



Published in final edited form as:

Vis cogn. 2009 ; 17(1-2): 180–194. doi:10.1080/13506280802211334.

Feature binding in attentive tracking of distinct objects

Tal Makovski¹ and Yuhong V. Jiang^{1,2}

¹*Department of Psychology, University of Minnesota*

²*Center for Cognitive Sciences, University of Minnesota*

Abstract

To what degree can attentive tracking of objects' motion benefit from increased distinctiveness in the objects' surface features? To address this question, we asked observers to track 4 moving digits among a total of 8 moving digits. By varying the distinctiveness of the digits' color and identity, we found that tracking performance improved when the 8 objects were all distinct in color, digit identity, or both, compared to when the 8 objects were identical. However, when the 8 objects were distinct in a combination of color and digit but targets and nontargets shared color or digit identity, performance enhancement was not observed. Four follow-up experiments extended the range of the feature dimensions generating the effect and ruled out alternative strategic accounts. We conclude that surface features can be used to enhance tracking performance. This enhancement is feature-based, revealing a limited degree of feature binding in attentive tracking.

Keywords

Multiple-object tracking; attentive tracking; feature and location binding

Introduction

The dynamics of the visual environment pose a significant challenge to human vision. How do we maintain a sense of temporal continuity in a constantly changing environment? One solution to this computational problem is to track objects as they move. Motion tracking is easily achieved when a single object must be tracked, or when the tracked objects differ from the other objects in a salient visual feature. Tracking becomes increasingly difficult when multiple objects are to be tracked among visually similar nontargets (Pylyshyn & Storm, 1988). In this study we investigate the degree to which attentive tracking can benefit from distinctiveness in object features. We asked observers to track a subset of moving chromatic digits and measured their performance as a function of the objects' distinctiveness in color, digit identity, and the combination of color and digit identity. The results are informative of feature binding in attentive tracking. In the following introduction, we review the current state of knowledge on this topic and identify a critical gap.

Standard multiple-object tracking tasks have deprived participants of the opportunity to use identity information. In Pylyshyn and Storm's (1988) original paradigm, observers are shown a field of identical objects. A subset of these objects are designated as tracking targets during a cue period, after which the targets and nontargets are identical in shape and surface features. Tracking is accomplished by processing where the targets are and by updating each target's location as it moves. Object identity plays no role in this paradigm.

Although it seems that the standard multiple-object tracking task is unrepresentative of everyday visual activity where moving objects are usually distinct, there are several reasons why identity information has rarely been incorporated (with the exception of a few studies reviewed below). First, spatiotemporal properties are used predominantly when visual continuity must be established. Infants rely primarily on spatiotemporal continuity, rather than shape or color constancy, to track objects after brief occlusions. At the age of ten-month old, infants were not surprised when a toy truck moved behind an occluder and reemerged as a toy duck (Xu & Carey, 1996). Infants also did not treat two objects that differ in surface properties as separate objects until the objects started to move independently (Spelke, Gutheil, & Van de Walle, 1995). The heavy reliance on spatiotemporal properties was also observed in visually deprived individuals who recovered vision as adults. These individuals also initially relied heavily on motion continuity to segregate a visual scene and only later incorporated surface properties such as color or shape discontinuity (Sinha, 2007).

Normal human adults also appear to prefer spatiotemporal information to identity information in updating visual objects. For example, in the object reviewing paradigm (Kahneman, Treisman, & Gibbs, 1992), two letters (e.g., A and B) were presented inside two squares positioned at the 3 and 9 o'clock locations. The letters then disappeared while the squares moved clockwise to the 6 and 12 o'clock locations. After the motion, a probe letter was displayed for participants to name as quickly as possible. Participants were faster to name a formerly presented letter if it appeared in its original square rather than in the other square. Kahneman et al. (1992) proposed that spatiotemporal continuity served to maintain the integrity of "object files", allowing participants to update object information after motion. Interestingly, object updating appears to rely primarily on spatiotemporal continuity. In a similar paradigm, if the two letters initially appeared against blue and green background and later the background switched colors, performance in reporting the previewed letters was unaffected (Mitroff & Alvarez, 2007).

