Perception, 2013, volume 42, pages 1281–1300

doi:10.1068/p7635

Visually tracking and localizing expanding and contracting objects

Piers D L Howe[1], Alex O Holcombe[2], Mark D Lapierre[1], Simon J Cropper[1]
[1] Melbourne School of Psychological Sciences, University of Melbourne, Melbourne, VIC 3010, Australia; [2] Department of Psychology, University of Sydney; e-mail: pdhowe@unimelb.edu.au
Received 21 October 2013, in revised form 13 December 2013

Abstract. The maintenance of attention on moving objects is required for cognition to reliably engage with the visual world. Theories of object tracking need to explain on which patterns of visual stimulation one can easily maintain attention and on which patterns one cannot. A previous study has shown that it is easier to track rigid objects than objects that expand and contract along their direction of motion, in a manner that resembles a substance pouring from one location to another (vanMarle and Scholl 2003 *Psychological Science* **14** 498–504). Here we investigate six possible explanations for this finding and find evidence supporting two of them. Our results show that, first, objects that expand and contract tend to overlap and crowd each other more, and this increases tracking difficulty. Second, expansion and contraction make it harder to localize objects, even when there is only a single target to attend to, and this may also increase tracking difficulty. Currently, there is no theory of object tracking that can account for the second finding.

1 Introduction

We are often inundated with more visual information than we can process at any one time. To avoid being overwhelmed, we ignore most of it and attend to only a small fraction. Early theories of attention implicitly or explicitly assumed that attention was location-based and that all stimuli that happened to be located within the attended region would receive enhanced processing with the degree of the enhancement decreasing with the distance from the focus of attention (LaBerge 1983; Posner 1980; Treisman and Gelade 1980). Subsequent research has contradicted this assumption. For example, Egly et al (1994) showed observers two rectangles, briefly cueing the end of one of them. Although a subsequent probe would most often appear at the cued location, occasionally it would appear either at the uncued end of the cued rectangle or at an end of the other rectangle. This inconsistent probe was more likely to be detected when on the previously cued rectangle than on the uncued one, even though both locations were equidistant from the original cue. This study implied that attention is partially object-based in that it tends to spread more throughout an object than between different objects.

Although a number of studies have now provided additional empirical support that attention is partially object-based (Behrmann et al 1998; Moore et al 1998; Reuter-Lorenz et al 1996; Scholl et al 2001; Tipper and Behrmann 1996), it is not clear what counts as an object for the purposes of attention. Several factors seem to influence the degree to which an observer can direct and confine his or her attention to an object (Feldman 2007). For example, it is easier to attend to an object whose boundaries are continuous rather than containing gaps (Marino and Scholl 2005), that remains as a single entity rather than separating into two or more entities (Cheries et al 2008; Mitroff et al 2004), and that does not change its 2-D topological properties (Zhou et al 2010). However, there is at least one additional manipulation that can also affect an observer's ability to direct attention to a given object: the degree to which the object expands and contracts, as has been demonstrated by the multiple object tracking (MOT) paradigm (vanMarle and Scholl 2003).

The MOT paradigm has proven to be a powerful way of studying the deployment of attention in dynamic scenes (for reviews see Cavanagh and Alvarez 2005; Scholl 2001, 2009). In this paradigm, observers are shown a number of identical, moving objects and asked to keep track of a subset of them (Pylyshyn and Storm 1988). While in most MOT studies the objects were rigid and did not deform, a study by vanMarle and Scholl (2003) investigated tracking in a display where the objects expanded and contracted parallel to their direction of motion, as this sort of motion often occurs in everyday life when a nonrigid substance such as sand or a liquid pours from one location to another. It was found that it was much more difficult to track these substance-like objects than rigid objects that moved in an otherwise equivalent manner. This result was inconsistent with the literature because the deformation did not break the boundaries, subdivide the targets, or change the target's 2-D topological properties, so there was no a priori reason to suspect that it would hinder the deployment of attention and the observers' ability to track the objects. The results of vanMarle and Scholl seem consistent with those of an earlier study that investigated the tracking ability of 8-month-old infants (Huntley-Fenner et al 2002). In that study infants viewed a puppet stage. An object was then either lowered or poured onto the stage, and a screen was raised to hide it. A second object was then seen to be either lowered or poured into the area behind the screen. In the consistent condition the screen was removed to reveal both objects. Conversely, in the inconsistent condition, when the screen was removed, only one of the objects was present. Previous studies have demonstrated that infants will gaze at a scene longer when their expectations are violated (Simon et al 1995; Wynn 1992). Consequently, infants were deemed to have successfully tracked the number of objects that should have been behind the screen if they gazed at the screen longer in the inconsistent condition than in the consistent condition. The results indicated that the infants were able to track the objects when they were rigid but not when they were piles of sand that were poured on the stage. It was concluded that infants can readily track rigid objects but have difficulties tracking deforming substances such as sand (Hespos and vanMarle 2012; Huntley-Fenner et al 2002; vanMarle and Wynn 2011). It seems that this deficit persists into adulthood (vanMarle and Scholl 2003).

vanMarle and Scholl (2003) were able to show that the observed decrement in tracking accuracy was not due simply to their objects changing shape. In particular, they showed that deformations that did not result in expansion and contraction had at most a minor effect on tracking accuracy. Conversely, a decrement in tracking accuracy was observed when objects expanded and contracted, even when each object maintained integrity, remaining a single coherent entity. On the basis of these two findings, they postulate that the tracking difficulty was caused by the expansion and contraction rendering the locations of the objects ambiguous. However, they provided no direct evidence to support this assertion. In fact, there are at least six possible alternative explanations for their results, not excluded by their experiments.

The leading and trailing object edges in their 'slinky' condition, the condition where the objects expanded and contracted, moved at double the speed of the edges in the baseline condition where the objects did not expand and contract. Since it is harder to track faster-moving items (Alvarez and Franconeri 2007; Feria 2013), this might account for the decrement in tracking accuracy in the slinky condition relative to the baseline condition. A second possibility is that the effect may occur only when the objects are perceived to pour from one location to the other. In all the situations considered by vanMarle and Scholl (2003), the objects always expanded and contracted parallel to the direction of motion so always appeared to pour from one location to another. A third possibility is that the effect may not have been caused by the expansion and contraction per se but rather because the expanding and contracting objects tended to overlap and crowd each other more than those in the baseline condition. It is known that overlaps (Scholl et al 2001) and crowding (Intriligator and Cavanagh 2001) decrease tracking accuracy. A fourth possibility is that the expansion and contraction created motion signals that conflicted with the true motion of the centers of each object. Conflicting motion signals have previously been shown to increase tracking difficulty (St. Clair and Seiffert 2011). A fifth possibility is that the motion signals created by the expansion and contraction bias estimates of the object locations (Nishida and Johnston 1999; Snowden 1998; Whitney 2002), thereby making it harder for the observers to accurately localize the targets. Finally, it could be that contracting objects may simply attract less attention (Franconeri and Simons 2003; von Muhlenen and Lleras 2007). Whereas expansion often signals the approach of an object, contraction often signals that the object is receding (Gibson et al 1969), implying less of a need to immediately attend to it. Since diverting attention from objects makes it harder to track them (Kunar et al 2008), if contraction causes observers to pay less attention to the objects, this might explain why it is more difficult to track contracting objects.

Here, we ran six experiments to test these six alternative explanations for the vanMarle and Scholl (2003) finding. Our seventh experiment directly tested for the first time their hypothesis that it is harder to localize objects that expand and contract. This series of experiments allowed us to determine which factors make it hard to track objects that expand and contract. These discoveries regarding the stimulus characteristics that lead to the breakdown of object-based attention provide constraints on theories of how object representations are created and maintained.

2 Experiment 1

This experiment was designed to test the first two alternative explanations of the vanMarle and Scholl (2003) finding. We did this by running three conditions. In one condition, the pouring condition, the objects expanded and contracted parallel to their direction of motion as in the previous studies (Huntley-Fenner et al 2002; vanMarle and Scholl 2003), whereas in another condition, the orthogonal condition, they expanded and contracted orthogonal to their direction of motion. We compared the tracking accuracy in these two conditions with that in a third condition, the baseline condition, where the objects did not deform but moved in an otherwise identical manner. This allowed us to determine the degree to which the angle of the expansion and contraction, relative to the direction of motion, affects the tracking difficulty.

According to the first alternative explanation, the tracking deficit is caused by the increased speed of movement of the leading and trailing edges in the pouring condition relative to the baseline condition. These edges will move at the same speed in the baseline and orthogonal condition but at twice the speed in the pouring condition, due to the expansion and contraction moving in the same direction as the direction of motion in that condition. Thus, according to this explanation, we would expect the same tracking accuracy in the baseline and orthogonal conditions but a significantly reduced tracking accuracy in the pouring condition.

