Endogenous attention biases transformational apparent motion based on high-level shape representations

Sharif Saleki

Department of Psychological and Brain Sciences, Dartmouth College, NH, USA ✉

Kirsten Ziman

Department of Psychological and Brain Sciences, Dartmouth College, NH, USA ✉

Kevin C. Hartstein

Department of Psychological and Brain Sciences, Dartmouth College, NH, USA ✉

Patrick Cavanagh

Centre for Vision Research, York University, Toronto, Ontario, Canada
Department of Psychology, Glendon College, Toronto, Ontario, Canada ✉

Peter U. Tse

Department of Psychological and Brain Sciences, Dartmouth College, NH, USA ✉

When two pre-existing, separated squares are connected by the sudden onset of a bar between them, viewers do not perceive the bar to appear all at once. Instead, they see an illusory morphing of the original squares over time. The direction of this transformational apparent motion (TAM) can be influenced by endogenous attention deployed before the appearance of the connecting bar. Here, we investigated whether the influence of endogenous attention on TAM results from operations over high-level feature-independent shape representations, or instead over lower level shape representations defined by specific visual features. To do so, we tested the influence of endogenous attention on TAM in first- and second-order displays, which shared common shapes but had different shape-defining attributes (luminance and texture contrast, respectively). In terms of both the magnitude of directional bias and timing, we found that endogenous attention exerted a similar influence on both first- and second-order objects. These results imply that endogenous attention biases the perceived direction of TAM by operating on high-level shape representations that are invariant to the low-level visual features that define them. Our results support a four-stage model of TAM, where a feature encoding stage passes a features-specific layout to a parsing stage that forms discrete, high-level meta-featural shapes, which are then matched and visually interpolated over time.

Introduction

When a figure abruptly appears adjacent to a pre-existing figure, an illusory percept of motion away from the pre-existing figure occurs (Hikosaka, Miyauchi, & Shimojo, 1993a, 1993b, 1993c). The original figure appears to change shape continuously into the new shape defined by the combination of the old and new figures (Tse, Cavanagh, & Nakayama, 1998). For instance, when a bar abruptly appears adjacent to a static square, it appears to smoothly shoot out of the square, and when the bar disappears all at once, it appears to contract back into the square as if it were an animation.

This illusory line motion or transformational apparent motion (TAM) can also be perceived when two pre-existing, statically presented objects ("squares") are suddenly connected by another stimulus appearing between them (a "bar"), as in Figure 1. In this case, the direction of TAM can be influenced by initial differences between the squares. For example, when the squares in the initial display are physically identical (with no attentional, temporal, figural, or featural differences; as illustrated in frame 1 of Figure 1a), the onset of the bar between them leads to an illusory percept of the squares extending toward each other
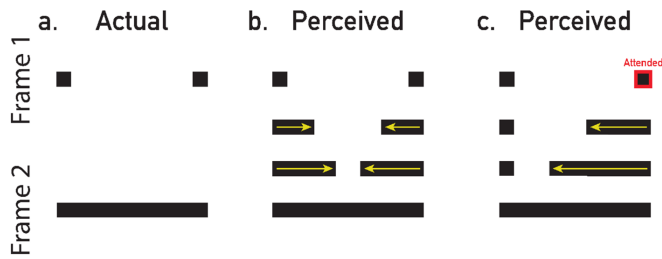
Figure 1. Transformational apparent motion. When a bar appears between adjacent, identical squares (a), participants perceive both squares as continuously changing shape—appearing to extend toward each other and collide in the middle (b). If attention is directed to one of the squares, motion is perceived away from the attended square (c). Yellow arrows indicate the direction the perceived motion.

and colliding in the middle (Figure 1b). However, when either exogenous (Hikosaka, 1993a) or endogenous (Hikosaka et al., 1993b, 1993c) attention is directed to one or the other of the two squares, the direction of the perceived motion along the bar appears to shoot away from the attended square (Figure 1c) (Hikosaka et al., 1993a, 1993b, 1993c; Faubert & von Grünau, 1995).

Hikosaka et al. (1993a) used an earlier offset of one of the two squares to draw exogenous attention and found that the illusory motion was seen to move more strongly away from that square. Faubert and von Grünau (1995) made one square turn on before the other, attracting exogenous attention to it. They also found that motion was seen predominantly away from the exogenously attended square.

The allocation of endogenous attention to one square also causes subsequent illusory motion to be perceived away from the attended square (similar to Figure 1c), toward the unattended square (Hikosaka et al., 1993b, 1993c; Schmidt, 2000). Therefore, both stimulus-driven and volitional attention can bias TAM to appear to extend away from the attended square.

This influence of attention on the illusory motion initially led researchers to theorize that the direction of illusory motion is mediated by attention-induced prior entry (Hikosaka et al., 1993a, 1993b, 1993c). Titchener (1908) defined prior entry as the phenomenon whereby information from an attended locus in the visual field is processed faster than information away from that locus, thus enjoying earlier access to subsequent stages of processing. This prior entry account suggests that a gradient of attention speeds processing for the parts of the bar stimulus that are nearer the attended square. Moreover, the farther away a portion of the bar is from the attended square, the later it enters awareness, creating an impression of motion along the bar. However, more recent results (Faubert & von Grünau, 1995; Tse et al., 1998; Tse & Logothetis, 2002; Tse 2006) have challenged this view.