Another reason why researchers have focused primarily on tracking "where" rather than "what" is because the use of identity information in multiple-object tracking is highly limited. When objects with different shapes were used and tracking was occasionally interrupted by a question about the characteristics of a target, observers accurately reported the target's direction and speed, but had difficulty reporting its shape (Scholl, Pylyshyn, & Franconeri, 1999). Further evidence that "what" is poorly bound to "where" comes from a study by Pylyshyn, who presented different digits, one in each circle in the cue period for participants to remember (Pylyshyn, 2004). The digits were taken away before the circles began to move. At the end of the motion period, participants were asked to select the target circles and to report the digit originally presented in each circle. This paradigm bears some similarity to Kahneman et al.'s object updating task, except that there were more objects (8 versus 2 moving objects), more identities (4 vs. 2), more complex motion, and an explicit instruction to bind what with where. Although participants in Kahneman et al.'s paradigm were able to update the identity of two objects after a brief rotation of the placeholders, those in Pylyshyn's more complex task were unable to do so. Even when they successfully tracked the target circles, memory for the associated digit identities was poor. Similar failures of "what-to-where" binding have been observed when participants tracked objects with apparent, rather than continuous, motion (Saiki, 2003).

Although the binding of what to where is difficult in attentive tracking, some degree of binding has been demonstrated when distinct objects are used for tracking, and when participants are explicitly required to bind what to where. The multiple-identity tracking task usually presents observers with a small number of objects and designate either a subset or the entire set for tracking. Unlike Pylyshyn's (2004) task where the distinct identities are presented only during the cue period, in the multiple-identity tracking task objects are always distinct in shape or

color for the entire tracking period. At the end of the motion period, all objects are masked. One of the locations is then probed for observers to retrieve the associated object identity (Oksama & Hyönä, 2004, in press). Observers' limited ability to bind identities to locations was dependent upon the speed and number of targets.

Horowitz and colleagues employed a further modification of the multiple-identity tracking task, investigating whether distinctive identities are addressable during tracking (Horowitz et al., 2007). Participants were asked to track cartoon animals (e.g., a tiger, a leopard) that moved in straight lines. The unique identities of the animals remained on the screen during tracking. At the end of the tracking period the animals hid behind cactuses. Observers were asked to report either the location of all the targets (*standard* question) or the location of a specific animal (*specific* question). Horowitz et al. found that the estimated capacity was higher when answering the *standard* question than when answering the *specific* question, with a capacity estimate of about one object in the *specific* question. From these data Horowitz et al. concluded that "what" information is accessible during tracking, but only to a limited degree. Consistent with this idea, tracking performance was higher when the moving animals were distinct rather than identical to one another (Experiment 5, Horowitz et al., 2007).

The main goal of the present study is to investigate the extent to which attentive tracking can be enhanced by distinctiveness in object identities. Horowitz et al. showed that tracking performance is higher when distinct, rather than homogeneous, objects must be tracked. However, their use of cartoon animals makes it difficult to dissociate the role of visual distinctiveness from semantic or name distinctiveness. In addition, it is difficult to isolate the visual features that contribute to tracking of distinct objects observed by Horowitz et al. (2007). To understand how object features are integrated with their location during tracking, it is necessary to test tracking of visual objects made of identifiable features.

We conducted a series of five experiments using colored digits (or colored shape/orientation) to study how distinctiveness in features or conjunction of features affects attentive tracking. In this article, we report one of the experiments in full and briefly describe the other experiments that provide converging support for the main findings. In the main experiment, we used an array of colored digits and asked participants to track a subset of those objects. Unlike the multiple-identity tracking task, participants of the present study were not required to remember object identities. Tracking can be achieved entirely by updating the motion trajectories of objects without remembering their identities. Importantly, we manipulated the objects' distinctiveness. The objects could be distinct in one or two features (e.g., all have different colors) or in a combination of features (e.g., the conjunction of color and shape made all objects unique). We also tested whether all objects or only the target subset needs to be distinct to facilitate tracking.

Consider two conditions shown in Table 1: the *homogeneous* condition and the *color-distinct* condition, where participants tracked 4 out of 8 objects that were either homogeneous or distinct in color. Motion trajectory tracking is sufficient for both conditions. If spatiotemporal continuity is the only source of information for attentive tracking, then performance should be comparable between the *homogeneous* and *color-distinct* conditions. In contrast, if distinctiveness in color is also used to separate targets from nontargets, then performance may be superior in the *color-distinct* condition. Similar predictions can be made when comparing the *homogeneous* condition with the *digit-distinct* condition. Furthermore, the distinctiveness in both color and digit may convey an added advantage in the *both-distinct* condition. Thus, the comparison of the *homogeneous* condition with the *color-distinct*, *digit-distinct*, and *both-distinct* conditions informs us whether featural information is helpful for tracking when the task does not require identity memory. An advantage in the feature distinct conditions may imply that participants can bind a feature with a moving location.

