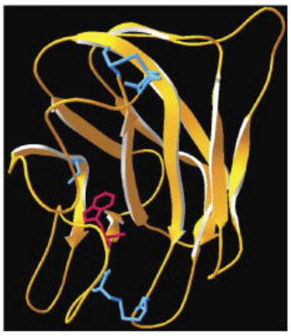
VISION RESEARCH

An International Journal for Functional Aspects of Vision



Biochemistry & Cell Biology • Molecular Biology & Genetics
Anatomy, Physiology, Pathology & Pharmacology • Optics, Accommodation & Refractive Error
Circuitry & Pathways • Psychophysics • Perception • Attention & Cognition
Computational Vision • Eye Movements & Visuomotor Control

ISSN 0042-6989 | Volume 46 | Number 22 | October 2006

Available online at www.sciencedirect.com
ScienceDirect
Vision Research 46 (2006) 3867–3880
Vision Research
www.elsevier.com/locate/visres

An inter-item similarity model unifying feature and conjunction search ☆

Steven Phillips [a,*], Yuji Takeda [b], Takatsune Kumada [b]

[a] *Neuroscience Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan*

[b] *Institute for Human Science and Biomedical Engineering, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan*

Received 11 April 2006; received in revised form 23 June 2006

Abstract

We propose a model of visual search to address the hitherto unresolved issue of reconciling serial deployment of attention accounts with inter-item similarity effects. Target–distractor and distractor–distractor similarity were systematically varied in 85 (17×5) set type-size conditions over seven experiments, including univariate feature and bivariate conjunction search. The model, a power (square root) function of dimension-specific target–distractor and distractor–distractor similarity in linear combination with set size, accounted for 98% of the variance on type-size means. It suggests that much of efficient and inefficient search can be unified under a single theory involving item similarity.

Keywords: Visual search; Visual attention; Item similarity; Parallel

1. Introduction

The visual search paradigm has been used extensively to probe the processes underlying vision and its possible interaction with attention and memory. Participants are typically given the task of finding a target among an array of display items. Search slope—response time as a function of the number of display items—is a measure of the efficiency of the search process. For example, a flat slope where search time is independent of the number of display items suggests, though not necessarily implies, efficient parallel processing of display items and a steep slope suggests that display items are processed in series. For a review of the types of features that afford efficient/inefficient search, see Wolfe and Horowitz (2004); and for its interpretation in terms of underlying processes, see Wolfe (2003).

Two major theoretical explanations for search efficiency that have been extensively tested and debated are Feature Integration Theory (Treisman & Gelade, 1980) and Attention Engagement Theory (Duncan & Humphreys, 1989). Efficient search is often observed for display sets of items varying along a single (e.g., color) dimension, where all distractors share the same feature—feature search. Inefficient search is often seen for sets of items varying along multiple (e.g., color and shape) dimensions, where some distractors share a feature with the target—conjunction search. Feature Integration Theory (FIT) says that feature dimensions are processed in parallel, but conjunctions of features are processed in series. A search along multiple dimensions means more items must be encoded as conjunctions (to distinguish targets from distractors), hence the difference in search difficulty. In contrast, Attention Engagement Theory (AET) says that search is efficient because similarity between distractors causes attention to them to be suppressed—*spreading suppression*; and search is inefficient because similarity between targets and distractors causes more items to be considered for further processing—

☆ We thank Adam Reeves and two anonymous reviewers for their thoughtful comments and suggestions on this work.

* Corresponding author. Fax: +81 29 861 5857.
E-mail address: steve@ni.aist.go.jp (S. Phillips).

0042-6989/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2006.06.016

template matching, where similarity is defined in terms of common item features (Duncan & Humphreys, 1989, 1992). Unidimensional display sets usually have greater distractor–distractor similarity than bidimensional display sets, hence the difference in search efficiency. An intermediate position between these two extremes is the Guided Search Model (Wolfe, Cave, & Franzel, 1989; Cave & Wolfe, 1990; Wolfe, 1994), where parallel processing of feature information determines the serial order in which items are matched to the target.

From an analysis of Duncan and Humphrey's claims and a further experiment, Treisman found no support for search efficiency based on the spreading suppression component of AET (Treisman, 1991, 1992). In the critical conjunction search condition over color-oriented bars, where 50% of the distractors were replaced with grey disks, FIT and AET made opposite predictions. FIT predicted a shallower slope relative to standard conjunction, because fewer distractors sharing a feature with the target means fewer items enter into a serial attention stage for matching against the target. AET predicted a steeper slope, because the increased distractor–distractor similarity between the additional grey disks should increase suppression and therefore reduce the number of items needing attention. The data supported FIT.

Although Treisman's results do not support spreading suppression, the general implications for a theory of visual search remain unclean. Quinlan (2003) noted in his review of FIT that although the debate appears to have ended with Treisman, a theory of visual search is not complete without an account of item similarity effects. The work reported here is a closer examination of the effect of feature sharing between targets and distractors on visual search. Over seven experiments, we systematically vary target–distractor and distractor–distractor similarity in a series of tasks that spans both feature (univariate dimensional) search and variations in conjunction (bivariate dimensional) search, including Treisman's "grey-disks" conditions that were modified for compatibility with the other experiments. Experiments 1 and 2 test the effects of target–distractor and distractor–distractor similarity (respectively) for each dimension. Experiment 3 tests the relative contributions of these two factors. Experiment 4 tests the relative contributions of distractor–distractor similarity on both versus either dimension. Experiment 5 mirrors Experiment 4 by testing the relative contributions of target-distractor similarity on one versus two dimensions and provides a further test for the relative effects of target–distractor versus distractor–distractor similarity. Experiment 6 tests distractor–distractor similarity for univariate display sets, and the relative influence of orientation and frequency dimensions on search time. Experiment 7 tests Treisman's "grey-disks" conditions. All conditions in Experiments 1–5, and 7 involve bivariate (conjunction) search, whereas all conditions in Experiment 6 involve univariate (feature) search. We then propose a simple model that accounts for all these