In particular, differences in shape correspondence between squares and the bar strongly affect the perceived direction of motion. For example, if a red bar appears all at once between a red and green square, the bar will appear to shoot out of the red square alone (Faubert & von Grünau, 1995). Likewise, if the bar shares continuous contours with one square and not the other, it will appear to undergo a dynamic shape change from the former square alone (Tse et al., 1998; Tse & Logothetis, 2002; Tse, 2006). These instances of illusory motion driven by shape cannot be explained by the prior entry account or attentional gradients. We have, therefore, suggested that TAM is instead driven by the parsing and matching of the successive shapes presented in the image sequence (Hartstein, Saleki, Ziman, Cavanagh, & Tse, 2021; Tse et al., 1998). These operations decide which of the figures had to change shape to account for their new arrangement. The change in that figure is then interpolated by the motion processing system as a continuous transformation even though the change was discrete in the physical stimulus.

Several studies have provided compelling evidence that TAM and its variants are high-level phenomena that involve an inference of a change in shape, regardless of how that shape was defined in the stimulus. For example, Hartstein et al. (2021) showed that the direction of motion was determined by shape correspondence between stimuli in a similar manner for both first- and second-order displays, indicating that TAM is invariant to low-level visual cues.

In the current study, we ask whether the common processing of first- and second-order shapes in TAM holds when endogenous attention biases the direction of the illusory motion. Although some studies have found effects of endogenous attention on second-order stimuli (Jigo & Carrasco, 2018; Barbot, Landy, & Carrasco, 2012) that are similar to first-order stimuli (e.g., Reynold & Chelazzi, 2004; Barbot & Carrasco, 2017), two studies that directly compared these different types of stimuli reported a difference (Allen & Ledgeway, 2003; Lu, Liu & Dosher, 2000). In particular, they found that endogenous attention reduced the direction threshold for second-order motion but not for first order motion.

We present TAM as a four-stage process (Figure 2): First, features are extracted from the image. Second, a shape parsing process determines what counts as a shape, based on cues, such as contour continuity, that can segment a shape away from the background or other occluding or occluded shapes. When two images follow each other in succession, a third matching stage attempts to map shapes from the first image onto shapes in the next frame in order to map objects to themselves over time. At the fourth stage, the motion system infers the analog transformation that must have taken place in the world to account for the change from the initial
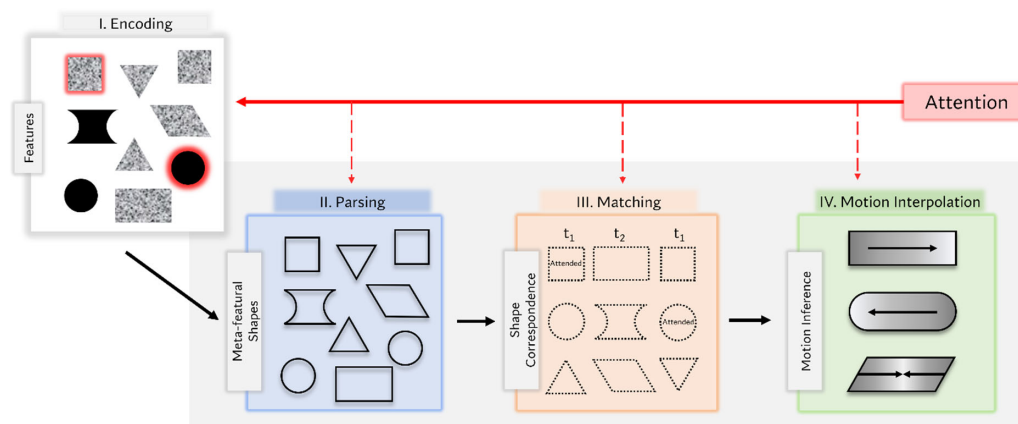
Figure 2. Four-stage model of TAM processing. (I) Encoding: low-level features and feature-specific shapes are detected at an early stage. Attention (red outlines) modulates first-order and second-order features differently (e.g., Allen & Ledgeway, 2003). The resulting activity is fed forward to the sequence of parsing, matching, and motion interpolation operations that result in TAM. (II) Parsing: meta-featural shapes are extracted from the input. Low-level feature information is discarded from the stimuli and feature specific attentional effects from the encoding stage (if any) may or may not carry over. The output of the parser is an abstract feature-independent shape representation. (III) Matching: shape contours that correspond with each other are matched across the two time intervals, t_1 and t_2. At this stage, the direction of motion could be biased by attentional differences (top and middle rows), or it could remain ambiguous in the absence of any attentional cues (bottom row). (IV) Motion interpolation: the motion that must have occurred in the world to give rise to the image sequence is reconstructed based on the correspondence between shapes.

shape at time 1 into its new form at time 2. The motion that we consciously experience as TAM is that analog interpolation.