Now consider the *conjunction-distinct* condition, where the 8 objects were made distinctive because they were distinct in the combination of color and digit shape. For instance, a target object, such as a blue 3, is the only object that is blue in color and is “3” in shape. In terms of features, it shares with a distractor (e.g., a blue 4) in color and with another distractor (e.g., a green 3) in digit identity. If the advantage conveyed by tracking distinct objects over tracking homogeneous objects operates upon a color-digit integrated object, then it should also enhance performance in the *conjunction-distinct* condition. In contrast, if the advantage for tracking distinct objects emerges primarily from distinguishing features that are not fully integrated together, then the *conjunction-distinct* condition may not yield superior performance to the *homogeneous* condition.

Whether distinctiveness in the conjunction of two features helps attentive tracking is an interesting question not addressed previously. When objects are static, binding of separate features to form an integrated object is easily achieved. Indeed, extensive research has shown that visual attention and visual working memory operate on the basis of integrated objects (Duncan, 1984; Scholl, 2001). For example, participants can remember all four colors and all four orientations as well as remembering only four colors or only four orientations, as long as the colors and orientations conjoin to form four colored oriented lines (Luck & Vogel, 1997). Feature binding may become much more difficult in attentive tracking, however, where objects are moving and where a subset of objects is not actively attended (Pylyshyn, 1989; but see Cavanagh & Alvarez, 2005). In turn, the *conjunction-distinct* condition may produce performance similar to one of the paired conditions, where targets and distractors share featural information (see Table 1, *color-paired*, *digit-paired*, and *conjunction-paired*). Sample trials can be viewed online at <http://jianglab.psych.umn.edu/MOTdigits/MOTdigits.htm>.

In short, by comparing performance in the different identity-distinct conditions with the *homogeneous* condition, this study informs us how feature binding operates in attentive tracking.

Method

Participants

All participants were volunteers from the University of Minnesota between 18 and 33 years old. They all had normal color vision and normal or corrected-to-normal visual acuity. 20 participants took part in the main experiment.

Equipment

Participants were tested individually in a room with normal interior lighting. They sat approximately 57cm away from a 19” computer monitor. The experiment was programmed with psychophysical toolbox (Brainard, 1997; Pelli, 1997) implemented in MATLAB (www.mathworks.com).

Stimuli

The moving objects were colored digits (each fit within a $1.5^{\circ} \times 1.5^{\circ}$ imaginary square) presented against a black background. There were eight possible colors (red, green, blue, yellow, orange, azure, white, and pink) and eight possible digits (1 to 8). The colors and digits were randomly sampled according to a trial’s condition.

Procedure

Participants were tested in 8 conditions with 15 trials each, presented in a randomly intermixed order. All conditions used similar trial sequences. Participants pressed the spacebar to initiate each trial. Eight colored digits were presented at randomly selected locations within an

imaginary presentation window of $21^\circ \times 21^\circ$. Four of the objects were surrounded by white frames ($1.9^\circ \times 1.9^\circ$), which identified their status as tracking targets. After 1330msec, the frames disappeared and the objects started moving randomly within the presentation window. The objects moved at a constant speed of $21.9^\circ/\text{s}$, bounced off the “wall” of the presentation window, and repelled one another when a minimal center-to-center distance of 2.1° was reached. Participants were asked to track the cued targets and were encouraged to maintain central fixation during tracking. After the objects moved for 7 seconds and stopped, they all turned to white disks (1.5° in diameter). Participants were asked to click on the four targets. The correctly selected targets turned green while the missed targets turned red to provide feedback for 1 second.

Design

The eight conditions differed in the distinctness of identity (Table 1). All eight objects were identical in the *homogenous condition* (e.g., they were all red 2s); the exact color and digit identity was randomly selected on each trial of that condition.