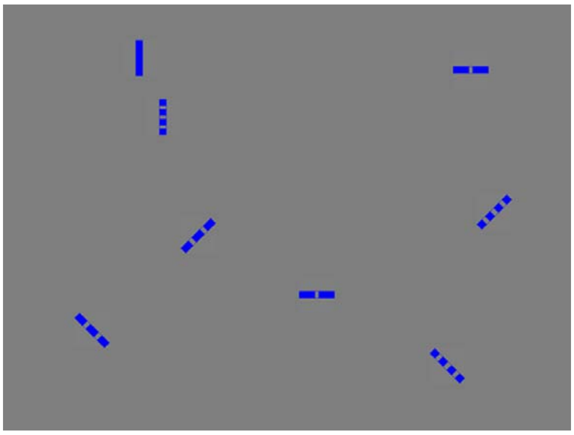


Fig. 1. Example display set. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

variations that is inspired by, but differs from the specific details of AET. We finish with a discussion of these results and implications for a more unified theory of visual search.

2. General methods

2.1. Participants

Participants were generally between 20 and 24 years old, about half were female, and all with good vision. They were paid for their participation.

2.2. Apparatus and stimuli

Stimuli were presented on a 17-in. CRT using a standard desktop computer and placed about 57 cm from the participant (i.e., 1 cm equals about 1° of the participant's field of view). The display area subtended an angle of approximately 29° horizontally and 23° vertically. Stimuli were rectangular bars, subtending 2° in length and 0.4° in width. An example display set of blue bars in various orientations and frequencies is shown in Fig. 1. Target and display bars were either all red, green, blue, or yellow within a trial. The background color for all trials was grey. The display was divided by invisible horizontal and vertical lines into four equal sized quadrants. Display stimuli including targets were uniformly distributed over the four quadrants.

2.3. Conditions

Display sets were constructed from a single target among several distractor types. Display set sizes were either 4, 8, 16, 24, or 32. Sets of more than four items were constructed by recycling through the distractor item types. Display conditions are specified each experiment section and summarized in Table 1, where 00 is the target. The one-to-one correspondence between digits and features was randomly chosen for each trial. So, for example, a horizontal bar with one gap may be a target in one trial, but a distractor for another.

2.4. Procedure

On each trial, participants were presented with a target stimulus at the center of the screen (1500 ms), followed by a delay (500 ms), followed by a display set. Targets varied across trials, and were randomly selected from the trial display set. We employed target-present trials only. Hence, upon identifying the stimulus in the display set that matched the target, participants pressed the *space bar*, which caused each stimulus in the display set

Table 1
Display set type conditions (search slopes)

UDDo (18.6)		UoBL (29.6)		UDDf (30.5)		UfBL (37.6)		DDof (13.3)		DDo (17.8)		DDf (20.5)		BL (26.8)		TD&DDf (31.4)	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	1	0	1	1	1	1	1	1	1	1	1	1	0
1	0	2	0	0	1	0	2	1	2	1	2	2	1	2	2	2	1
2	0	3	0	0	2	0	3	2	2	3	3	3	2	3	3	3	1

TDf (36.4)		TD&DDo (36.8)		TDo (39.5)		TDof (50.0)		SC[a] (55.9)		XC[b] (31.1)		TDoM[a] (34.3)		TDfM[a] (21.8)	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	1	0	1	0	1	0	1	0	1	0	1	1	0
2	1	1	2	1	2	1	0	1	0	1	0	2	2	2	2
3	2	1	3	2	3	2	2			1	1				

Note. The condition labels are coded as follows: BL, no similarity baseline; TD, target–distractor similarity; DD, distractor-distractor similarity; o, orientation dimension; f, frequency dimension; U, univariate search condition; SC, standard conjunction; XC, extended conjunction and M, matched distractor ratio. For example, UDDo is the *univariate distractor–distractor orientation similarity* condition; TDof is the *target–distractor orientation and frequency similarity* condition; and TDoM is the *target–distractor orientation similarity with matched distractor ratio* condition.

[a] Indicates a 0.5:0.5 ratio for distractors.

to be replaced by a digit from one to eight. Participants then responded by pressing the key with the digit that replaced the matched stimulus. To avoid the strategy of identifying the target by the uniqueness of its corresponding digit for set sizes greater than eight, the target and three other digits appeared only once in the digit display. Thus, some distractors were associated with more than one digit, but this arrangement does not affect target identification. After pressing a digit, or 10,000 ms (whichever came first), a fixation cross was displayed at the center of the screen to signify the end of the current trial. Response time was taken as the delay from display set presentation to *space bar* response. Participants were instructed to respond quickly and accurately to the display set. Speed was not stressed for digit responses. Participants were prompted to start each block by a key press to allow a short break between blocks if needed. There was one block per display set type condition. For Experiments 1–5, with three display set conditions, each block consisted of 50 (=10 × 5 [set size]) trials. Thus, every participant received 600 (=4 [session] × 3 [block] × 50) experiment trials. For Experiments 6 and 7, with four display set conditions, there were 40 trials per block, totalling 640 trials for each participant. The arrangement of features, locations and digits; and the order of blocks was randomized across trials, sessions, and participants. Participants were assigned one of eight session sets (i.e., each set contained one practice and four experiment sessions) so that no more than two participants took exactly the same sequence of trials.

Analyses were conducted on the response time data excluding error and outlier trials. An incorrect digit response, or failure to respond within the 10 second time limit was regarded as an error. Outliers were determined by the *modified recursive* method (Selst & Jolicoeur, 1994) on error-free trials. Between 0.9% and 1.2% of trials were outliers in each experiment. Unless stated otherwise, two-way (set type by size) repeated measures analyses of variance (ANOVA) were conducted on response times and error rates. Participant-specific linear regressions of response times were performed on each set type-size condition, and the slopes[1] were entered into a one-way repeated measures ANOVA. Errors were subjected to the arcsine transformation before conducting analyses of variance.