Although this four-stage model has ruled out the attentional gradient as the source of TAM (Tse et al., 1998; Tse & Logothetis, 2002; Tse 2006), attention can nevertheless influence each stage of the process, biasing the direction that is seen when the direction would otherwise be ambiguous (Hikosaka et al., 1993b, 1993c; Downing & Treisman, 1997; Schmidt, 2000). Once the shapes have been encoded in a meta-featural format, these attentional biases will be identical for first- and second-order stimuli. However, previous studies have indicated that there is a difference in the effect of endogenous attention on first- and second-order stimuli (Lu et al., 2000; Allen & Ledgeway, 2003). This can only happen at the feature encoding stage as later stages no longer represent at first- and second-order stimuli independently (see Figure 2). As such, it remains an open question whether the effects of endogenous attention on the first stage of encoding can be inherited by higher levels, biasing the perceived direction of illusory motion in TAM differently for first- and second-order stimuli.

Here, we compared first- and second-order TAM in a case where the direction of motion was influenced solely by endogenous attention. We predicted that the effects of endogenous attention on initial shape extraction (featural level) would be lost once the shapes were encoded into a meta-featural representation, in the sense that it defines a figure via multiple possible

shape-defining features, known to be used for the parsing and matching stages of TAM (Hartstein et al., 2021). Therefore, the high-level representations of the second-order stimuli would show the same perceived direction as that seen for first-order stimuli across all stimulus onset asynchronies (SOAs). In line with our prediction, we found comparable effects of endogenous attention on both luminance- and texture-defined TAM. The time course of this effect roughly followed the time course of endogenous attention (stronger effect at longer cue-stimulus SOAs) (Nakayama & Mackeben, 1989). We conclude that the shape representations that go into the computation of TAM are meta-featural, and that attention then biases the matching and motion interpolation processes that go into the construction of the TAM percept.

Methods

Participants ($N = 15$) were Dartmouth College students enrolled in an introductory psychology course. All participants provided informed consent as required by the Committee for the Protection of Human Subjects at Dartmouth College and were compensated with course credit. Stimuli were presented using Psychtoolbox (Brainard & Vision, 1997; Pelli, 1997), in MATLAB (The MathWorks, Natick, MA) on an LCD monitor (15-in, 40.0° × 30.0°, 60 Hz).
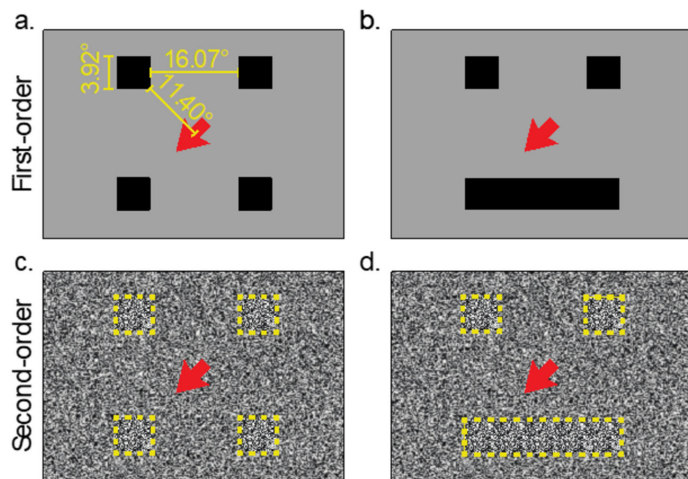
Figure 3. Square, arrow, and bar stimuli. At the start of each trial, participants saw four squares with a side length of 3.92°, positioned 11.40° away from a central arrow cue and 16.07° from each other, adjacently (a, first order; c, second order). Then, they saw an additional stimulus appear: a long bar bridging adjacent squares, 16.07° length (b, first order; d, second order). Note that the drawings in yellow do not reflect what participants saw, but have been added for illustration purposes.

Participants observed the display from a chin rest at a viewing distance of 57 cm and central fixation was monitored using an Eyelink II eyetracker (SR Research, Ontario, Canada). Trials during which gaze deviated more than 3° from central fixation were excluded from the analysis.

First-order stimuli were black on a medium gray background (Figures 3a, b). Second-order stimuli were defined by dynamically updating black and white textures on a statically presented black and white background (Figures 4d–f). In second-order trials, black and white background and stimulus textures were created using the same procedure as described in Hartstein et al. (2021).

Participants completed four 80-trial blocks for each stimulus type. Presentation order was counterbalanced between participants. Each trial began with a central fixation dot. After a jittered delay interval, four squares appeared (3.92° side length)—one in each quadrant of the display— equidistant (11.40°) from fixation (Figures 3 and 4). A central red arrow also appeared simultaneously, subsuming the central fixation dot, pointing at one of the squares. Participants covertly attended the square to which the arrow was pointing. Following a variable SOA of 33 to 366 ms, a bar (16.07° length) appeared that bridged two adjacent squares. In 80% of trials (valid trials) the bar bridged two squares, one of which had been indicated by the preceding arrow cue. The participant was instructed to indicate the direction of motion perceived along the connecting bar using the four arrow keys on a standard keyboard. Importantly, the bar could connect squares horizontally, on the top or bottom of the display (across hemifields), or vertically, to the left or to the right of fixation (within a single hemifield). The bar, squares, and arrow remained on the screen until the participant entered a response.