According to the second explanation, the tracking deficit should occur only when the expansion and contraction occur parallel to the direction of motion, as this is what is required for the objects to appear to pour from one location to another. This explanation therefore predicts that tracking accuracy should be the same in the baseline and orthogonal conditions, but much worse in the pouring condition.

Both explanations thus make the same prediction, allowing them to be tested simultaneously. Even if only one of them is correct, we would expect a significant deficit in tracking accuracy in the pouring condition relative to the orthogonal condition.

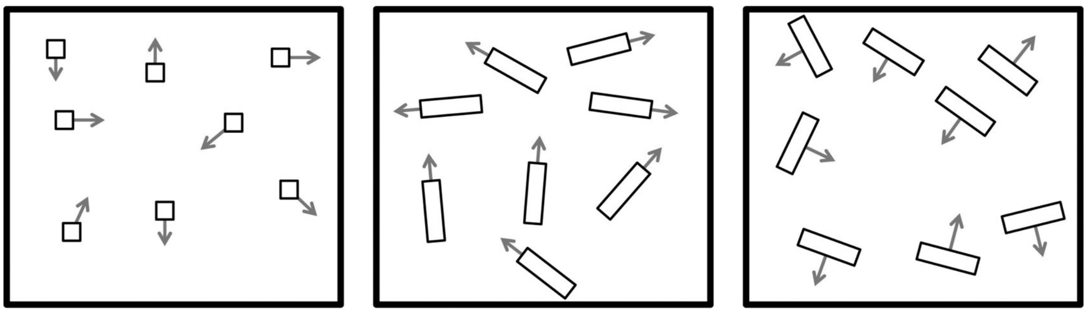
2.1 *Method*

2.1.1 *Participants.* There were twelve participants. Ages ranged from 18 to 41 years; six were male. A near-vision (40 cm) Good-Lite® eye chart was used to verify that all observers had normal or corrected-to-normal visual acuity. In addition, the observers reported not being color deficient. As with all the experiments reported here, the observers provided informed written consent, and the study was approved by the Department Human Ethics Advisory Group in the School of Psychological Sciences at the University of Melbourne. The experiments were conducted in accord with the World Medical Association Helsinki Declaration as revised in October 2008.

2.1.2 *Apparatus.* The participants viewed a 21-inch CRT monitor at a resolution of 1280 by 1024 pixels with a frame rate of 85 Hz at a distance of 60 cm. The usable display subtended 33 deg by 27 deg. Stimuli were constructed in MATLAB (Mathworks, Natick, MA) using the Psychophysics toolbox (Brainard 1997; Pelli 1997).

2.1.3 *Stimuli.* There were three conditions: *baseline*, *pouring*, and *orthogonal* (figure 1). The baseline condition was very similar to the *objects* condition of vanMarle and Scholl (2003). In this condition there were eight hollow white squares on a black background. Each square subtended 0.5 deg × 0.5 deg. The center of the display was divided into a 10 by 10 grid of imaginary (ie nonvisible) points subtending in total 14.2 deg by 11.5 deg.

For the sake of comparison, we used a similar motion algorithm to that which has been used previously (vanMarle and Scholl 2003). Each trial was subdivided into 1.25 s segments. At the start of each segment, eight of the grid points were selected, subject to the constraint that each grid point could be selected only once. Each of the eight grid points was then designated as the destination of a different square. Each square then moved to its designated grid point at a speed such that it arrived at its designated grid point exactly at the time the segment ended. Thus, in each segment each square would move a different distance at a different speed, but all squares would complete their movements at the same time. Immediately after one segment was completed, a new segment was initiated. During these movement phases, the squares could move over and occlude each other. See the experiment 1 baseline condition movie in the online supplementary material (http://dx.doi.org/10.1068/p7635) for a demonstration.



(a) (b) (c)

Figure 1. The three conditions of experiment 1. (a) Baseline condition: the objects do not change size. (b) Pouring condition: the objects expand and contract in their direction of motion, all in phase with each other. (c) Orthogonal condition: the objects expand and contract orthogonally to their direction of motion, again all in phase with each other. Grey arrows indicate the direction of motion. See the movies in the online supplementary material (http://dx.doi.org/10.1068/p7635) for a dynamic demonstration. Please note that the stimuli used in the experiments were of higher quality than those in the demonstration movies. In particular, their motion was not jerky.

The pouring condition was very similar to the baseline condition, except that each object expanded and contracted in its direction of motion, which gave the appearance of them pouring from one location to another. At the start of each segment each object was assigned a grid point. Each square would deform, extending into a rectangle until its leading edge reached its destination grid point. Then the trailing edge would follow, causing the rectangle to contract into a square. Thus, at the start and end of each segment all objects were squares, but in between they were rectangles. Note that the centers of the objects moved in an identical manner in this condition as in the baseline condition: during the course of each segment the center of each object started at one grid point and moved at a constant velocity to its destination grid point. The only difference between the two conditions was that in the pouring condition each object also expanded and contracted by an amount that depended on the distance it traveled. In this way, our pouring condition mimicked the slinky condition of vanMarle and Scholl (2003), in which the objects also appeared to pour from one location to another.

The orthogonal condition was identical to the pouring condition, except that the expansion and contraction occurred orthogonally to the direction of motion. As with the previous two conditions, during the course of a segment the center of each object moved at a constant velocity from its starting grid point to its destination grid point. Also, as before, each object expanded and contracted by the amount of distance traveled in that segment. Thus, the amount of expansion and contraction was exactly the same in both the orthogonal condition and the pouring condition. The only difference was that in the orthogonal condition the direction of the expansion and contraction was orthogonal to the motion of the object.

2.1.4 *Procedure.* Observers performed 15 practice trials before performing 120 trials in the main experiment. The three conditions were presented in equal numbers and randomly interleaved. Each trial comprised eight 1.25 s movement segments for a total trial duration of 10 s. For the first 2.5 s, four of the objects were colored red to indicate that they were the targets to be tracked. Then all objects reverted to white and continued to move about the screen. At the end of the trial the observers were required to use the mouse to click on the four targets from among the eight objects present. Accuracy was defined as the average percentage of targets that were correctly selected. Chance performance was therefore 50%.

2.1.5 *Results.* The results are shown in figure 2. A one-way repeated measures ANOVA showed that accuracy was significantly different in the three conditions ($F_{2,22} = 94.3$, $p < 0.001$, $\eta_p^2 = 0.896$). Planned *t*-tests revealed a significant difference between the baseline condition and the pouring condition ($t_{11} = 10.5$, $p < 0.001$, Cohen's $d = 3.23$). Like vanMarle and Scholl (2003), we found tracking accuracy to be much reduced in the pouring condition relative to the

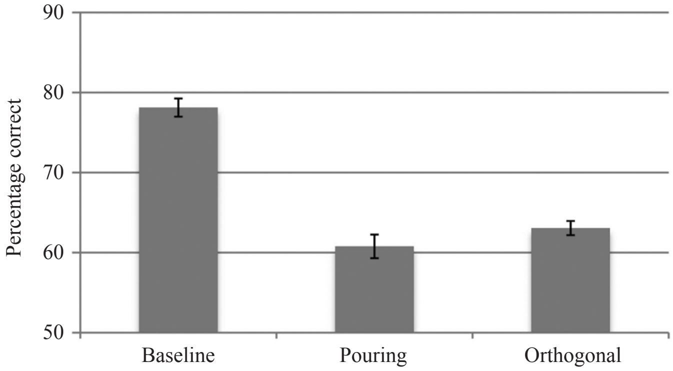


Figure 2. Tracking accuracy for experiment 1. Error bars are within-subject standard error of the mean (Cousineau 2005; Morey 2008).

baseline condition. Similarly, tracking accuracy was much lower in the orthogonal condition than in the baseline condition ($t_{11} = 16.2$, $p < 0.001$, Cohen's $d = 4.687$). Accuracy was not significantly different between the pouring and the orthogonal condition ($t_{11} = 1.6$, $p = 0.138$). These results demonstrate that the angle of the expansion and contraction has little if any effect on the tracking difficulty. Objects that repeatedly expand and contract are hard to track regardless of whether the expansion and contraction is parallel or orthogonal to their direction of motion.

These results argue against both the first and second alternative explanations listed above. Note in particular that the first explanation, the edge-based explanation, cannot be rescued by assuming that observers track one of the corners instead of the leading or trailing edges. The corners in the orthogonal condition move more slowly than those in the pouring condition, by a factor of $1/\sqrt{2}$. This hypothesis would therefore still predict performance to be significantly greater in the orthogonal condition than in the pouring condition.