3. Experiment 1

A basic prediction of AET and FIT is that search slope increases with greater target–distractor similarity (shared features) in the display set. The purpose of this experiment is to calibrate the size of this effect for our stimuli.

3.1. Method

3.1.1. Participants

Sixteen participants performed the experiment.

3.1.2. Conditions

Three display set conditions were employed in the first experiment. They were the no similarity *baseline* (BL), where all display item types, both target and distractors, had unique feature types on both orientation and frequency dimensions; *target–distractor orientation similarity* (TDo), where a distractor type shared a common orientation feature type with the target type; and *target–distractor frequency similarity* (TDf), where a distractor type shared a common frequency feature type with the target type. Throughout the paper, we refer to conditions by their item

[1] The R^2 values for each participant-condition slope were greater than 0.85 in most (about 90%) of cases, indicating a linear relationship between set size and response time.

type (not instance) similarity, although subsequent references to type are omitted for brevity.

3.2. Results and discussion

The mean response times are shown in Fig. 2. There was a significant interaction between set type and size, $F(8,120) = 8.82$, $p < .00001$, and post hoc analysis (Newman–Keuls) showed significant differences between BL and TDf means at set sizes 16 ($p < .002$), 24 ($p < .0002$) and 32 ($p < .0002$); a significant difference between BL and TDo at set sizes 16 ($p < .0002$), 24 ($p < .0002$) and 32 ($p < .0002$); and a significant difference between TDf and TDo means at set size 24 ($p < .002$), with marginally significant differences at set sizes 16 ($p < .07$) and 32 ($p < .1$). There was also a significant effect of set type on response time slope, $F(2,30) = 19.57$, $p < .00001$. The response time slopes were 26.5 ms/item (BL), 36.4 ms/item (TDf), and 39.5 ms/item (TDo). Post hoc analysis also showed a significant difference between BL and TDf slopes ($p < .0004$), but not TDf and TDo slopes ($p < .34$). There was a significant effect of set type on errors $F(2,30) = 3.81$, $p < .04$. The error rates were 0.018 (BL), 0.019 (TDf), and 0.027 (TDo). There were no other significant effects.

The significantly steeper response time slopes for the target–distractor orientation similarity and target–distractor frequency similarity conditions in contrast to the no similarity baseline condition confirms the prediction that greater target–distractor similarity induces longer response times. The difference between target–distractor orientation similarity and target–distractor frequency similarity suggests that participants preferred the orientation dimension. In a post-experiment questionnaire, most participants indicated that they felt they searched first on the basis of orientation.

4. Experiment 2

In this experiment, we test the second basic prediction that increased distractor–distractor similarity results in shallower search slopes, by varying distractor–distractor similarity with target-distractor similarity held constant. AET, but not FIT, explicitly predicts this effect.

4.1. Method

4.1.1. Participants

Sixteen participants did the experiment. Two participated in the previous experiment.

4.1.2. Conditions

Participants performed search over three display type conditions: the no similarity baseline (BL), same as Experiment 1; *distractor–distractor orientation similarity* (DDo), where distractors shared a common orientation feature with each other; and *distractor–distractor frequency similarity* (DDf), where distractors shared a common frequency feature with each other.

4.2. Results and discussion

Fig. 3 shows the mean response times. There was a significant interaction between set type and size, $F(8,120) = 7.23$, $p < .00001$, and post hoc analysis showed significant differences between BL and DDf means at set sizes 16 ($p < .02$), 24 ($p < .0003$) and 32 ($p < .03$); a significant difference between BL and DDo means at set sizes 16 ($p < .02$), 24 ($p < .0002$) and 32 ($p < .0002$); and significant differences between DDf and DDo means at set sizes 24 ($p < .0003$) and 32 ($p < .0002$). There was a significant effect of set type on response time slopes, $F(2,30) = 11.89$, $p = .0002$. The mean slopes were 25.7 ms/item (BL), 21.2 ms/item (DDf), and 18.0 ms/item (DDo). Post hoc analysis revealed significant differences between BL and DDf ($p < .01$), and

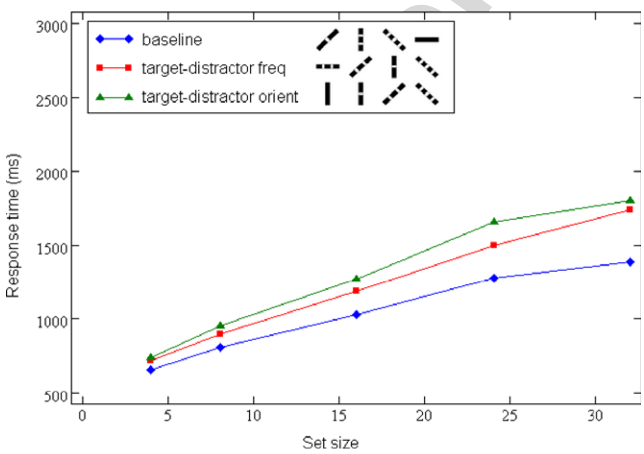


Fig. 2. Display type by size response times and sample target (left) and distractor items for Experiment 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

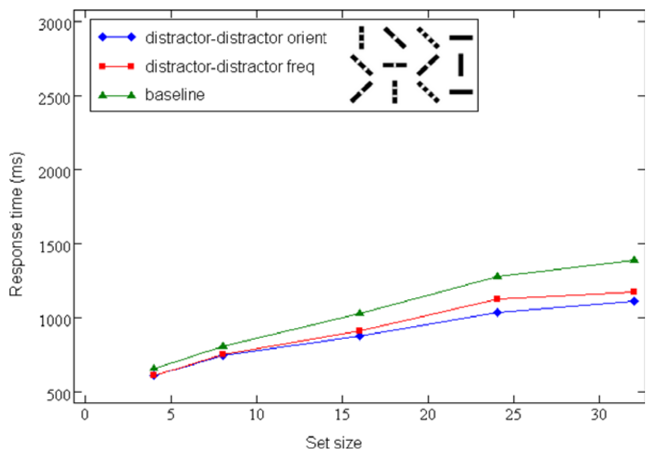


Fig. 3. Display type by size response times and sample target (left) and distractor items for Experiment 2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

BL and DDo ($p < .0005$); and the difference between DDf and DDo was marginally significant ($p = .052$). There were no significant effects on errors. The error rates were 0.013 (BL), 0.013 (DDf), and 0.016 (DDo).