Catch trials (20%) were included to assess whether participants perceived motion in a systematic direction when the cue arrow was misleading or when incremental motion appeared in the display. In invalid catch trials (10% of total trials), the cue arrow pointed to a square that was not subsequently bridged by
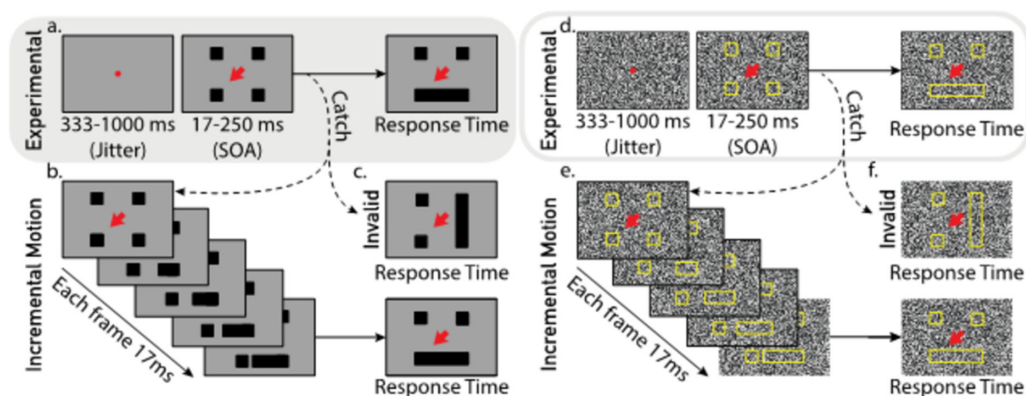

Figure 4. Experimental paradigm. Both first-order (left, a–c) and second-order (right, d–f) trials followed the same procedure. In experimental trials, we presented a central fixation dot (red), then a display of four squares with a central cue arrow, then a connecting bar between two adjacent squares (a, d). In incremental-motion catch trials, we presented a connecting bar incrementally, extending from an adjacent square toward the cued square (b, e). In invalid catch trials, we presented a connecting bar between two uncued squares (c, f). At the end of each trial, participants pressed one of four arrow keys indicating the direction of perceived motion. Note that yellow lines in (d–f) do not reflect what participants saw but have been added to illustrate the edges of second-order, texture-defined objects. SOA, stimulus onset asynchrony.

the bar stimulus. In incremental motion catch trials (10% of total trials), the bar stimulus was added incrementally, extending from an adjacent square to the square the arrow was pointing at (in the opposite direction from that predicted for TAM). Incremental motion occurred over 5 frames at 60 Hz presentation, for 83.3 ms total duration, mimicking the perception of TAM. As in experimental trials, the bar remained onscreen until the participant pressed a response key.

Results

One participant reported TAM outside the testing room, in the absence of any stimulus, and their data were excluded without review. We analyzed data from the remaining 14 participants. We compared how often participants perceived motion in the predicted direction (away from the cued square: congruent perception) in first- versus second-order experimental trials, across SOAs, using a two-way repeated measures analysis of variance. There was a significant main effect of SOA on congruent motion perception, $F(7,91) = 18.39$, $p < 0.001$, $\eta^2 = 0.59$, but the main effect of stimulus type and the interaction between SOA and stimulus type were not significant, $F(1,13) = 0.50$, $p = 0.49$, $\eta^2 = 0.04$ and $F(7,91) = 1.10$, $p = 0.37$, $\eta^2 = 0.08$, respectively. We further conducted Bayesian analyses and found that our data are best modeled solely by the main effect of SOA on congruent motion percepts; type III comparison of Bayes factors showed that the data pattern was more likely to be observed when an effect of SOA was added to the model ($BF_{10} = 10^{\wedge}13$). Furthermore, there was substantial evidence for the lack of the main effect of stimulus type ($BF_{10} = 0.18$), and strong evidence for the lack of an interaction ($BF_{10} = 0.03$).

This finding indicates that endogenous attention influenced the perceived direction of motion differently at different SOAs, but in a way that was similar across first- and second-order stimuli. To explore further, we compared the proportion of congruent motion percepts at each SOA to the proportion we would expect by chance (50%) for first- and second-order stimuli, using FDR correction for multiple comparisons. We found that participants perceived congruent motion significantly more than would be expected by chance at SOAs of 250 ms or longer (250, 300, and 366 ms; significance of $p < 0.01$ for each) for both first- and second-order displays. For four of the SOAs shorter than 250 ms (216, 183, 150, and 100 ms), congruent motion was not reported at a level significantly different from chance. The only SOA under 250 ms that showed a significant difference was the shortest SOA of 33 ms. Notably, this effect was in the opposite direction: in trials with a 33-ms SOA, participants reported
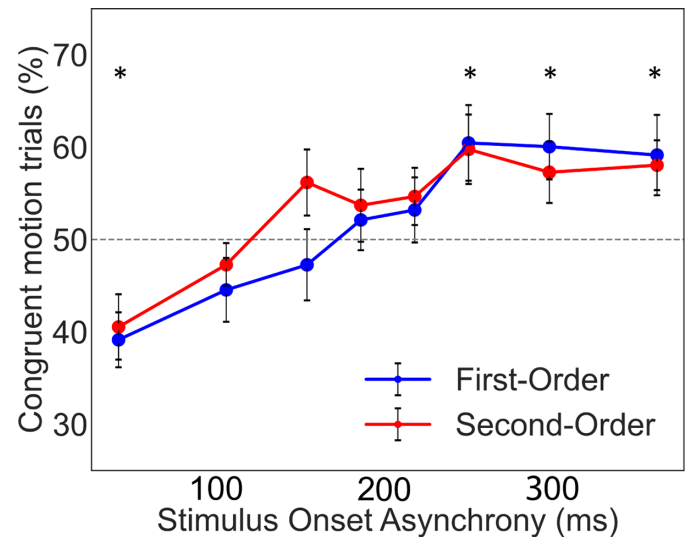