There were four “distinct” conditions. The eight objects were uniform in digit identity but distinct in color in the *color-distinct condition* (e.g., they were all 4s, but in 8 different colors). The objects were the same in color but distinct in digit identity in the *digit-distinct condition* (e.g., they were all red, but in 8 different digit identities). Or the objects were distinct in both color and digit identity in the *both-distinct condition* (e.g., all 8 digits and all 8 colors were used in random combination). The *conjunction-distinct condition* involved distinct objects produced by the conjunction of four colors and four digits. In this condition, no two objects were the same in the combination of color and digit, but a given target (e.g., a blue 3) always shared color with one of the distractors (e.g., a blue 4) and digit identity with another distractor (e.g., a green 3).

Finally, there were three “paired” conditions in which all four targets were distinct and all four distractors were distinct, but each of the targets was the same as one of the distractors. In the *color-paired condition*, for instance, all objects were the same in digit identity, but the four targets were distinct in color, and the four distractors were distinct in color, and a target always shared a distractor’s color. Similarly, in the *digit-paired condition*, the targets and distractors were paired in digit identity but each subset was distinct. In the *conjunction-paired condition*, the targets and distractors were paired in both digit identity and color, but each subset was distinct. Note that the *conjunction-paired condition* included 4 digits and 4 colors which were paired to make one target identical to one distractor. The *conjunction-distinct condition* also included 4 digits and 4 colors, however the features were combined such that all 8 objects were distinct in color-digit combination.

Results

Figure 1 depicts mean tracking accuracy as a function of stimulus condition in the main experiment.

(1) *Homogeneous versus distinct identities*

Does distinctiveness in feature identities enhance tracking? We addressed this question by comparing performance in the *homogeneous*, *color-distinct*, *digit-distinct*, and *both-distinct* conditions using a repeated-measures ANOVA. Distinctiveness in color enhanced performance, $F(1, 19) = 14.60$, $p < .01$, as did distinctiveness in digit identity, $F(1, 19) = 6.70$, $p < .05$. The interaction between color distinctiveness and digit distinctiveness was marginally significant, $F(1, 19) = 3.20$, $p = .09$, as the two factors were slightly under-additive. Pairwise t -tests showed that each of the three feature-distinct conditions was significantly higher in

accuracy than the *homogeneous* condition, all $ps < .001$. These results clearly showed that attentive tracking was influenced by objects' distinctive identities, even when participants were not required to remember what the objects were (see also Horowitz et al., 2007).

Did the enhancement from distinct identities extend to a condition where the objects were distinct in terms of a combination of color and digit, but not in color or digit alone? Our results clearly showed that it did not. To the contrary, the *conjunction-distinct* condition had significantly lower accuracy than the *homogeneous* condition, $t(19) = 3.30, p < .004$. The accuracy of the *conjunction-distinct* condition was also significantly lower than any of the three feature-distinct conditions, all $ps < .001$.

(2) *Homogeneous versus paired identities*

The three paired conditions were designed such that objects had distinct features within each subset, but one target was always identical to one distractor. In these cases, tracking on the basis of distinct identity (e.g., a red 7) can only help distinguish between different targets, or between each specific target and three of the distractors. It does not help distinguish a target and its paired distractor. Our results showed that accuracy in the *paired* condition was lower than the accuracy in the *homogeneous* condition. The impairment was significant in the *digit-paired* condition, $t(19) = 4.10, p < .001$, and the *conjunction-paired* condition, $t(19) = 5.40, p < .001$, but was insignificant in the *color-paired* condition, $t(19) = 1.60, p > .10$. These results suggest that the advantage afforded by distinguishing the target from other targets or from three of the distractors was outweighed by the disadvantage of confusing the target with an identical distractor.

(3) *Conjunction-distinct vs. conjunction-paired*

Both the *conjunction-distinct* and the *conjunction-paired* conditions were produced by conjoining four colors with four digits to form eight objects. The only difference was that the combination of digits and colors produced eight distinct objects in the former but four pairs of target-distractor in the latter. Does distinctiveness in feature conjunction produce any advantage in the former? No, the *conjunction-distinct* condition was not significantly better than the *conjunction-paired* condition, $t(19) = 1.30, p > .20$.

Follow-up Experiments

In addition to the main experiment, we conducted four control experiments to verify the generality of the results. The control experiments were similar in procedure to the main experiment, so we included a brief description of the methods and results. Descriptive statistics and p-values for these experiments are provided in Table 2. Our main focus was to show that objects distinct in feature conjunction (but not in single features) are not tracked better than homogeneous objects.