The faster response times for distractor–distractor orientation similarity and distractor–distractor frequency similarity conditions in contrast to the no similarity baseline condition is consistent with the prediction by AET of shallower search slope with increased distractor–distractor similarity, which is not accounted for by FIT. In addition, the advantage for target–distractor feature sharing along the orientation dimension observed in Experiment 1 also extends to distractor–distractor feature sharing.

5. Experiment 3

The first two experiments confirm separately the effects of target–distractor and distractor–distractor similarity. In this experiment, we test the relative influence of these two components on search time. If target–distractor similarity has a stronger influence than distractor–distractor similarity, then we expect the slopes for display set types with shared target–distractor and distractor-distractor features to be steeper than display set types with no shared features. Conversely, if distractor-distractor similarity has greater influence then we expect the slopes to be shallower. AET is neutral on this issue, since response time is simply taken to be a ratio of these two factors. FIT predicts that slopes will be greater, because it does not recognize any effect for distractor–distractor similarity.

5.1. Method

5.1.1. Participants

Sixteen participants were recruited for the experiment. One participant did Experiment 1 and three participants did Experiment 2.

5.1.2. Conditions

Participants performed search over three display type conditions: the no similarity baseline (BL), same as Experiments 1 and 2; *target–distractor and distractor–distractor orientation similarity* (TD&DDo), where a distractor shared a common orientation feature with the target and other distractors shared a common orientation feature with each other; and *target–distractor and distractor–distractor frequency similarity* (TD&DDf), where a distractor shared a common frequency feature with the target and other distractors shared a common frequency feature with each other.

5.2. Results and discussion

Fig. 4 shows the mean response times. Analysis indicated a significant interaction between set type and size, $F(8,120) = 5.30$, $p < .00001$, and post hoc analysis showed significant differences between BL and TD&DDf means at set sizes 8 ($p < .05$), 16 ($p < .002$) and 32 ($p < .0002$); significant differences between BL and TD&DDo means at set sizes 16 ($p < .03$), 24 ($p < .0002$) and 32 ($p < .0002$), and a marginally significant difference at size 8 ($p < .07$); and significant differences between TD&DDf and TD&DDo means at set sizes 24 ($p < .02$) and 32 ($p < .0002$). There was a significant effect of set type on response time slopes, $F(2,30) = 11.22$, $p < .0005$. The mean slopes were 28.0 ms/item (BL), 31.4 ms/item (TD&DDf), and 36.8 ms/item (TD&DDo). Post hoc analysis revealed a significant difference between TD&DDf and TD&DDo ($p < .008$); and the difference between BL and TD&DDf was marginally significant ($p = .08$). There was a marginally significant effect of display type on errors $F(2,30) = 3.28$, $p < .06$, but no significant difference between means, nor set by type interaction. Errors rates were 0.02 (BL), 0.03 (TD&DDf), and 0.03 (TD&DDo).

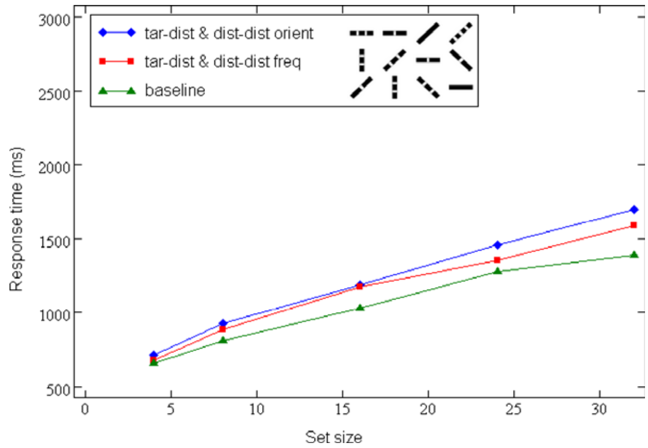


Fig. 4. Display type by size response times and sample target (left) and distractor items for Experiment 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

The results indicate that the influence of target–distractor feature similarity is stronger than distractor–distractor similarity. This result is consistent with both AET and FIT, but for different reasons as mentioned above. Furthermore, the relative advantage for orientation over frequency was replicated here.

6. Experiment 4

This experiment is an extension of Experiment 2 to include shared distractor features on both dimensions. AET (but not FIT) predicts that shallower slopes with shared distractor features on both dimensions than either dimension.

6.1. Method

6.1.1. Participants

Fifteen participants did the experiment. Four participants did Experiment 3, and two of those participants did Experiment 2.

6.1.2. Conditions

Participants performed search over three display type conditions: *distractor–distractor orientation and frequency similarity* (DDof), where distractors shared a common orientation and a common frequency feature with each other; and distractor–distractor orientation (DDo) and distractor–distractor frequency (DDf) conditions from Experiment 2.

6.2. Results and discussion

Fig. 5 shows the mean response times. There was a significant interaction between set type and size, $F(8,112) = 13.8$, $p < .00001$, and post hoc analysis showed significant differences between DDof and DDo means at set sizes 8 ($p < .04$), 24 ($p < .0002$) and 32 ($p < .0005$); significant differences between DDof and DDf means at set sizes 8 ($p < .02$), 24 ($p < .0002$) and 32 ($p < .0002$); and significant differences between DDo and DDf means at set sizes 24 ($p < .002$) and 32 ($p < .01$). There was also a significant effect of set type on response time slopes, $F(2,28) = 36.850$, $p < .00001$. The mean slopes for the three display types were 13.3 ms/item (DDof), 17.5 ms/item (DDo) and 19.7 ms/item (DDf). Post hoc analysis showed that differences between means were significant: DDof < DDo, $p < .0002$; and DDo < DDf, $p < .007$. There was no effect of display type on errors. The error rates were 0.02 for all display types.