3 Experiment 2

The purpose of this experiment was to test the third alternative explanation, that tracking objects that expand and contract is difficult because such objects tend to overlap and crowd each other more. The longer an object is, the greater the likelihood it will overlap with another object. Given that joining targets to distractors increases tracking difficulty by marking it harder to distinguish the targets from distractors (Keane et al 2011; Scholl et al 2001), which would be expected from Palmer and Rock's (1994) principle of uniform connectedness, overlaps would also be expected to increase tracking difficulty. In addition, the longer the objects are, the closer they will be to each other, so the greater the expected crowding. Crowding has also been shown to decrease tracking accuracy (Intriligator and Cavanagh 2001). The second experiment addressed these issues by using conditions where objects were long but did not change in length. This allowed us to measure the degree to which overlapping and increased crowding increased tracking difficulties in our displays, in the absence of expansion and contraction.

3.1 *Method*

3.1.1 *Participants and apparatus.* There were twelve participants, and again ages ranged from 18 to 41 years. Nine were male. All had normal or corrected-to-normal visual acuity and reported that they were not color deficient. The apparatus was the same as experiment 1.

3.1.2 *Stimuli.* There were four conditions: *baseline*, *expansion/contraction*, *average length*, and *max length*. The baseline condition was very similar to the baseline condition of the previous condition. As before there were eight hollow white squares, each of which subtended 0.5 deg × 0.5 deg, presented on a black background. However, the squares now moved in a continuous fashion, bouncing off the walls of an imaginary square, so that their centers were confined to a 15 deg × 15 deg region of space. Each square moved at a constant 8 deg s^{-1}.

In the expansion/contraction condition the objects were rectangles that continuously varied their length in a sinusoidal fashion with a time period of 1.5 s. At their point of maximum contraction they had dimensions of 0.5 deg × 0.5 deg, and at their maximum expansion they had dimensions of 14.5 deg × 0.5 deg. For each rectangle its orientation was chosen at random, was fixed for the duration of the trial, and was therefore not correlated with its direction of motion. As with the baseline condition, the centers of the rectangles did not leave the central 15 deg × 15 deg region of the screen, bouncing if they reached the edge of this area. Thus, the centers of the rectangles and the centers of the baseline condition's squares moved in identical fashion.

In the average-length condition the objects were rectangles that had fixed dimensions of 7.5 deg × 0.5 deg. Thus, their lengths were the average of those of the expansion/contraction condition. Their orientations were chosen randomly and fixed for the duration of the trial. As for the other conditions, their centers were confined to the central 15 deg × 15 deg region of the screen.

Finally, the max-length condition was identical to the average-length condition, except that the rectangles had dimensions of 14.5 deg × 0.5 deg, which corresponds to the maximum length present in the expansion/contraction condition.

3.1.3 *Procedure.* Observers performed 15 practice trials before performing 160 trials in the main experiment. The four conditions were presented in equal numbers and were randomly interleaved with each other. Each trial lasted 10 s, and for the first 2.5 s four of the objects were colored red to indicate that they were the targets to be tracked. Then all objects reverted to white and continued to move about the screen. At the end of the trial, the observers were required to use the mouse to click on the four targets.

3.1.4 *Results.* The results are shown in figure 3. A one-way repeated measures ANOVA showed that accuracy was significantly different in the four conditions ($F_{3,33} = 41.3$, $p < 0.001$, $\eta_p^2 = 0.790$). Planned *t*-tests were then used to examine the differences between the four conditions. Consistent with the previous experiment, it was found that accuracy in the expansion/contraction condition was much less than that in the baseline condition ($t_{11} = 7.49$, $p < 0.001$, Cohen's $d = 2.60$). However, accuracy in the average-length condition was significantly greater than that in the expansion/contraction condition ($t_{11} = 7.23$, $p < 0.001$, Cohen's $d = 2.263$). Conversely, accuracy in the max-length condition was not significantly greater than that in the expansion/contraction condition ($t_{11} = 1.05$, $p = 0.316$). As expected, accuracy was significantly greater in the average-length condition than in the max-length condition ($t_{11} = 7.72$, $p < 0.001$, Cohen's $d = 2.32$). These results show that longer objects are harder to track than shorter objects, even when the objects do not change length. This suggests that part of the difficulty associated with the expansion/contraction conditions of the previous experiment is that the expansion led to the objects having longer length during part of the trial than the objects in the baseline condition.

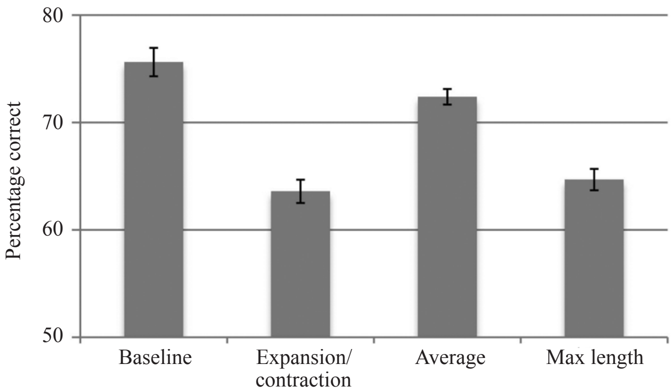


Figure 3. Tracking accuracy for experiment 2. Error bars are within-subject SEM (Cousineau 2005; Morey 2008).

4 Experiment 3

The previous experiment showed that tracking difficulty increased as the length of the rectangles increased. While this could be because longer rectangles tend to overlap and crowd each other more, it could also be due to the fact that it is harder to perceive the motion direction of extended objects (Lorenceau et al 1993).

Motion processing involves multiple stages and the first signal extracted is the direction of the motion of an edge (Lennie 2003; Marr 1982). The directional signal relating to any one-dimensional edge is known to within only ±90°, a property known as the 'aperture problem' (Hildreth and Koch 1987; Horn and Schunk 1981; Marr and Ullman 1981). Determination of the velocity more precisely is generally through the combination of the outputs of multiple local motion detectors. Thus, the overall motion percept is affected by several factors, and attention has a significant role (Cavanagh 1992).

Because attention tends to be directed to the centers of extended objects (Alvarez and Scholl 2005; Feria 2008), potentially overweighting the overall motion to be orthogonal to the extended edge, the direction of motion of extended objects may be misperceived. Because motion direction is at least sometimes used to improve tracking performance, incorrect registration of the current direction of motion may impair performance (Howe and Holcombe 2012), even if the object's changing position is correctly perceived. If this is the reason for poorer tracking performance with longer rectangles, then we should be able to improve tracking performance by placing disks at the center of each rectangle, as they should disambiguate the overall object motion. Experiment 3 tested this prediction.

4.1 *Method*

4.1.1 *Participants and apparatus.* There were 12 participants whose ages ranged from 18 to 25 years. Nine were male. The apparatus and participant selection procedure was the same as experiment 1.

4.1.2 *Stimuli.* There were three conditions: *baseline*, *max length*, and *disks* (figure 4). The first two conditions were identical to the corresponding conditions of the previous experiment. The third condition was identical to the max-length condition, except that a white disk with a diameter of 0.5 deg was placed at the center of each rectangle, providing each rectangle with a feature at its center that would generate unambiguous motion signals, thereby minimizing the impact of the aperture problem on these stimuli.

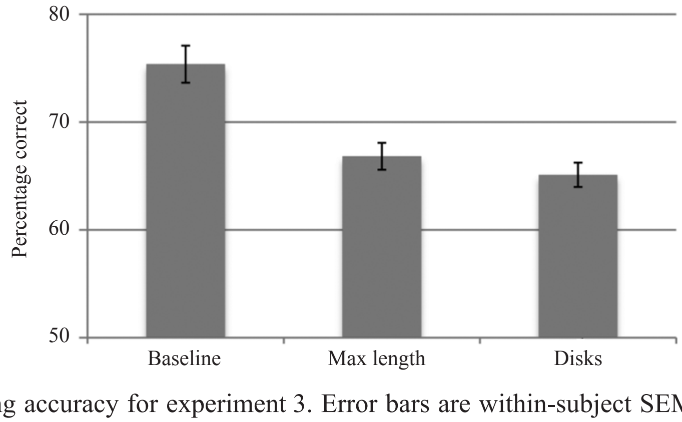


Figure 4. Tracking accuracy for experiment 3. Error bars are within-subject SEM (Cousineau 2005; Morey 2008).

4.1.3 *Procedure.* The procedure was identical to that in the previous experiment, except that, as there were only three conditions, there were only 120 trials in the main part of the experiment (40 trials per condition).