Control Experiment 1: variable trial duration

The unpredictable trial duration was used to discourage participants from using a strategy of not tracking until the final moments of a trial [footnote 1]¹. Seven participants were tested in this experiment which was a replication of the main experiment, except that the duration of tracking was unpredictable. Each trial ended after 3.0, 3.8, 4.9, 6.3, or 8.0s of motion. Similar to the main experiment, each distinctiveness condition contained 15 trials, divided randomly and evenly into the 5 possible trial durations. The basic pattern of results (Table 2) was the same as the main experiment: distinctness in features enhanced performance but distinctness in a combination of features did not. Compared with the *homogeneous* condition, accuracy was

¹We thank Todd Horowitz for suggesting this possibility.

enhanced in the *color-distinct* and *digit-distinct* conditions, $ps < .03$, but impaired in the *conjunction-distinct* condition, $p < .02$.

Control Experiment 2: Blocked design

This experiment was designed to investigate whether effects of feature and conjunction distinctiveness held with a blocked (rather than mixed) design. Eleven participants were tested using colored digits as in the main experiment. They completed five conditions presented in separate blocks of trials; the order of blocks was randomized. The conditions were homogeneous, color-distinct, digit-distinct, both-distinct, and *conjunction-distinct*. There were 15 trials in each condition. Blocking the distinctive conditions did not produce an advantage in the *conjunction-distinct* condition compared with the *homogeneous* condition ($p > .40$), which was significantly lower in accuracy than the *feature-distinct* conditions ($ps < .05$).

Control Experiment 3: color and novel shapes

Thirteen participants were tested in this experiment where digits were replaced by novel shapes that were difficult to name. Otherwise the experiment was identical to the main experiment. Again, the results were qualitatively similar to those of the main experiment. Distinctiveness in a combination of features failed to facilitate performance compared with the homogeneous condition, $p > .55$.

Control Experiment 4: color and line orientations

Line orientation is a simple visual feature that may be better integrated with color than digits or novel shapes [footnote 2]². Ten participants were tested in this experiment where 8 oriented lines (20°-160° in units of 20°) replaced the digits. Still, the use of two simple features (color and line orientation) did not change the results. The *conjunction-distinct* condition again did not yield higher accuracy than the *homogeneous* condition, $p > .43$.

Discussion

Everyday visual experience suggests that we track moving objects not only on the basis of their motion trajectories but also on the basis of their surface features. Our study confirms this intuition. Attentive tracking of a subset of moving objects was more accurate when all objects were distinct in color or digit identity than if they were identical in color and digit identity. The enhancement by distinctive features was confined to conditions where at least a single feature differentiated all objects, as when all objects had different colors, but not when a combination of features differentiated all objects. In addition, even when the targets were distinct from one another, tracking was not enhanced if the distractors shared the targets' colors or digit identities. Four follow-up experiments demonstrated the robustness of these effects by ruling out strategic biases (Control experiments 1-2) and by extending the range of the dimensions generating the effect (Control experiments 3-4). Together these findings have significant implications for feature binding in attentive tracking.

First, the finding that distinctiveness in features enhances tracking apparently contradicts previous results, where participants were found to have poor memory for surface properties in a multiple-object tracking task (Scholl et al., 1999). Indeed, trying to remember object identities may actively interfere with the ability to track (Fougnie & Marois, 2006). The discrepancy can be accounted for by an important difference between our study and the previous studies. Namely, we did not require participants to remember object identities. If we had stopped our participants occasionally to probe their memory of object identities, we might also have

²We thank Todd Horowitz for this suggestion. Digits, rather than orientations, were used in the main experiment because the 8 digits were more distinctive than the 8 orientations.

observed poor performance. Nonetheless, distinctiveness in features may help segregate targets from nontargets, a critical component of attentive tracking (Yantis, 1992).