The difference between the distractor–distractor orientation and frequency similarity condition and the distractor–distractor orientation similarity condition was predicted by AET. The significant difference in distractor–distractor orientation similarity and distractor–distractor frequency similarity conditions also replicates the advantage for orientation observed in previous experiments.

7. Experiment 5

This experiment mirrors Experiment 4 by contrasting shared target–distractor features on one dimension against both dimensions. Both AET and FIT predict steeper slopes with shared features on both dimensions. It also provides a further test of the effect of shared distractor–distractor features by contrasting display sets with shared target–distractor features on a single dimension with shared target–distractor and distractor–distractor features on the same dimension. AET predicts shallower slopes in the latter case, but FIT does not.

7.1. Method

7.1.1. Participants

Fifteen participants undertook the experiment. Two participants did Experiment 4 and one participant did Experiment 3.

7.1.2. Conditions

Participants performed search over three display type conditions: *target–distractor orientation and frequency similarity* (TDof), where distractors shared a common orientation and a common frequency feature with the target; target–distractor orientation similarity (TDo) from Experiment 1; and target–distractor and distractor-distractor orientation similarity (TD&DDo) from Experiment 3.

7.2. Results and discussion

The mean response times are shown in Fig. 6. Analysis revealed a significant interaction between set type and size, $F(8,112) = 15.2$, $p < .00001$, and post hoc analysis showed significant differences between TD&DDo and TDo means

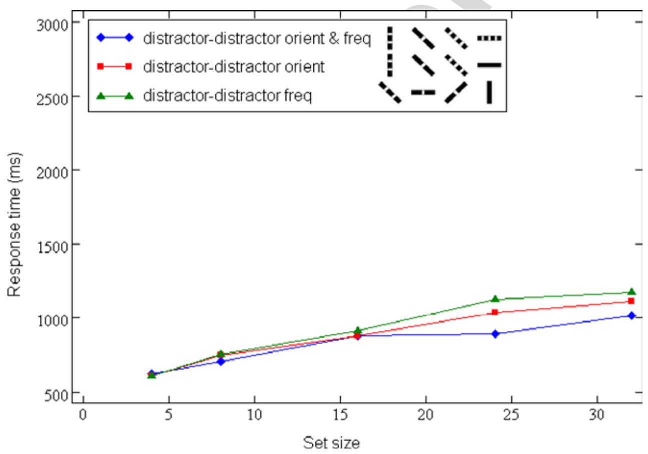


Fig. 5. Display type by size response times and sample target (left) and distractor items for Experiment 4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

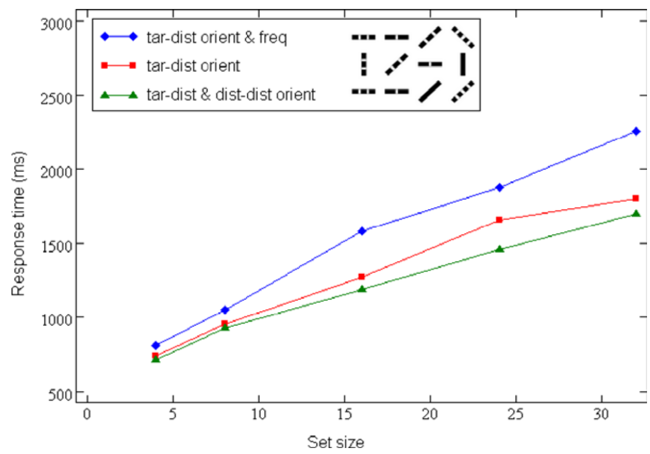


Fig. 6. Display type by size response times and sample target (left) and distractor items for Experiment 5. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

at set sizes 16 ($p < .05$), 24 ($p < .0002$) and 32 ($p < .006$); and significant differences between TDo and TDof means at set sizes 16 ($p < .0002$), 24 ($p < .0003$) and 32 ($p < .0002$). There was a significant effect of set type on response time slopes, $F(2,28) = 35.10$, $p < .00001$. The mean slopes were 32.1 ms/item (TD&DDo), 38.4 ms/item (TDo), and 50.1 ms/item (TDof). Post hoc analysis showed that differences between means were significant: TD&DDo < TDo, $p < .02$; and TDo < TDof, $p < .0002$. There was a marginally significant effect of display type on errors $F(2,28) = 2.87$, $p < .08$, but no significant difference between means, nor set by type interaction. The error rates were 0.04 (TD&DDo), 0.03 (TDo), and 0.06 (TDof).

The steeper slope for the target–distractor orientation and frequency similarity condition confirms the predictions of AET and FIT that shared target–distractor features on both dimensions results in slower response times than with sharing on one dimension. The shallower slope for the target–distractor and distractor–distractor orientation similarity condition than the target–distractor orientation similarity condition confirms the advantage for distractor similarity as predicted by AET.

8. Experiment 6

Experiments 1–4 suggested a consistent advantage for orientation over frequency and many participants indicated a preference for this dimension. In this experiment, we explicitly test this difference.

8.1. Method

8.1.1. Participants

Sixteen participants were recruited for the experiment. (The data for two participants were discarded because of failure to respond for contiguous sequences of trials.) One participant did Experiment 5, two participants did both Experiments 4 and 3, and two other participants did Experiment 2.