4.1.4 *Results.* The results are shown in figure 4. A one-way repeated measures ANOVA showed that accuracy was significantly different in the three conditions ($F_{2,22} = 23.7$, $p < 0.001$, $\eta_p^2 = 0.758$). Planned *t*-tests were then used to examine the differences between the three conditions. Consistent with the previous experiment, it was found that accuracy in the baseline condition was much higher than that in the max-length condition ($t_{11} = 4.63$, $p = 0.001$, Cohen's $d = 1.81$). Similarly, accuracy was much greater in the baseline condition than in the disks condition ($t_{11} = 5.87$, $p < 0.001$, Cohen's $d = 1.76$). However, accuracy was not significantly different between the max-length condition and the disks condition ($t_{11} = 1.57$, $p = 0.144$). Ensuring that the centers of the rectangles provided unambiguous motion signals did not improve tracking accuracy. This suggests that the decrement in tracking is not caused by an aperture problem engendered by the observers directing their attention to the center of each object rather than integrating the motion signals generated at the end of the rectangle to calculate the correct motion direction of each object.

5 Experiment 4

The previous experiment has shown that long rectangles are hard to track, even when their centers contain a feature that is easy to track, suggesting that part of the tracking difficulty of long rectangles is instead caused by them overlapping with each other, disrupting their identity as an independent 'object'. This in turn suggests that part of the reason why it is harder to track objects that expand and contract is that such objects tend to overlap more with each other than those that remain as small squares throughout the trial. The purpose of experiment 4 was to determine whether expansion and contraction provide any additional decrement in tracking accuracy, beyond that which be expected due to the increase in overlaps. Experiment 4 investigated this issue by arranging for objects to expand and contract but not to overlap with each other.

5.1 *Method*

5.1.1 *Participants and apparatus.* There were twelve participants, and their ages ranged from 18 to 33 years. Three were male. The apparatus and participant selection procedure was the same as before.

5.1.2 *Stimuli.* There were three conditions: *baseline*, *no overlap*, and *overlap*. For the sake of comparison, we purposely used a similar algorithm to that employed in vanMarle and Scholl (2003), modified so that in two of the conditions the objects would not overlap with each other or even touch each other. The first condition was similar to the baseline condition of experiment 1. The stimuli comprised eight white hollow squares, each of which subtended 0.5 deg × 0.5 deg. The center of the display was divided into a 10 by 10 grid of (imaginary) points subtending in total 14.2 deg by 11.5 deg. At the start of the trial, each square would be assigned to a random grid point, subject to the condition that no two squares could be on either the same row or the same column. The trial was 10.5 s long and divided into 12 segments, each of which lasted 0.875 s. On alternate segments all the squares would either move horizontally or vertically. Because all objects started on different rows and columns, this motion sequence guaranteed that they would never overlap, or even touch each other, during the movement phase. For a horizontal segment, each square would be assigned a destination grid point on the same row, with the caveat that at the end of the movement sequence no two objects would be on the same column. During the course of the segment, each square would move from its starting grid point to its destination grid point. Thus, each square could move a different distance and at a different speed, but all squares would arrive at their destination grid points at the same instant. The vertical segments operated in the same fashion, except that each square was confined to move vertically and the destination grid points were selected so that no two objects finished the sequence sharing the same row. In this fashion, over time each square could move anywhere in the display, but no two squares would ever pass over each other or even touch each other.

The objects in the no-overlap condition moved in exactly the same way as those in the baseline condition, except that they expanded and contracted in their direction of motion. Both the expansion and contraction were symmetrical about the center point of each object and had an amplitude equal to half the distance traveled during the segment. Consequently, in the first half of the segment only the leading edge of the rectangle would move while the trailing edge would be stationary. Then in the second half of the segment the leading edge would be stationary whereas the trailing edge would now move. This condition was therefore very similar to the pouring condition of experiment 1. The main difference was that, unlike in experiment 1, here all the objects were guaranteed not to overlap.

The overlap condition was identical to the no-overlap condition, except that the expansion and contraction occurred orthogonal to the direction of motion. As before, the expansion and contraction was symmetrical about the center of each object, with an amplitude equal to half the distance traveled in that segment. As a consequence, the objects in this condition would often overlap each other in passing.

5.1.3 *Procedure.* The procedure was identical to that in experiment 3

5.1.4 *Results.* The results are shown in figure 5. Owing to the different motion sequence and different display parameters, the accuracy in the baseline condition in this experiment was higher than that in baseline conditions in the previous experiments. Consequently, we performed only within-experiment statistical analyses. A one-way repeated measures ANOVA showed that accuracy was significantly different in the three conditions ($F_{2,22} = 60.5$, $p < 0.001$, $\eta_p^2 = 0.846$). Planned *t*-tests were used to examine the differences between the three conditions. It was found that accuracy in the baseline condition was much greater than that in the no-overlap condition ($t_{11} = 6.63$, $p < 0.001$, Cohen's $d = 2.92$), demonstrating that expansion and contraction in the absence of overlap (or touching) results in a decrement in tracking accuracy. Similarly, accuracy was much greater in the baseline condition than in the overlap condition ($t_{11} = 13.7$, $p < 0.001$, Cohen's $d = 4.37$). Accuracy was also greater in the no-overlap condition than in the overlap condition ($t_{11} = 2.81$, $p = 0.017$, Cohen's $d = 0.97$). The no-overlap and overlap conditions differed not only in whether or not there was overlap but also in the direction of expansion and contraction. However, experiment 1 showed that the direction of expansion and contraction has no significant effect on tracking accuracy. Thus, the only pertinent difference between these two conditions is the degree of overlap, supporting our finding from experiment 2 that overlap per se reduces tracking accuracy. In conclusion, while this experiment confirms that overlap decreases tracking accuracy, it additionally shows that, even in the absence of overlap, expansion and contraction inhibit tracking.

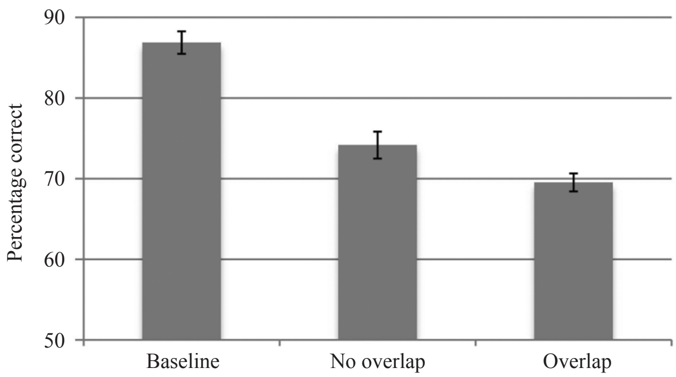


Figure 5. Tracking accuracy for experiment 4. Error bars are within-subject SEM (Cousineau 2005; Morey 2008).

6 Experiment 5

It remains unclear why expansion and contraction per se disrupt tracking. Clearly, tracking involves motion perception, since a stimulus without motion would not require tracking. However, the MOT literature has provided few insights into the interaction between the spatiotemporal (motion) properties of the stimulus elements and tracking performance. Some models of MOT do not even include a parameter describing the spatiotemporal structure of the elements (Vul et al 2010). Despite this, it is clear that observers can make use of objects' motion information when tracking (Fencsik et al 2007; Horowitz and Cohen 2010; Horowitz et al 2006; Howe and Holcombe 2012; Shooner et al 2010; St. Clair and Seiffert 2011; St. Clair et al 2010), though this depends on a number of parameters—for example, tracking load (Fencsik et al 2007; Howe and Holcombe 2012). Expansion and contracting is a nonrigid motion (Fleet et al 1996; Marr 1982); the contours' vectors have a more complex relationship with the vector specifying the displacement of the object's overall position (its translation), and this may impair tracking performance. For instance, the trailing contour of an expanding object may actually be moving opposite the overall displacement direction. If this is the reason for the tracking deficit, then the decrement in tracking accuracy should occur only if the expansion and contraction occurs while the objects are actually translating. In this experiment we tested this prediction.

6.1 *Method*

6.1.1 *Participants and apparatus.* There were twelve participants, and ages ranged from 20 to 41 years. Six were male. The apparatus and participant selection procedure was the same as before.

6.1.2 *Stimuli.* There were three conditions: *baseline*, *together*, and *separate*. The first condition was similar to the baseline condition of experiment 4, except that the squares moved for only part of the time. All the squares moved for 0.5 s and then paused for 1.25 s before moving again. This process was repeated for a total of 8 cycles in a single trial.