We also found that performance was impaired in the *paired*-identity conditions compared with the *homogeneous* condition, a finding that was observed in one but not another of Horowitz et al.'s (2007) experiments. The effect was not always statistically reliable (Table 2), but it had been observed several times (see also Makovski & Jiang, submitted). The impairment was somewhat surprising because the *paired-identity* conditions were more distinctive than the *homogeneous* condition, as each object was distinctive in feature from all but one other object. One might expect that distinctiveness would enhance, rather than impair, performance in the *paired-identity* condition. In a subsequent study, we found that the impairment for *paired* conditions was observed only when the *paired* condition was randomly intermixed in presentation with the *distinct* and *homogeneous* conditions (Makovski & Jiang, submitted). Testing these conditions in separate blocks of trials removed impairment in the *paired* condition. Our current conclusion is that the use of identity information in multiple-object tracking is not obligatory, but is under strategic control. A tendency to rely on identity was also applied to the *paired* condition when trials were randomly intermixed, because identity information helped performance on *distinct* trials. This strategy sets up a competition between two types of perceptual grouping: Grouping on the basis of attentional set (tracking targets versus nontargets), and grouping on the basis of surface features (red objects vs. green objects and so on). The competition from feature grouping impaired the effectiveness of attentional grouping. Competition for the two kinds of grouping cannot explain the lack of a benefit in the *conjunction-distinct* condition, however, as blocking conditions did not facilitate its performance compared with the *homogeneous* condition.

The results of the current study clearly demonstrate that feature rather than conjunction distinctiveness enhances tracking performance. Does this benefit reflect improved tracking abilities, or that distinctiveness provides additional information for recovering lost targets (Horowitz et al., 2007)? On the one hand, tracking ability may be directly improved because observers can tolerate closer distance between a target and a distractor when they have distinct identities. On the other hand, identity information may be held in a separate system, such as visual working memory, which allows participants to recover a lost target. To disassociate these possibilities, in a recent study we asked participants to track distinct objects whose color changed once every 260msec (Makovski & Jiang, submitted). The objects were always distinct from one another at any given moment, but the continuity of identity information over time was destroyed. In this setup, participants could no longer track the distinct condition better than the homogeneous condition. These data are more consistent with the second hypothesis, which asserts that identity information is kept in a system (e.g., visual working memory) that operates in parallel to motion tracking.

Even though features held in visual working memory are typically considered to be bound together to form coherent objects (Luck & Vogel, 1997), our results suggest that they are not properly conjoined during attentive tracking. Distinctiveness in feature-conjunction would have produced as much advantage as distinctiveness in features had they have been properly conjoined. Apparently, motion made such feature binding more difficult to retain, probably because the basis for binding - shared location (Treisman, 1988) - is no longer a perceptually stable property.

In summary, we have shown that multiple-object tracking is sensitive to the distinctness in object identities, even in a task where identity memory is never probed. The enhancement is largely feature-based, revealing a limited degree of feature binding in attentive tracking.

Acknowledgements

This research was supported in part by NSF 0733764 and NIH 071788. We thank Leah Watson, Khená Swallow, Carrick Williams, Todd Horowitz and James Brockmole for comments and suggestions.

References

- Brainard DH. The psychophysics toolbox. *Spatial Vision* 1997;10(4):433–436. [PubMed: 9176952]
- Duncan J. Selective Attention and the Organization of Visual Information. *Journal of Experimental Psychology-General* 1984;113(4):501–517. [PubMed: 6240521]
- Fougnie D, Marois R. Distinct capacity limits for attention and working memory - Evidence from attentive tracking and visual working memory paradigms. *Psychological Science* 2006;17(6):526–534. [PubMed: 16771804]
- Horowitz TS, Klieger SB, Fencsik DE, Yang KK, Alvarez GA, Wolfe JM. Tracking unique objects. *Perception & Psychophysics* 2007;69(2):172–184. [PubMed: 17557588]
- Kahneman D, Treisman A, Gibbs BJ. The Reviewing of Object Files - Object-Specific Integration of Information. *Cognitive Psychology* 1992;24(2):175–219. [PubMed: 1582172]
- Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. *Nature* 1997;390(6657):279–281. [PubMed: 9384378]
- Makovski T, Jiang YV. The role of visual working memory in attentive tracking of unique objects. submitted
- Mitroff SR, Alvarez GA. Space and time, not surface features, guide object persistence. *Psychonomic Bulletin & Review* 2007;14(6):1199–1204. [PubMed: 18229497]
- Oksama L, Hyönä J. Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition* 2004;11:631–671.
- Oksama L, Hyönä J. Dynamic binding of identity and location information: A serial model of multiple-identity tracking. *Cognitive Psychology*. in press
- Pelli DG. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision* 1997;10(4):437–442. [PubMed: 9176953]
- Pylyshyn ZW. Some puzzling findings in multiple object tracking: I. Tracking without keeping track of object identities. *Visual Cognition* 2004;11(7):801–822.
- Pylyshyn ZW, Storm RW. Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision* 1988;3:179–197. [PubMed: 3153671]
- Saiki J. Feature binding in object-file representations of multiple moving items. *Journal of Vision* 2003;3(1):6–21. [PubMed: 12678621]
- Scholl BJ. Objects and attention: the state of the art. *Cognition* 2001;80(12):1–46. [PubMed: 11245838]
- Scholl BJ, Pylyshyn ZW, Franconeri SL. When are featural and spatiotemporal properties encoded as a result of attentional allocation? *Investigative Ophthalmology & Visual Science* 1999;40(4):S797–S797.
- Sinha, P. From fragments to objects: Mechanisms of visual integration; Paper presented at the 2007 Summer Institute on the Visual Perception & Cognition Conference; Minneapolis, MN. 2007;
- Spelke, E.; Gutheil, G.; Van de Walle, G. The development of object perception. In: Osherson, D., editor. *Invitation to Cognitive Science*, 2nd Ed., Vol 2: *Visual Cognition*. MIT Press; Cambridge, MA: 1995.
- Treisman A. Features and Objects - the 14th Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology* 1988;40(2):201–237.
- Xu F, Carey S. Infants' metaphysics: The case of numerical identity. *Cognitive Psychology* 1996;30(2):111–153. [PubMed: 8635312]
- Yantis S. Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology* 1992;24:295–340. [PubMed: 1516359]