Figure 5. Proportion of experimental trials with congruent motion percept, by stimulus onset asynchrony (SOA). At longer SOAs ($\geq$250 ms), participants perceived motion away from the cued square (congruent) in a significantly higher proportion of trials than would be expected by chance (50%). This was true for both first-order stimuli (blue) and second-order (red) stimuli. Each SOA at which participants perceived motion in a particular cue-relative direction in more than half of trials (significant at $p < 0.05$) is denoted with an asterisk. Standard error of the mean for each stimulus type, at each SOA, is shown in black.

congruent motion significantly less than we would expect by chance, for both first and second-order displays, first-order: $t(13) = -3.93$, $p < 0.01$, $d = -1.48$; second-order: $t(13) = -2.55$, $p < 0.05$, $d = -0.96$.

For the SOAs that consistently yielded congruent motion percepts (250, 300, and 366 ms) (see asterisks in Figure 5), we proceeded to compare motion perception in experimental trials versus catch trials. First, we compared TAM percepts in experimental trials with real motion percepts in catch trials (Figure 6). A two-way, repeated measures analysis of variance revealed a significant difference between the proportion of congruent motion percepts in experimental (TAM) trials versus real motion catch trials, $F(1,13) = 52.96$, $p < 0.001$, $\eta^2 = 0.80$, with a higher proportion of congruent motion percepts in real-motion catch trials, $t(13) = 6.96$, $p < 0.001$, $d = 1.58$. However, there was no significant difference between first- and second-order stimuli, $F(1,13) = 2.60$, $p = 0.13$, $\eta^2 = 0.17$, and no significant interaction between stimulus type (first order vs. second order) and trial type, $F(1,13) = 1.53$, $p = 0.24$, $\eta^2 = 0.11$. So, at SOAs where participants consistently perceived TAM in the expected (congruent) direction, they did so comparably for both first- and second-order stimuli, with TAM percepts weaker than real motion percepts (fewer congruent trials) in both cases.
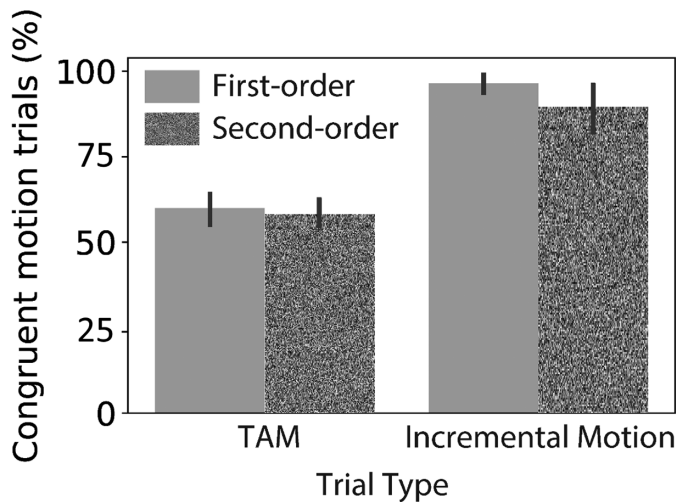
Figure 6. Proportion of congruent percepts in experimental versus real-motion trials, at longest stimulus onset asynchrony (SOA). At the longest SOAs ($\geq$250 ms), where participants perceived congruent motion in the majority of trials, they perceived congruent motion comparably for first- and second-order displays. In both cases, congruent motion was perceived in a significantly higher proportion of real-motion catch trials versus experimental transformational apparent motion (TAM) trials.

Next, we explored motion percepts in invalidly cued catch trials (Figures 4c, 4f) at the SOAs of interest. This measure was included to ensure that participants were not favorably perceiving or reporting motion in a particular direction in the absence of attentional influences (i.e., when they were not volitionally attending either of the squares connected by the bar stimulus, as in Hartstein et al 2021). We found that participants did not significantly favor a particular direction in invalidly cued catch trials where the direction of motion was vertical, for both first- and second-order displays. The proportion of trials where vertical motion was perceived in the upward (vs. downward) direction was not significantly above the level of chance, 50%; first order: $t(13) = 0.67$, $p = 0.51$, d = 0.26, and second-order: $t(13) = 0.22$, $p = 0.83$, d = 0.09. However, for invalidly cued catch trials where the direction of motion was horizontal, participants reported motion in the rightward (vs. leftward) direction in the significant majority of trials, for both first- and second-order displays, first order: $t(13) = 2.95$, $p < 0.05$, d = 1.16; and second order: $t(13) = 3.97$, $p < 0.01$, d = 1.56. This result raised the question of whether a rightward motion bias may have influenced participants' responses in experimental trials.