The together condition was similar to the baseline condition, except that the objects expanded and contracted in the same direction of motion, reaching a maximum extent of 12 deg before contracting back to 0.5 deg by 0.5 deg squares. The expansion and contraction were confined to the portion of the trial where the objects were moving. As in the baseline condition, the objects did not overlap with each other at any point in the trial.

The separate condition was identical to the together condition, except that the expansion and contraction did not occur during the movement phase. Instead, the objects expanded for 0.25 s, paused for 0.25 s, moved for 0.5 s, paused for 0.25 s, contracted for 0.25 s, and then paused for another 0.25 s. Thus, there were pauses between the expansion and contraction phases and the movement phase.

6.1.3 *Procedure.* The procedure was identical to that in experiment 3.

6.1.4 *Results.* The results are shown in figure 6. Because the motion sequence was different in this experiment from that in the previous experiment, the accuracy in the baseline condition was also different from that in the previous experiment. Consequently, we will again confine ourselves to a within-experiment statistical analysis. A one-way repeated measures ANOVA showed that accuracy was significantly different in the three conditions ($F_{2,22} = 46.5$, $p < 0.001$, $\eta_p^2 = 0.809$). Planned *t*-tests were then used to examine the differences between the three conditions. It was found that accuracy in the baseline condition was much greater than that in the together condition ($t_{11} = 8.34$, $p < 0.001$, Cohen's $d = 3.62$). Similarly, accuracy was much greater in the baseline condition than in the separate condition ($t_{11} = 6.59$, $p < 0.001$, Cohen's $d = 2.01$). Accuracy was also greater in the separate condition than in the

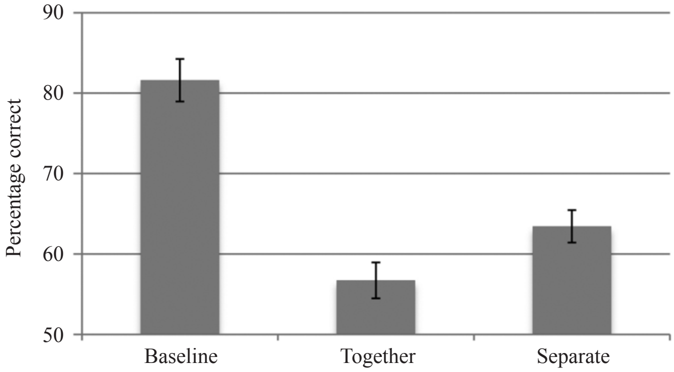


Figure 6. Tracking accuracy in experiment 5. Error bars are within-subject SEM (Cousineau 2005; Morey 2008).

together condition ($t_{11} = 3.03$, $p = 0.012$, Cohen's $d = 1.08$). These results show that while expansion and contraction impair tracking even when they do not occur simultaneously with the movement, they impair tracking the most when they do. While the former result suggests that expansion and contraction per se can directly disrupt tracking, perhaps by disrupting the localization of the objects, the latter result suggests that they can also hinder tracking indirectly by interfering with the motion signals that could otherwise aid tracking (Fencsik et al 2007; Howe and Holcombe 2012).

7 Experiment 6

Previous studies have found that contracting stimuli are less attentionally engaging than those that expand (Franconeri and Simons 2003; von Muhlenen and Lleras 2007), possibly because shrinking often denotes an object that is moving away from an observer while expansion often denotes an object approaching the observer, so is in more immediate need of attention (Gibson et al 1969). This phenomenon provides a potential explanation for our findings: stimuli that expand and contract may be harder to track than those that remain at a fixed length because the former stimuli are ignored and thus lost when they are in their contracting phase. This hypothesis makes a clear prediction. Only contraction should hinder tracking. Expansion should either have no effect on tracking performance or possibly even increase tracking performance by attracting additional attentional resource to the targets (Franconeri and Simons 2003; von Muhlenen and Lleras 2007). Experiment 6 tested this prediction. In one condition the objects only expanded, whereas in the other they only contracted. By comparing the tracking difficulty in both these conditions relative to a condition where the objects neither expanded nor contracted, we could determine to what degree expansion and contraction hinder tracking.

7.1 *Method*

7.1.1 *Participants and apparatus.* There were twelve participants, and ages ranged from 19 to 43 years. Five were male. The apparatus and participant selection procedure was the same as before.

7.1.2 *Stimuli.* There were four conditions: *baseline*, *expand*, *contract*, and *long bars*. The first condition was similar to the baseline condition of experiment 2. There were eight hollow white squares, each of which subtended 0.5 deg × 0.5 deg, presented on a black background. The squares in this experiment moved in a continuous fashion, bouncing off the walls of an imaginary square, so that their centers were confined to a 15 deg × 15 deg region of space. Each square moved at 5 deg s^{-1}.

The expand condition was similar to the baseline condition, except that the objects suddenly expanded halfway through the trial. At the start of the trial all the objects were small squares, identical to those in the baseline condition. Halfway through the trial the objects all suddenly expanded, becoming hollow rectangles of dimensions 14.5 deg × 0.5 deg. The orientation of each rectangle was random and the duration of the expansion phase was 0.2 s.

In the contract condition the reverse occurred. The objects started off as long rectangles and halfway through the trial they suddenly contracted into squares (0.5 deg × 0.5 deg). In the long-bars condition the objects remained as long bars for the duration of the trial.

7.1.3 *Procedure.* The procedure was similar to that in the previous experiment. At the start of the trial four of the objects turned red for 2.5 s to indicate that these were the targets to be tracked. Then all the objects reverted to white and moved about the monitor for 5 s. At the end of the trial the observer used the mouse to indicate the locations of the targets.

7.1.4 *Results.* The results are shown in figure 7. In both the contract and expand conditions each object spent exactly half the trial as a long rectangle and the other half as a small square. Consequently, if expansion/contraction per se did not affect tracking performance, we would expect tracking accuracy in these two conditions to be approximately halfway between the tracking accuracies in the baseline and long-bars conditions. Thus, for each observer we took the mean of the baseline and long-bars conditions to create a comparison point. By comparing the tracking accuracy in the contract and expand conditions to this comparison point we could determine whether either expansion or contraction reduces tracking accuracy. Planned t-tests found that tracking accuracy in the contract condition ($t_{11} = 8.54$, $p < 0.001$, Cohen's $d = 3.08$) was significantly less than the average of the baseline and long-bars condition. Furthermore, tracking accuracy in the contract condition was significantly less than that in the expand condition ($t_{11} = 4.94$, $p < 0.001$, Cohen's $d = 6.04$). However, contrary to the above hypothesis, tracking accuracy in the expand condition ($t_{11} = 6.47$, $p < 0.001$, Cohen's $d = 1.89$) was also significantly less than the average of the baseline and long-bars condition. This showed that even expansion impairs tracking accuracy.

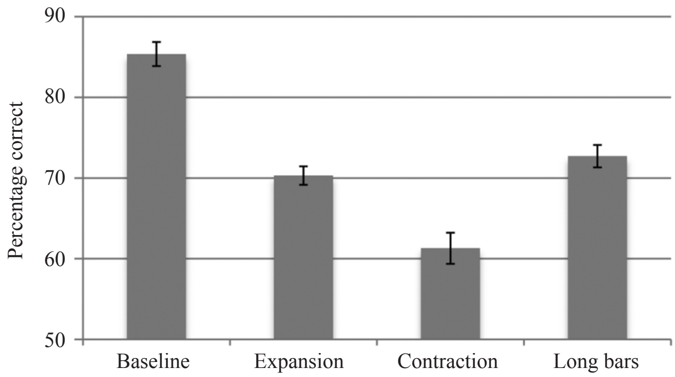


Figure 7. Tracking accuracy in experiment 6. Error bars are within-subject SEM (Cousineau 2005; Morey 2008).

8 Experiment 7

The previous experiment showed that both expansion and contraction hinder tracking. Experiment 5 showed that this hindrance occurs even when the expansion and contraction occurs while the objects are not moving. This suggests that expansion and contraction may disrupt tracking by hindering the ability of observers to accurately localize the objects.

Many studies have found that the perceived position of an object is biased in its direction of motion (for a review see Whitney 2002). However, the bias in the motion direction effect alone would not be expected to increase the average localization error for expanding or contracting objects. Consider that, for an expanding or contracting object, the position bias at opposite ends of the object should be equal and opposite and thus cancel out.