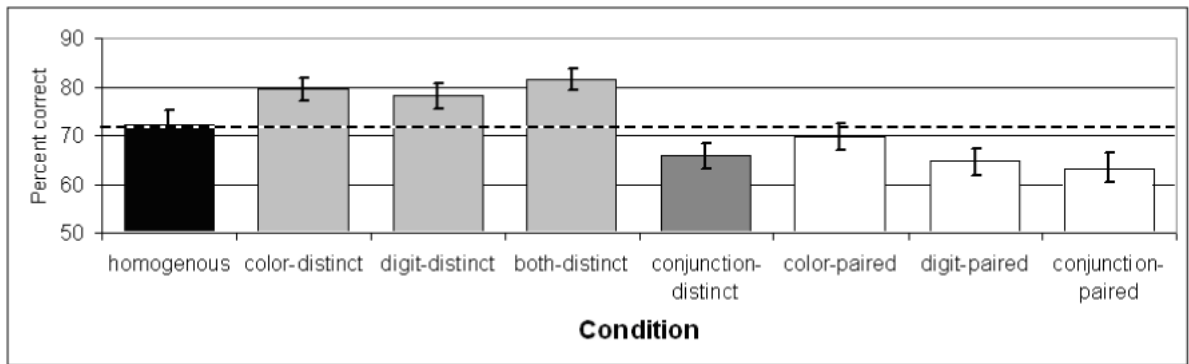


Figure 1. Tracking accuracy as a function of stimulus conditions (Error bars show ± 1 standard error of the mean). The dashed line represents performance in the homogenous condition.

Table 1
A schematic illustration of tracking objects used in Experiment 1

Targets	
1, 1, 1	1
3, 3, 3	3, 3
2, 4, 5, 7	1, 3
5, 7, 8	2, 3
2, 3, 4, 6	2
4, 4, 4	4, 4

Targets

2, 4, 7

1, 2,

3, 4, 6

2

Table 2
Mean percent correct in the four control experiments (Standard Error in parenthesis). The p-values correspond to the t-tests comparing each condition with the homogenous condition

Control Experiment	Homogeneous	Color-distinct	Shape-distinct	Both-distinct	Conjunction-distinct	Color-paired	Shape-paired	Conjunction-paired
1: Variable trial duration	73 (4)	83 (4) p=0.03	78 (5) p=0.02	79 (5) p=0.09	65 (5) p=0.02	73 (4) p=0.93	68 (5) p=0.17	70 (5) p=0.40
2: Blocked design	70 (4)	78 (4) p=0.02	79 (4) p<0.01	77 (4) p=0.04	68 (4) p=0.42	N/A		
3: Novel shapes	80 (3)	86 (3) p<0.01	82 (3) p=0.22	87(3) p<0.01	80 (3) p=0.57	76 (3) p=0.03	79 (3) p=0.50	78 (3) p=0.21
4: Line orientations	70(5)	77(4) p<0.01	69(5) p=0.62	77(5) p<0.01	71(5) p=0.43	71(6) p=0.41	73(5) p<0.01	69(6) p=0.57