To investigate whether any rightward motion bias meaningfully influenced our key findings in experimental trials, we analyzed the proportion of congruent motion percepts in first- and second-order experimental trials, with respect to the orientation of the bar stimulus (horizontal or vertical) and SOA. If participants were heavily influenced by a rightward motion bias, regardless of attentional cuing, we would expect a lower proportion of cue-relative congruent motion percepts in horizontal versus vertical bar trials. However, we did not find a significant effect of bar orientation, first order: $F(1,13) = 2.66$, $p = 0.13$, $\eta^2 = 0.17$; and second order: $F(1,13) = .24$, $p = 0.63$, $\eta^2 = 0.02$, on the proportion of congruent motion percepts, nor a significant interaction between bar orientation and SOA, first-order: $F(7,91) = 1.92$, $p = 0.08$, $\eta^2 = 0.13$; and second order: $F(7,91) = .39$, $p = 0.91$, $\eta^2 = 0.03$. We did, again, find a significant main effect of SOA on the proportion of congruent motion percepts for first- and second-order displays, first order: $F(7,91) = 12.78$, $p < 0.001$, $\eta^2 = 0.50$; and second order: $F(7,91) = 6.87$, $p < 0.001$, $\eta^2 = 0.35$. This result indicates that our key metric in experimental trials, the proportion of congruent motion percepts, varied significantly with respect to SOA and was not significantly influenced by a rightward motion bias.

Discussion

Our results showed a comparable effect of endogenous attention on first- and second-order TAM. The time course of this effect roughly followed the time course of endogenous attention (stronger at longer cue-stimulus SOAs) (Nakayama & Mackeben, 1989). First- and second-order TAM were more likely to be perceived moving away from the endogenously attended figure at longer SOAs, consistent with previous studies that used exogenous cueing (e.g., Hikosaka et al., 1993a).

Surprisingly, at the shortest tested SOA (33 ms), the trend was reversed; participants perceived motion in the unexpected direction. This difference may have resulted from the time needed to focus endogenous attention onto one square, versus the entire set of squares, once the endogenous cue appeared at fixation. If correct, during the shift of endogenous attention to the cued figure, attention at noncued figures may have remained relatively dominant, accounting for the motion direction toward the cued figure at the shortest duration. Although we did not predict this result, it is consistent with previous findings of Christie and Klein (2005), where central cues, as opposed to peripheral cues, were used to direct top–down attention. Given that both first- and second-order TAM showed this same reversal at the shortest SOA, this finding supports our hypothesis that both are similarly affected by endogenous attention.

We also found that the frequency of congruent first- and second-order TAM reports were comparable,

although both forms of TAM were weaker than the perception of real motion in catch trials. These smaller effects are generally the case with reports of the influence of endogenous attention on TAM (e.g., Schmidt, 2000).

Furthermore, we observed that TAM in invalidly cued catch trials (where the bar connected two unattended stimuli) was perceived with a rightward motion bias: when the motion trajectory was horizontal, participants perceived rightward (vs. leftward) motion in a significant majority of trials. Tse and Cavanagh (2000) have previously shown that there may be a rightward bias in perceiving horizontal TAM, perhaps arising from how Western subjects read. For vertical motion trajectories, there was no such bias: participants perceived upward and downward motion in a comparable proportion of trials. This finding raised questions about whether rightward motion bias may have influenced participants' responses in experimental trials. However, an additional analysis of experimental trials showed no significant difference in the proportion of congruent motion percepts between horizontal and vertical motion trials, suggesting that rightward motion bias, if any, did not meaningfully impact our primary measure (proportion of congruent motion percepts).

Because endogenous attention seems to bias perceived motion direction in TAM figures in the same way, regardless of the features used to define those figures, we conclude that endogenous attention has little or no effect at the featural stage that could be passed along to subsequent processing. After the feature encoding stage, figural representations become meta-featural, or cue invariant and any effects of endogenous attention will influence first- and second-order stimuli equally. Our results therefore suggest that endogenous attention biases perceived motion, determining its direction at this higher level for TAM stimuli.