In addition to measuring the positional bias induced by motion, De Valois and De Valois (1991) also assessed the effect of motion on the precision of localization, by measuring thresholds in their Vernier paradigm. Thresholds were 31 arcsec in the motion conditions against 24 arcsec in the stationary condition. Expressed as a proportion, this constitutes a substantial increase (29%), but the absolute precision is still extraordinarily high. Errors on the order of magnitude of arcsec would not be expected to significantly impair tracking in our present conditions. However, it is far from clear how an effect on a Vernier judgment with two drifting Gabors would translate to localization errors of an expanding or contracting bar. To investigate the possibility that motion might impair localization enough to impair tracking, in experiment 7 observers were asked to report the location of the center of an expanding, contracting, or unchanging bar.

8.1 *Method*

8.1.1 *Participants and apparatus.* There were twelve participants, and ages ranged from 19 to 43 years. Five were male. The apparatus and participant selection procedure was the same as before.

8.1.2 *Stimuli and procedure.* There were two conditions: *fixed length* and *expansion/contraction*. In the fixed-length condition a white hollow bar was flashed on the computer monitor for a period of 200 ms. The length of the bar did not change and it had dimensions of 7.5 deg × 0.5 deg. Thus, it very closely resembled the stimuli of the previous experiments. Immediately after it disappeared, a pointer appeared and the observer was prompted to use the mouse to position the pointer where the center of the bar had been. The expansion/contraction condition was identical to the fixed-length condition, except that while the bar was visible its length varied in a sinusoidal fashion with a mean of 7.5 deg and an amplitude of 3.5 deg. The time period of the sinusoidal movement was 200 ms, and the initial phase was either 0 or π radians. Consequently, the bar always started and finished its presentation at its mean length. Observers performed 20 practice trials before performing 300 experimental trials, split equally between the two conditions, in a random, interleaved order.

8.1.3 *Results.* For each trial we calculated the error—the distance between the reported position and the actual position. The results are shown in figure 8. Average error was significantly worse in the expansion/contraction condition (0.88 deg) than in the fixed-length condition (0.77 deg) ($t_{11} = 4.92$, $p < 0.001$, Cohen's $d = 1.43$). This poor localization performance in the expansion/contraction condition was not significantly affected by whether the oscillation began with expansion (average error 0.86 deg) or with contraction (0.90 deg) ($t_{11} = 0.81$, $p = 0.436$).

To assess the predominant direction of the errors, we calculated the average component of the error along the same axis as the expansion/contraction (parallel to the orientation of the bar) and the component orthogonal to the motion. Errors were significantly worse in the expansion/contraction condition along the axis of motion ($t_{11} = 4.72$, $p = 0.001$, Cohen's $d = 1.36$), but not for the axis orthogonal to the motion ($t_{11} = 0.939$, $p = 0.368$). In summary, these results show that expansion and contraction do decrease an observer's ability to accurately localize an object, and this may be specific to the axis of expansion/contraction. Although the size of the localization error is fairly small in comparison with the size of the objects, this error could still be enough to cause observers to confuse the targets with the distractors, given that the objects would often pass close by or even over each other.

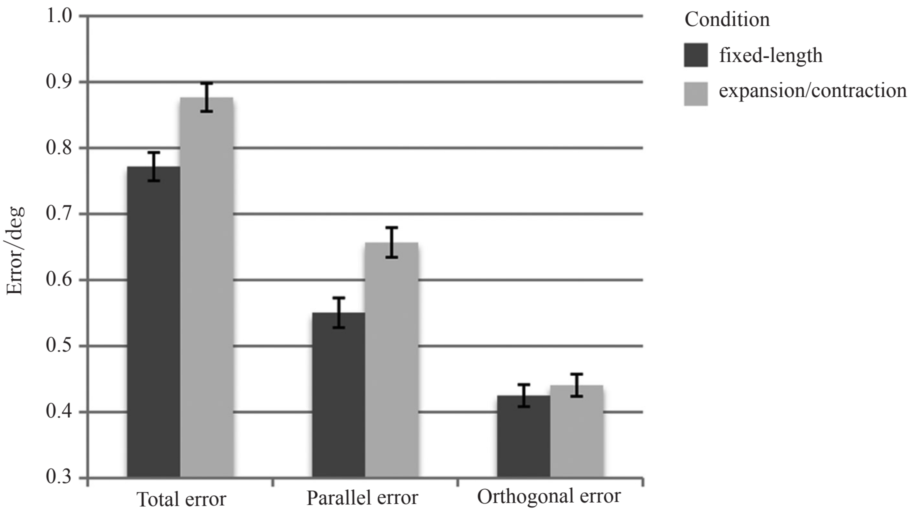


Figure 8. Localization error in degrees of visual angle for experiment 7, for both the fixed-length condition and the expansion/contraction condition. The leftmost pair of columns shows the total error, the middle pair shows the component of the error measured parallel to the orientation of the stimulus bar, and the rightmost pair shows the component of the error measured orthogonal to the orientation of the stimulus bar. Error bars are within-subject SEM (Cousineau 2005; Morey 2008).

9 General discussion

It has been shown that objects that expand and contract in their direction of motion in a manner that resembles substances pouring from one location to another are harder to track than those that are rigid but otherwise move in an identical fashion (vanMarle and Scholl 2003). It was argued that the impairment in tracking was due to the dynamic expansion and contraction rendering the location of the objects uncertain, but no direct evidence was provided to support this conclusion. In our study we evaluated six alternative explanations for these data and directly tested the hypothesis that expansion and contraction hinder the localization of objects.

In our first experiment we investigated to what extent tracking is hindered when the expansion and contraction occur orthogonally to the direction of motion. These stimuli do not appear to pour from one location to another, unlike the stimuli of the original vanMarle and Scholl (2003) study that expanded and contracted in their direction of motion. This first experiment thus assessed whether the appearance of pouring was critical by using expansion and contraction orthogonal to the motion. Compared with the stimuli that expand and contract orthogonally to their direction of motion, those stimuli that expand and contract parallel to their direction of motion should also yield poorer performance if tracking primarily uses objects' leading edges, trailing edges, or their corners. Since tracking difficulty increases with speed (Alvarez and Franconeri 2007), such a mechanism would predict tracking accuracy to be higher in the orthogonal condition, as in that condition those landmarks all move less quickly than they do in the parallel condition. Our result was that performance was equivalent in both conditions, allowing us to discard both alternative explanations.

Our second experiment investigated a third possible explanation for the poorer tracking performance in displays where objects expand and contract. Objects that expand and contract are necessarily elongated at times. This elongation should result in greater crowding or lateral masking as well as greater contour masking due to more instances of overlap among the objects. Such factors can, of course, impair tracking (Franconeri et al 2010; Intriligator and Cavanagh 2001; Scholl et al 2001) just as they impair object perception generally (Pelli and Tillman 2008). This experiment verified that longer objects are indeed harder to track, even when they do not expand and contract, thus providing a partial explanation for vanMarle and Scholl's finding. This experiment did not determine whether the increased tracking difficulty was due to crowding or instances of overlap, as both would be increased for longer objects.

Our third experiment tested the importance of having cues to object motion near their centers. When tracking objects, attention seems to be concentrated at the objects' centers (Alvarez and Scholl 2005; Feria 2008). For an elongated object the contours at its center signal only the component of the object's motion perpendicular to its orientation. It might have been that elongated objects are hard to track not just because they tend to overlap and crowd each other more but also because it is harder to discern their direction of motion. Experiment 3 tested this suggestion by adding contours at the centers of the elongated objects so as to make clear their direction of motion. As this did not increase tracking accuracy, it appears that tracking makes effective use of remote contour motion.

The previous experiments suggest that elongated objects are harder to track in part because they tend to overlap each other. Experiment 4 documented the influence of an additional, more interesting factor—the expansion and contraction themselves. By comparing a condition where expanding and contracting objects would overlap with one where they did not, we were able to determine that most of the tracking impairment was caused by the expansion and contraction of the objects per se. This result is in agreement with what vanMarle and Scholl (2003) concluded on the basis of their experiment 1b. However, experiment 4 here went beyond this previous experiment in at least two ways. First, it compared a condition with expansion and contraction and no overlap with an otherwise identical condition with overlap. This allowed us to directly compare the relative contributions to tracking difficulty of overlap and expansion and contraction per se. Second, in the present study, in the no-overlap condition the objects never touched each other. Conversely, in the nonoverlap condition of experiment 1b of vanMarle and Scholl (2003) the target and distractor objects would sometimes come in contact as they passed by each other. This contact might substantially impair tracking, as a previous study found marked impairments when distractors were connected to targets (Scholl et al 2001). Thus, it might have been that the decrement in tracking observed by vanMarle and Scholl (2003) and attributed to expansion and contraction was instead due to contact among the objects. Experiment 4 avoided this potential confound.