8.1.2. Conditions

Participants performed search over four display type conditions: *univariate no similarity orientation baseline* (UoBL), where all items had a unique orientation feature and the same frequency feature; *univariate distractor-distractor orientation similarity* (UDDo), where distractors shared a common orientation feature with each other and all items had the same frequency feature; *univariate no similarity frequency baseline* (UfBL), where all items had a unique frequency feature and the same orientation feature; and *univariate distractor-distractor frequency similarity* (UDDf), where distractors shared a common frequency feature with each other and all items had the same orientation feature.

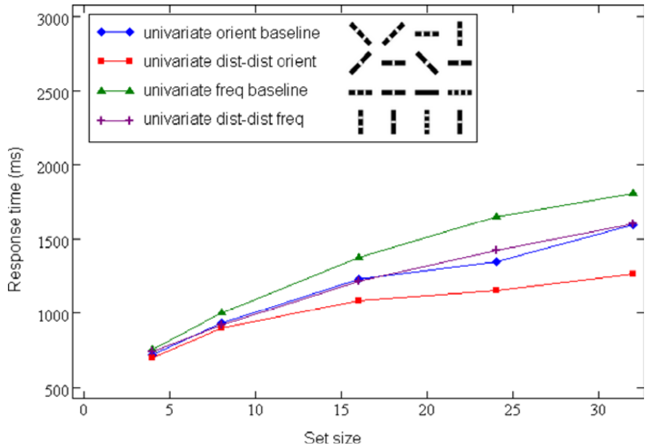


Fig. 7. Display type by size response times and sample target (left) and distractor items for Experiment 6. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

8.2. Results and discussion

The mean response times are shown in Fig. 7. A three-way repeated measures ANOVA indicated a significant effect of feature dimension $F(1,13) = 34.73$, $p < .00005$; and similarity $F(1,13) = 163.12$, $p < .00001$, but no dimension by similarity interaction. A two-way ANOVA of response time slopes indicated a significant effect on dimension, $F(1,13) = 38.66$, $p < .00003$; and similarity, $F(1,13) = 56.6$, $p < .00001$, and a marginally significant interaction, $F(1,13) = 3.12$, $p < .1$. The response time slopes for the four display types were 29.6 ms/item (UoBL), 37.6 ms/item (UfBL), 18.6 ms/item (UDDo), and 30.5 ms/item (UDDf). There were no significant effects on errors. The error rates were 0.02 (UoBL), 0.03 (UfBL), 0.02 (UDDo), and 0.03 (UDDf).

The results show that participants found the orientation dimension more salient than the frequency dimension in both univariate baseline and distractor–distractor conditions. Thus, participants found the orientation dimension more salient than the frequency dimension. Both AET and FIT are neutral with respect to dimension saliency, though it may not be difficult to incorporate this component in models derived from either theory.

9. Experiment 7

Experiments 2, 4, and 5 provide direct support for AET over FIT in the form of distractor–distractor similarity effects. Nonetheless, Treisman (1992) provided direct evidence of a search effect that was predicted by FIT, but not AET. In Treisman's experiment there were four display type conditions: search for a blue vertical bar among blue tilted and pink vertical bars (standard conjunction); standard conjunction search but with 50% of the blue tilted bars and pink vertical bars replaced by pink tilted bars (extended conjunction); and two "grey-disk" conditions where participants searched for a blue vertical bar among equal numbers of blue tilted bars and grey disks (grey-disks-1), or pink vertical bars and grey disks (grey-disks-2). AET predicted a shallower slope for the extended conjunction than standard conjunction, or either of the grey-disk conditions because the increase in shared distractor features should cause greater suppression. FIT also predicted shallower slopes for the two grey-disk conditions than standard conjunction, because conjunction search time is a function of the number of distractors sharing a feature with the target. But, FIT predicts the slopes for grey-disk conditions should also be shallower than extended conjunction, because identifying the target object only requires search along one dimension (either color, or orientation), whereas search along both dimensions is required in the extended conjunction condition.

We modified Treisman's design to make use of the common set of items used in the previous experiments. For "grey disks," we used bars with orientation and frequency features not shared by either target or other distractors. However, the predictions for both theories are unchanged. AET predicts a shallower slope for extended conjunction than either grey-disks condition, because of the increase in distractor–distractor similarity. FIT predicts shallower slopes for both grey-disks conditions, because search only requires one feature dimension instead of two.

9.1. Method

9.1.1. Participants

Sixteen participants performed the experiment. None of the participants did any of the previous experiments.

9.1.2. Conditions

The four display type conditions in this experiment were: *standard conjunction* (SC), where distractors shared common orientation and frequency features with the target; *extended conjunction* (XC), where some distractors shared a common orientation or frequency feature with the target and other distractors shared no features with the target; *target–distractor orientation similarity with matched distractor ratio* (TDoM), where the proportion of distractors sharing a common orientation feature with the target was matched to the proportion of distractors sharing no feature with the target or the other distractors; and *target–distractor frequency similarity with matched distractor ratio* (TDfM), where the proportion of distractors sharing a common frequency feature with the target was matched to the proportion of distractors sharing no feature with the target or the other distractors. The matched distractor ratio conditions correspond to Treisman's "grey disks" conditions.

9.2. Results and discussion

The mean response times are shown in Fig. 8. Analysis indicated a significant interaction between set type and size, $F(12,180) = 36.2$, $p < .00001$, and post hoc analysis showed that means for SC sets were significantly greater than the other sets at set sizes 8 ($p < .005$), 16 ($p < .0001$), 24 ($p < .0001$) and 32 ($p < .0001$); and significant differences between TDfM and TDoM at set sizes 16 ($p < .003$), 24 ($p < .003$) and 32 ($p < .00002$). There was a significant difference between XC and TDoM at set size 24 ($p < .05$), and between XC and TDfM at set size 32 ($p < .0001$), but not at other set sizes. There was also a significant effect of set type on response time slopes, $F(3,45) = 96.11$, $p < .00001$. The mean slopes were 55.9 ms/item (SC), 31.1 ms/item (XC), 34.3 ms/item (TDoM) and 21.8 ms/item (TDfM). Post hoc analysis showed significant differences between SC and the other set types ($p < .0002$); between XC and TDfM ($p < .0002$); and TDoM and TDfM ($p < .0002$); but not XC and TDoM. There was an effect of display set type on errors, $F(3,45) = 4.88$, $p < .006$, but the pairwise differences between means were not significant by post hoc analysis. The error rates were 0.025 (SC), 0.028 (XC), 0.016 (TDoM) and 0.016 (TDfM). There was a marginally significant type by size interaction, $F(12,180) = 1.71$, $p < .07$, but post hoc analysis did not show any significant differences between same set size means.