Extensive literature has shown that first-order and second-order stimuli can be processed by different mechanisms that exhibit distinct spatial tuning properties and temporal dynamics (Ledgeway & Smith, 1994, 1997; Allen & Derrington, 2001; Ledgeway & Hutchinson, 2005; Bressler & Whitney, 2006; Hutchinson & Ledgeway, 2007; Pavan & Mather, 2008; Pavan et al., 2009). Some second-order stimuli seem to drive low-level motion processes (Johnston, McOwan, & Buxton, 1992; Lu & Sperling, 1999) whereas others drive high-level (feature-tracking) motion systems, like those that respond to apparent motion which also show meta-featural responses (Cavanagh, Arguin, & von Grünau, 1989). Because we find similar temporal dynamics for the processing of first- and second-order stimuli in the case of TAM, we conclude that TAM relies on high-level motion systems that operates on high-level, meta-featural, shape representation

(Hartstein et al, 2021). Moreover, endogenous attention biases matching operations at the level of this high-level shape representation, rather than at the level of the low-level features associated with first- or second-order motion stimuli, because if low-level features played a role, we would expect to see different biases for the two types of stimuli, given that endogenous attention can bias first- and second-order motion direction sensitivity differently (Allen & Ledgeway, 2003; Lu et al., 2000).

We suggest that the meta-featural shape representation that serves as input to matching and motion processing stages is the output of a distinct stage of shape parsing. How a form is parsed influences the magnitude of motion signals generated by its perceived shape (Caplovitz & Tse, 2007b; Hsieh & Tse, 2007), suggesting that a form analysis stage precedes the computation of motion vectors. That the form processing stage and motion processing stage may involve different underlying neural populations is revealed by the finding that motion can adapt to a point that motion perception ceases, in a phenomenon known as "motion fading," although the form remains statically visible despite it having changed its location in the visual input (Hsieh & Tse, 2009; Kohler, Caplovitz, Hsieh, Sun, & Tse, 2010). Moreover, form outputs may interact with translational and rotational motion computations independently (Porter, Caplovitz, Kohler, Ackerman, & Peter, 2011). The stage of form analysis may take place in V3A (Caplovitz & Tse, 2007a) and/or in the lateral occipital complex before form outputs are passed on to motion processing areas, including hMT+ (Tse, 2006). Other areas, in particular V3v, V3B, and V4V, may also play a role in parsing operations (Caplovitz & Tse, 2010), whereas the integration of form cues over time may involve areas KO and hMT+ (McCarthy, Kohler, Tse, & Caplovitz, 2015).

In the matching and interpolation steps after the parsing stage, attention biases the direction of motion. Endogenous attention can prioritize the processing of one of the squares, such that the bar is attributed to only one of them in the matching stage. Therefore, it is regarded as a shape change of the attended square rather than the unattended square. The perceived motion is then an inference of the shape change that must have happened in the world to account for the transition from the prebar to postbar states. Given that the temporal resolution of the visual system is limited, there are many cases when discrete shape changes in the image sequence are seen as smooth motion, for example, the rapid sequences of still images in movies and television (Watson, Ahumada, & Farrell, 1986). Similarly for TAM, the assumption that objects change shape in an analog fashion in the world, discrete changes in the image are rendered as analog where possible.

Together, our results demonstrate that endogenous attention affects first- and second-order stimuli

similarly for shape-matching operations. The fact that endogenous attention seems to operate in a similar fashion for shape changes defined over first- and second-order stimuli suggests that attentional processes that influence TAM act on high-level representations of shapes and objects that are invariant to low-level information. That is, shape parsing and matching operations appear to operate over shape per se, regardless of how it is defined in the stimulus (luminance, texture, color, outlines etc.), and endogenous attention appears to act on these higher level representations to bias the perceived direction.

Keywords: transformational apparent motion, psychophysics, motion perception, second-order stimulus

Acknowledgments

Commercial relationships: none.
Corresponding author: Sharif Saleki.
Email: sharif.saleki.gr@dartmouth.edu.
Address: Department of Psychological and Brain Sciences, Dartmouth College, Hanover NH 03755, USA.

References

Allen, H. A., & Derrington, A. M. (2001). Distracting attention from contrast-defined motion. *Investigative Ophthalmology and Visual Science, 42*(4), 5061.

Allen, H. A., & Ledgeway, T. (2003). Attentional modulation of threshold sensitivity to first-or der motion and second-order motion patterns. *Vision Research, 43*(27), 2927–2936.

Barbot, A., & Carrasco, M. (2017). Attention modifies spatial resolution according to task demands. *Psychological Science, 28*(3), 285–296.

Barbot, A., Landy, M. S., & Carrasco, M. (2012). Differential effects of exogenous and endogenous attention on second-order texture contrast sensitivity. *Journal of Vision, 12*(8), 6–6.

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Bressler, D. W., & Whitney, D. (2006). Second-order motion shifts perceived position. *Vision Research, 46*(6–7), 1120–1128.

Caplovitz, G. P., & Tse, P. U. (2007a). V3A processes contour curvature as a trackable feature for the perception of rotational motion. *Cerebral Cortex, 17*(5), 1179–1189.

Caplovitz, G. P., & Tse, P. U. (2007b). Rotating dotted ellipses: Motion perception driven by grouped figural rather than local dot motion signals. *Vision Research, 47*(15), 1979–1991.

Caplovitz, G. P., & Tse, P. U. (2010). Extrastriate cortical activity reflects segmentation of motion into independent sources. *Neuropsychologia, 48*(9), 2699–2708.

Cavanagh, P., Arguin, M., & von Grünau, M. (1989). Interattribute apparent motion. *Vision Research, 29*(9), 1197–1204.