Having shown that expansion and contraction per se impair tracking we then asked what it was about the expansion and contraction that so impairs tracking performance. Part of the problem may have been that the expansion-related and contraction-related motion signals conflict with or mask the motion signals indicating each object's true direction of movement. However, experiment 5 showed that this cannot be the whole explanation, as it found that expansion and contraction hinder performance substantially even if the expansion and contraction occur when the objects are not otherwise moving.

Experiment 6 tested the expansion and contraction phases in separate conditions to investigate whether each could separately hinder tracking and, if so, which hindered tracking the most. The results showed that contraction hinders tracking much more than does expansion. Although this might in part be because contracting objects are not as attentionally engaging as expanding ones (Franconeri and Simons 2003; von Muhlenen and Lleras 2007), expansion also hindered tracking. A possible reason that changing the size of an object (for both directions of change, expansion and contraction) impairs tracking is that it hinders localization of the objects (Nishida and Johnston 1999; Snowden 1998; Whitney 2002).

Experiment 7 tested this hypothesis. Participants were presented with a single object that briefly changed in size before disappearing, and the participants then tried to click on the object's center. Localization errors were much greater in the expansion/contraction condition than in the fixed-length condition. This apparently novel finding might explain much of the difficulty in tracking objects that change in size. The results of experiment 5 in particular suggest this. In experiment 5, despite the objects not otherwise moving when they were changing in size, the expansion/contraction still impaired tracking markedly.

9.1 *Theoretical implications*

The above seven experiments have revealed that observers find it hard to track objects that expand and contract, likely for two reasons. First, objects that expand and contract tend to overlap and crowd each other more, and this increases tracking difficulty. It is not surprising that overlap impairs tracking because the different objects then ought to mask each other (Breitmeyer and Öğmen 2006). To some degree it is still an open question whether overlap and proximity disrupt tracking more than would be expected from simpler, nontracking perception tasks. However, previous studies involving connecting targets to distractors with line segments suggest that contact among objects is disruptive to tracking (Howe et al 2012; Scholl et al 2001). In fact, target-distractor grouping in general tends to disrupt tracking (Erlikman et al, in press; Suganuma and Yokosawa 2006). Although Scholl et al (2001) and Howe et al (2012) did not fully control for low-level masking, the size of the tracking impairment was so large that it seems safe to conclude, as those researchers did, that contact among objects is particularly disruptive to tracking. A likely explanation is that tracking processes operate on the entire connected assemblage as a single unit to be tracked and does not differentiate among its parts (for review see Scholl 2001). Tracking, like certain other processes of attention, may be object-based.

A second factor that may impair tracking is that objects that expand and contract are harder to localize, although this impairment in localization has not been directly shown to hinder tracking. The reason for this impairment is unclear, and our data do not allow us to draw any definite conclusions. As noted above, the decrement in localization performance for contracting objects as compared with expanding objects might in part be because contracting objects are less likely to attract attention (Franconeri and Simons 2003; von Muhlenen and Lleras 2007), perhaps because they resemble objects that are receding (Gibson et al 1969). However, this explanation is incomplete because expansion also hindered localization. The effect on localization might arise because motion can bias position estimates (Nishida and Johnston 1999; Snowden 1998; Whitney 2002), thereby making them less reliable.

No published theory of tracking appears able to explain this deleterious effect of expansion and contraction. Each of the published theories was designed to explain tracking findings that seem completely unrelated to expansion and contraction.

The theory of Franconeri et al (2013), for example, emphasizes spatial competition among objects. It would not predict that expanding or contracting objects are particularly difficult to track or localize. Kazanovich and Borisyuk (2006) provided a full neural network model that uses a temporal oscillator code and a single network layer assigned to each target. While this network might well confuse targets and distractors when they contact each other, there is no reason for contraction or expansion to disrupt the network otherwise. Tracking theories with a serial component (Oksama and Hyona 2008; Pylyshyn and Storm 1988; Yantis 1992) propose that targets' positions are updated too infrequently for tracking to always succeed. Vul et al (2010) added Bayesian inference, so that objects are localized in an optimal way given the uncertainty of sensory measurements of their positions. Holcombe and Chen (2013) documented temporal interference between targets and distractors, but again nothing to suggest that expansion and contraction would be deleterious.

Because we found expansion and contraction to impair explicit perceptual localization of even a lone object, the deficit may arise in perceptual stages prior to tracking. Tracking appears to operate on object-level representations (Scholl 2001) and may occur in the parietal regions (Culham et al 1998, 2001; Howe et al 2009; Jovicich et al 2001), while low-level cortical areas trade predominantly in edges or contours.

Future studies should assess the extent to which localization of even single contours is impaired by displacement. Because of simple temporal integration, a moving contour will be represented in several locations (blurred), which might be expected to impair localization. Yet a previous study found very good localization for moving contours, although this was strongly modulated by eccentricity (De Valois and De Valois 1991). That study may have been limited to the Vernier paradigm. Additional psychophysics will be needed to assess whether any associated impairment is enough to explain the large impairment found here for localizing an expanding object. If it is not large enough, the localization deficit may arise at object-based or shape-based processes, and studying it will provide insights into these mid-level representations.

Acknowledgments. This manuscript was funded by research grants to PH from the University of Melbourne and ARC FT0990767 to AOH. We thank Charlotte Hudson and Anthea Blunden for assistance in the data collection.

References

Alvarez G A, Franconeri S L, 2007 "How many objects can you track? Evidence for a resource-limited attentive tracking mechanism" *Journal of Vision* **7**(13):14, 1–10

Alvarez G A, Scholl B J, 2005 "How does attention select and track spatially extended objects? New effects of attentional concentration and amplification" *Journal of Experimental Psychology: General* **134** 461–476

Behrmann M, Zemel R S, Mozer M C, 1998 "Object-based attention and occlusion: evidence from normal participants and a computational model" *Journal of Experimental Psychology: Human Perception and Performance* **24** 1011–1036

Brainard D H, 1997 "The Psychophysics Toolbox" *Spatial Vision* **10** 433–436

Breitmeyer B, Öğmen H, 2006 *Visual Masking: Time Slices through Conscious and Unconscious Vision* (Oxford: Oxford University Press)

Cavanagh P, 1992 "Attention-based motion perception" *Science* **257** 1563–1565

Cavanagh P, Alvarez G A, 2005 "Tracking multiple targets with multifocal attention" *Trends in Cognitive Sciences* **9** 349–354

Cheries E W, Mitroff S R, Wynn K, Scholl B J, 2008 "Cohesion as a constraint on object persistence in infancy" *Developmental Science* **11** 427–432

Cousineau D, 2005 "Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method" *Tutorial in Quantitative Methods for Psychology* **1** 42–45

Culham J C, Brandt S A, Cavanagh P, Kanwisher N G, Dale A M, Tootell R B, 1998 "Cortical fMRI activation produced by attentive tracking of moving targets" *Journal of Neurophysiology* **80** 2657–2670

Culham J C, Cavanagh P, Kanwisher N G, 2001 "Attention response functions: characterizing brain areas using fMRI activation during parametric variations of attentional load" *Neuron* **32** 737–745

De Valois R L, De Valois K K, 1991 "Vernier acuity with stationary moving Gabors" *Vision Research* **31** 1619–1626

Egly R, Driver J, Rafal R D, 1994 "Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects" *Journal of Experimental Psychology: General* **123** 161–177

Erlikman G, Keane B P, Mettler E, Horowitz T S, Kellman P J, in press, "Automatic feature-based grouping during multiple object tracking" *Journal of Experimental Psychology: Human Perception and Performance*

Feldman J, 2007 "Formation of visual 'objects' in the early computation of spatial relations" *Perception & Psychophysics* **69** 816–827

Fencsik D E, Klieger S B, Horowitz T S, 2007 "The role of location and motion information in the tracking and recovery of moving objects" *Perception & Psychophysics* **69** 567–577

Feria C S, 2008 "The distribution of attention within objects in multiple-objects scenes: Prioritization by spatial probabilities" *Perception & Psychophysics* **70** 1185–1196

Feria C S, 2013 "Speed has an effect on multiple-object tracking independently of the number of close encounters between targets and distractors" *Attention, Perception, & Psychophysics* **75** 53–67

Fleet D J, Wagner H, Heeger D J, 1996 "Neural encoding of binocular disparity: energy models, position shifts and phase shifts" *Vision Research* **36** 1839–1857