The interaction between set type and size provides only partial support for both theories. FIT, but not AET, predicted the advantage for standard conjunction over extended conjunction. However, the significantly faster response time for the extended conjunction condition than the target–distractor orientation similarity with matched distractor ratio condition at set size 24 supports AET, not FIT. Clearly, dimension saliency is a contributing factor. Although the task here is logically equivalent to Treisman's, the general difficulty of searching along the frequency dimension may preclude possible parallel feature search advantage.

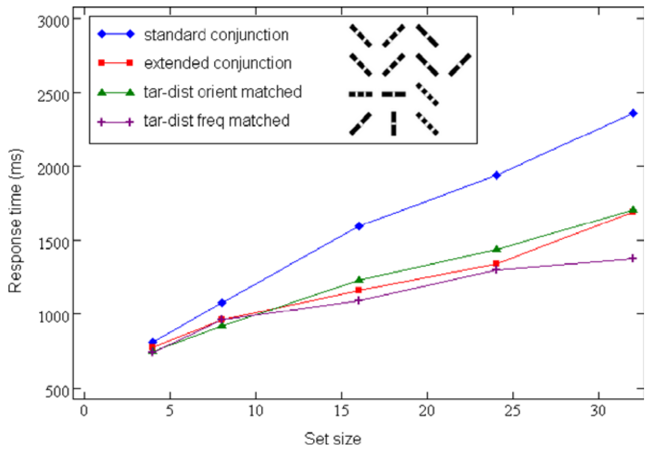


Fig. 8. Display type by size response times and sample target (left) and distractor items for Experiment 7. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

10. Model

The data showed three types of effects: target–distractor similarity, distractor–distractor similarity, and dimension saliency. Similarity combined with saliency, and together with set size constitute five factors that go into our model for search time. That is, target–distractor and distractor–distractor similarity for the orientation dimension, target–distractor and distractor–distractor similarity for the frequency dimension, and set size. In our model, *target–distractor similarity is defined as the number of items, including the target, that share the same feature as the target on the specified dimension*. That is

$$s^{\mathrm{T}} = \sum_{i} (I_t = I_i), \tag{1}$$

where I_t is the target item, and I_i are the items in the display set. For example, the display set {(0 0),(0 1),(1 1),(2 2)} has a target–distractor similarity measure $\langle s^{\mathrm{TO}}, s^{\mathrm{TF}} \rangle = \langle 2,1 \rangle$, where pair (0 0) is the target; and s^{TO} and s^{TF} are the target–distractor measures along the orientation and frequency dimensions, respectively. *Distractor–distractor similarity is defined as the number of distractor–distractor pairs that share a common feature along the specified dimension*. That is,

$$s^{\mathrm{D}} = \sum_{i,j} \left(I_i^{\mathrm{D}} = I_j^{\mathrm{D}} \right), \tag{2}$$

where I_i^{D} and I_j^{D} are distractor items in the display set. By this measure, $(I_i^{\mathrm{D}}, I_i^{\mathrm{D}})$, $(I_i^{\mathrm{D}}, I_j^{\mathrm{D}})$, and $(I_j^{\mathrm{D}}, I_i^{\mathrm{D}})$ are regarded as distinct pairs, where $i \neq j$. For example, the distractor subset {(0 1),(1 1),(2 2)} has a distractor–distractor similarity measure $s = \langle s^{\mathrm{DO}}, s^{\mathrm{DF}} \rangle = \langle 3,4 \rangle$, where s^{DO} and s^{DF} are the distractor–distractor measures along the orientation and frequency dimensions, respectively. Table 1 summarizes the display set type conditions. Note that similarity may also depend on set size, because in some cases the number of shared features will increase with more distractors.

The idea of feature-sharing as the basis for predicting search time comes from AET. Nonetheless, the model proposed here differs from AET in several important ways. First, in AET, similarity is the sum of the number of shared features between items. Thus, sets {(0 0),(0 1)} and {(0 0),(1 0)} have the same measure of similarity. However, the data indicate that the effects of feature sharing are dependent on the saliency of the dimension. Accordingly, our model assumes the effects of feature sharing along each dimension are independent. Second, an overall measure is obtained by a ratio of target–distractor to distractor–distractor similarity measures in AET. Similarity increases as approximately the square of set size. Under constant target–distractor similarity, this measure implies search time as a nonlinear function of set size, contrary to what is generally observed in the data. We assume a diminishing influence with increased sharing. Hence, search time is modeled as a power (square root) function of the amount of feature sharing.

Search time (t) to find a target is modeled by the following equation:

$$t = A \cdot s^{1/2} + Bn + C, \tag{3}$$

where $s = \langle s^{\mathrm{TO}}, s^{\mathrm{DO}}, s^{\mathrm{TF}}, s^{\mathrm{DF}} \rangle$ is the four-dimensional vector of similarity measures, n is the number of items in the display set, A and B are parameters and C is the constant offset term. ("·" is the vector dot product.) Factors s and n are computed for all 85 (i.e., 17 type by 5 size) conditions. A linear regression of $s^{1/2}$ and n onto the mean response times accounted for 98.2% of the variance of set-size means. Fig. 9 shows the fit of the model to the search time data. Linear regressions of set size onto obtained (participant) and estimated (model) response time means were computed for each display set condition. Participant and model response time slopes are shown in Fig. 10. The fitted parameters (and t-values) of the model are $A = \langle A^{\mathrm{TO}}, A^{\mathrm{DO}}, A^{\mathrm{TF}}, A^{\mathrm{DF}} \rangle = \langle 220, -47, 151, -41 \rangle$ (29, 14, 20, 12); $B = 78$ (31); and $C = 145$ (8), where A^{TO} and A^{DO} are the target–distractor and distractor-distractor parameters for the orientation dimension; and A^{TF} and A^{DF} are the parameters for the frequency dimension. For the purpose of regression analysis, mean response times were calculated after pooling conditions that occurred in multiple experiments.