Christie, J., & Klein, R. M. (2005). Does attention cause illusory line motion? *Perception & Psychophysics, 67*(6), 1032–1043.

Downing, P. E., & Treisman, A. M. (1997). The line–motion illusion: Attention or impletion? *Journal of Experimental Psychology: Human Perception and Performance, 23*(3), 768.

Faubert, J., & Von Grünau, M. (1995). The influence of two spatially distinct primers and attribute priming on motion induction. *Vision Research, 35*(22), 3119–3130.

Hartstein, K. C., Saleki, S., Ziman, K., Cavanagh, P., & Tse, P. U. (2021). First-and second-order transformational apparent motion rely on common shape representations. *Vision Research, 188*, 246–250.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993a). Focal visual attention produces illusory temporal order and motion sensation. *Vision Research, 33*(9), 1219–1240.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993b). Voluntary and stimulus-induced attention detected as motion sensation. *Perception, 22*(5), 517–526.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993c). Visual attention revealed by an illusion of motion. *Neuroscience Research, 18*(1), 11–18.

Hsieh, P. J., & Tse, P. U. (2007). Grouping inhibits motion fading by giving rise to virtual trackable features. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 57.

Hsieh, P. J., & Tse, P. U. (2009). Motion fading and the motion aftereffect share a common process of neural adaptation. *Attention, Perception, & Psychophysics, 71*(4), 724–733.

Hutchinson, C. V., & Ledgeway, T. (2007). Asymmetric spatial frequency tuning of motion mechanisms in human vision revealed by masking. *Investigative Ophthalmology & Visual Science, 48*(8), 3897–3904.

Jigo, M., & Carrasco, M. (2018). Attention alters spatial resolution by modulating second-order processing. *Journal of Vision, 18*(7), 2–2.

Johnston, A., McOwan, P. W., & Buxton, H. (1992). A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 250*(1329), 297–306.

Kohler, P. J., Caplovitz, G. P., Hsieh, P. J., Sun, J., & Tse, P. U. (2010). Motion fading is driven by perceived, not actual angular velocity. *Vision Research, 50*(11), 1086–1094.

Ledgeway, T., & Smith, A. T. (1994). Evidence for separate motion-detecting mechanisms for first-and second-order motion in human vision. *Vision Research, 34*(20), 2727–2740.

Ledgeway, T., & Smith, A. T. (1997). Changes in perceived speed following adaptation to first-order and second-order motion. *Vision Research, 37*(2), 215–224.

Ledgeway, T., & Hutchinson, C. V. (2005). The influence of spatial and temporal noise on the detection of first-order and second-order orientation and motion direction. *Vision Research, 45*(16), 2081–2094.

Lu, Z. L., Liu, C. Q., & Dosher, B. A. (2000). Attention mechanisms for multi-location first-and second-order motion perception. *Vision Research, 40*(2), 173–186.

Lu, Z. L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics, 61*(6), 1075–1088.

McCarthy, J. D., Kohler, P. J., Tse, P. U., & Caplovitz, G. P. (2015). Extrastriate visual areas integrate form features over space and time to construct representations of stationary and rigidly rotating objects. *Journal of Cognitive Neuroscience, 27*(11), 2158–2173.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 29*(11), 1631–1647.

Pavan, A., & Mather, G. (2008). Distinct position assignment mechanisms revealed by cross-order motion. *Vision Research, 48*(21), 2260–2268.

Pavan, A., Campana, G., Guerreschi, M., Manassi, M., & Casco, C. (2009). Separate motion-detecting mechanisms for first-and second-order patterns revealed by rapid forms of visual motion priming and motion aftereffect. *Journal of Vision, 9*(11), 27–27.

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

Porter, K. B., Caplovitz, G. P., Kohler, P. J., Ackerman, C. M., & Peter, U. T. (2011). Rotational and translational motion interact independently with form. *Vision Research, 51*(23–24), 2478–2487.

Reynolds, J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annual Review of Neuroscience, 27*, 611–647.

Titchener, E. B. (1908). *Lectures on the elementary psychology of feeling and attention*. New York: Macmillan Publishers.

Schmidt, W. C. (2000). Endogenous attention and illusory line motion reexamined. *Journal of Experimental Psychology: Human Perception and Performance, 26* (3), 980.

Tse, P. U. (2006). Neural correlates of transformational apparent motion. *Neuroimage, 31*(2), 766–773.

Tse, P. U., & Cavanagh, P. (2000). Chinese and Americans see opposite apparent motions in a Chinese character. *Cognition, 74*(3), B27–B32.

Tse, P., Cavanagh, P., & Nakayama, K. (1998). The role of parsing in high-level motion processing. *High-Level Motion Processing: Computational, Neurobiological, and Psychophysical Perspectives*, 249–266.

Tse, P. U., & Logothetis, N. K. (2002). The duration of 3-D form analysis in transformational apparent motion. *Perception & Psychophysics, 64*(2), 244–265.

Watson, A. B., Ahumada, A. J., & Farrell, J. E. (1986). Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays. *Journal of the Optical Society of America A, 3*(3), 300–307.