 | 1.26 | 16.39 | 8.21 | 19.62 | 15.71 |
| 3 | 11.75 | 4.71 | 33.34 | 10.16 | 19.84 | 19.46 |
| 4 | 5.38 | 1.39 | 80.63 | 14.55 | 21.59 | 16.68 |
| 5 | 15.38 | 7.92 | 47.01 | 10.75 | 24.29 | 25.60 |
| 6 | 25.54 | 16.62 | 58.26 | 21.10 | 29.83 | 23.02 |
| 7 | 19.37 | 9.35 | 24.09 | 10.03 | 21.71 | 11.24 |
| 8 | 36.81 | 5.46 | 20.04 | 8.06 | 12.45 | 10.33 |
| 9 | 16.07 | 9.22 | 37.25 | 33.23 | 23.61 | 14.61 |
| 10 | 53.64 | 15.79 | 54.27 | 17.69 | 27.92 | 25.39 |
| 11 | 26.70 | 12.66 | 72.35 | 16.18 | 19.01 | 22.37 |
| 12 | 33.04 | 12.07 | 38.14 | 8.80 | 18.09 | 15.64 |
| 13 | 18.59 | 5.95 | 30.20 | 9.19 | 26.66 | 23.37 |
| 14 | 5.35 | .99 | 54.38 | 20.26 | 22.57 | 14.30 |
| 15 | 268.94 | 80.73 | 109.35 | 30.42 | 34.73 | 33.94 |

Table A2. Experiment 4b.

| Observer number | Total radians adjusted | | Number of intervals with adjustment | | Trial duration (seconds) | |
|---|---|---|---|---|---|---|
| | Mean | *SD* | Mean | *SD* | Mean | *SD* |
| 1 | 5.47 | 1.16 | 38.30 | 5.54 | 8.50 | 4.59 |
| 2 | 54.61 | 28.79 | 50.41 | 25.60 | 45.00 | 34.20 |
| 3 | 16.01 | 6.46 | 42.48 | 20.84 | 52.86 | 28.50 |
| 4 | 40.87 | 13.54 | 197.00 | 49.69 | 88.75 | 66.24 |
| 5 | 12.41 | 1.92 | 96.82 | 40.98 | 59.97 | 36.40 |
| 6 | 29.21 | 16.84 | 85.58 | 32.20 | 67.81 | 38.01 |
| 7 | 12.73 | 7.89 | 59.03 | 30.05 | 118.25 | 71.31 |
| 8 | 181.23 | 38.18 | 164.39 | 40.63 | 48.84 | 32.43 |
| 9 | 25.71 | 19.80 | 12.93 | 7.92 | 45.76 | 26.48 |
| 10 | 37.48 | 12.67 | 14.50 | 4.90 | 26.59 | 12.47 |
| 11 | 15.48 | 10.19 | 48.87 | 24.78 | 33.57 | 22.13 |
| 12 | 12.56 | 4.68 | 19.62 | 15.69 | 33.58 | 9.98 |
| 13 | 2.77 | 2.60 | 15.93 | 6.19 | 20.65 | 12.67 |
| 14 | 5.83 | 0.59 | 24.17 | 4.65 | 22.02 | 15.33 |
| 15 | 8.86 | 7.57 | 27.39 | 11.74 | 27.19 | 14.98 |

Table A3. Experiment 4c.

adjusted on average and the trial length to understand how consistently observers adjusted the stimulus throughout the trial. Finally, we report trial length in seconds. This also represented the stimulus duration, as the stimulus was on screen until the observer was satisfied with their adjustment decision.

These control measures can help identify observers who did not adjust the dial during the trial or made a very poor effort at the task by adjusting the dial by the same amount each trial. These types of behavior indicate failure to properly attempt to complete the task. We did not exclude any participants from our analyses but include these values to add information to our overall findings. Mainly, the differences between observers in these values indicate different methods of completing the task, despite similar instructions. It is important to note that many tasks likely have variations in performance such as this but do not have access to this information due to the structure of the task.