Franconeri S L, Alvarez G A, Cavanagh P, 2013 "Flexible cognitive resources: Competitive content maps for attention and memory" *Trends in Cognitive Sciences* **17** 134–141
Franconeri S L, Simons D J, 2003 "Moving and looming stimuli capture attention" *Perception & Psychophysics* **65** 999–1010
Franconeri S, Sumeeth J, Scimeca J, 2010 "Tracking multiple objects is limited only by object spacing, not speed, time or capacity" *Psychological Science* **21** 920–925
Gibson J J, Kaplan G A, Reynolds J H, Wheeler K, 1969 "The change from visible to invisible: A study of optical transitions" *Perception & Psychophysics* **5** 113–116
Hespos S J, vanMarle K, 2012 "Physics for infants: characterizing the origins of knowledge about objects substances, and number" *Wiley Interdisciplinary Reviews: Cognitive Science* **3** 19–27
Hildreth E, Koch C, 1987 "The analysis of visual motion: From computational theory to neuronal mechanisms" *Annual Review of Neuroscience* **10** 477–533
Holcombe A O, Chen W-Y, 2013 "Splitting attention reduces temporal resolution from 7 Hz for tracking one object to <3 Hz when tracking three" *Journal of Vision* **13**(1):12, 1–19
Horn B K P, Schunk B G, 1981 "Determining optical flow" *Artificial Intelligence* **17** 185–203
Horowitz T S, Birnkrant R S, Fencsik D E, Tran L, Wolfe J M, 2006 "How do we track invisible objects?" *Psychonomic Bulletin & Review* **13** 516–523
Horowitz T S, Cohen M A, 2010 "Direction information in multiple object tracking is limited by a graded resource" *Attention, Perception, & Psychophysics* **72** 1765–1775
Howe P D, Holcombe A O, 2012 "Motion information is sometimes used as an aid to the visual tracking of objects" *Journal of Vision* **12**(13):10, 1–10
Howe P D, Horowitz T S, Morocz I A, Wolfe J M, Livingstone M S, 2009 "Using fMRI to distinguish components of the multiple object tracking task" *Journal of Vision* **9**(4):10, 1–11
Howe P D, Incledon N C, Little D R, 2012 "Can attention be confined to just part of a moving object? Revisiting target-distractor merging in multiple object tracking" *PLoS ONE* **7**(7) e41491
Huntley-Fenner G, Carey S, Solimando A, 2002 "Objects are individuals but stuff doesn't count: perceived rigidity and cohesiveness influence infants' representations of small groups of discrete entities" *Cognition* **85** 203–221
Intriligator J, Cavanagh P, 2001 "The spatial resolution of visual attention" *Cognitive Psychology* **43** 171–216
Jovicich J, Peters R J, Koch C, Braun J, Chang L, Ernst T, 2001 "Brain areas specific for attentional load in a motion-tracking task" *Journal of Cognitive Neuroscience* **13** 1048–1058
Kazanovich Y, Borisyuk R, 2006 "An oscillatory neural model of multiple object tracking" *Neural Computation* **18** 1413–1440
Keane B P, Mettler E, Tsoi V, Kellman P J, 2011 "Attentional signatures of perception: multiple object tracking reveals the automaticity of contour interpolation" *Journal of Experimental Psychology: Human Perception and Performance* **37** 685–698
Kunar M A, Carter R, Cohen M, Horowitz T S, 2008 "Telephone conversation impairs sustained visual attention via a central bottleneck" *Psychonomic Bulletin & Review* **15** 1135–1140
LaBerge D, 1983 "Spatial extent of attention to letters and words" *Journal of Experimental Psychology: Human Perception and Performance* **9** 371–379
Lennie P, 2003 "The cost of cortical computation" *Current Biology* **13** 493–497
Lorenceau J, Shiffrar M, Wells N, Castet E, 1993 "Different motion sensitive units are involved in recovering the direction of moving lines" *Vision Research* **33** 1207–1217
Marino A C, Scholl B J, 2005 "The role of closure in defining the 'objects' of object-based attention" *Perception & Psychophysics* **67** 1140–1149
Marr D 1982 *Vision* (New York: W H Freeman & Co)
Marr D, Ullman S, 1981 "Directional selectivity and its use in early visual processing" *Proceedings of the Royal Society of London, Series B* **211** 151–180
Mitroff S R, Scholl B J, Wynn K, 2004 "Divide and conquer: How object files adapt when a persisting object splits into two" *Psychological Science* **15** 420–425
Moore C M, Yantis S, Vaughan B, 1998 "Object-based visual selection: Evidence from perceptual completion" *Psychological Science* **9** 104–110
Morey R D, 2008 "Confidence intervals from normalized data: A correction to Cousineau (2005)" *Tutorial in Quantitative Methods for Psychology* **4** 61–64
Nishida S, Johnston A, 1999 "Influence of motion signals on the perceived position of spatial pattern" *Nature* **397** 610–612

Oksama L, Hyona J, 2008 "Dynamic binding of identity and location information: a serial model of multiple identity tracking" *Cognitive Psychology* **56** 237–283
Palmer S, Rock I, 1994 "Rethinking perceptual organization: The role of uniform connectedness" *Psychonomic Bulletin & Review* **1** 29–55
Pelli D G, 1997 "The VideoToolbox software for visual psychophysics: transforming numbers into movies" *Spatial Vision* **10** 437–442
Pelli D G, Tillman K A, 2008 "The uncrowded window of object recognition" *Nature Neuroscience* **11** 1129–1135
Posner M I, 1980 "Orienting of attention" *Quarterly Journal of Experimental Psychology* **32** 3–25
Pylyshyn Z W, Storm R W, 1988 "Tracking multiple independent targets: evidence for a parallel tracking mechanism" *Spatial Vision* **3** 179–197
Reuter-Lorenz P A, Drain M, Hardy-Morais C, 1996 "Object-centered attentional biases in the intact brain" *Journal of Cognitive Neuroscience* **8** 540–550
Scholl B J, 2001 "Objects and attention: the state of the art" *Cognition* **80** 1–46
Scholl B J, 2009 "What have we learned about attention from multiple object tracking (and vice versa)?", in *Computation, Cognition, and Pylyshyn* Eds D Dedrick, L Trick (Cambridge, MA: MIT Press) pp 49–78
Scholl B J, Pylyshyn Z W, Feldman J, 2001 "What is a visual object: Evidence from target-merging in multiple-object tracking" *Cognition* **80** 159–177
Shooner C, Tripathy S P, Bedell H E, Öğmen H, 2010 "High-capacity, transient retention of direction-of-motion information for multiple moving objects" *Journal of Vision* **10**(6):8, 1–20
Simon T J, Hespos S J, Rochat P, 1995 "Do infants understand simple arthimetic? A replication of Wynn (1992)" *Cognitive Development* **10** 253–269
Snowden R, 1998 "Shifts in perceived position following adaptation to visual motion" *Current Biology* **8** 1343–1345
St. Clair R, Huff M, Seiffert A E, 2010 "Conflicting motion information impairs mulitiple object tracking" *Journal of Vision* **10**(4):18, 1–13
St. Clair R, Seiffert A E, 2011 "Misrepresentation of motion direction causes prediction errors in multiple object tracking" *Journal of Vision* **11**(11) 291 (abstract), doi:10.1167/11.11.291
Suganuma M, Yokosawa K, 2006 "Grouping and trajectory storage in multiple object tracking: Impairments due to common item motions" *Perception* **35** 483–495
Tipper S P, Behrmann M, 1996 "Object-centered not scene-based visual neglect" *Journal of Experimental Psychology: Human Perception and Performance* **22** 1261–1278
Treisman A M, Gelade G, 1980 "A feature-integration theory of attention" *Cognitive Psychology* **12** 97–136
vanMarle K, Scholl B J, 2003 "Attentive tracking of objects versus substances" *Psychological Science* **14** 498–504
vanMarle K, Wynn K, 2011 "Tracking and quantifying objects and non-cohesive substances" *Developmental Science* **14** 502–515
von Muhlenen A, Lleras A, 2007 "No-onset looming motion guides spatial attention" *Journal of Experimental Psychology: Human Perception and Performance* **33** 1297–1310
Vul E, Frank M C, Tenenbaum J B, Alvarez G, 2010 "Explaining human multiple object tracking as resource-constrained approximate inference in a dynamic probabilistic model" *Advances in Neural Information Processing Systems* **22** 1–9
Whitney D, 2002 "The influence of visual motion on perceived position" *Trends in Cognitive Sciences* **6** 211–216
Wynn K, 1992 "Addition and subtraction by human infants" *Nature* **358** 749–750
Yantis S, 1992 "Multielement visual tracking: attention and perceptual organization" *Cognitive Psychology* **24** 295–340
Zhou K, Luo H, Zhou T, Chen L, 2010 "Topological change disturbs object continuity in attentive tracking" *Proceedings of the National Academy of Sciences of the USA* **107** 21920–21924