For comparison, a model without the power term (i.e., linear function of number of feature-sharing pairs and number of items) accounted for 92.2% of the variance; a model without the similarity term (i.e., linear function of number of display items only) accounted for 66.5% of variance; a model without the set size term (i.e., power function of number of feature-sharing pairs) accounted for 77.3% of variance; and a model based on the ratio of target–distractor to distractor-distractor similarity (as used in AET) for each dimension accounted for 66.7% of the variance.

In addition to accounting for most of the variance on display type-size means, the relative differences between response time slopes for specific sets were also in the same direction as the significant differences observed in the data.

The parameters obtained above were determined by fitting the model to all the 85 type-size conditions. We also tested predictability by fitting the model to 8 of the possible 17 set type conditions and using the derived parameters as the basis for the regressions of estimated onto obtained response times for the remaining 9 type by 5 size (45 type-size) conditions. The percentage of variance in the remaining 45 response time means accounted for by the model (R^2) was used as the measure for the model's predictability. R^2 values were computed for all 24310 combinations of 8 conditions minus 308 combinations for which there was either no deviation in one or more of the similarity measures, or the covariation (Pearson's R) in the distractor–distractor measure for orientation and frequency dimensions was greater than 0.98. Fig. 11 shows the number of combinations of 8 conditions yielding a given R^2 value. The mean R^2 over the 24002 combinations was

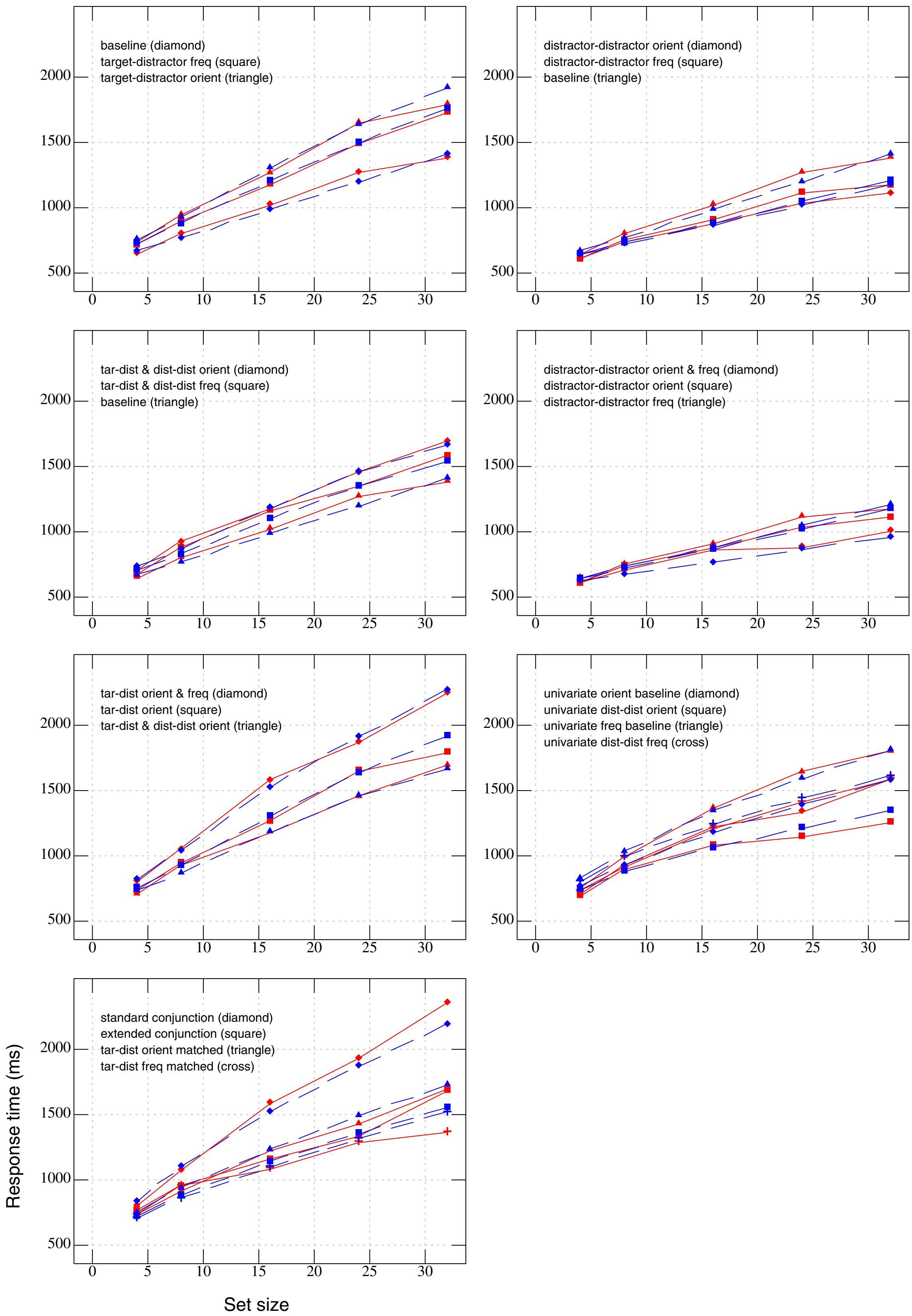


Fig. 9. Participant (solid lines) and model (dashed lines) response times. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

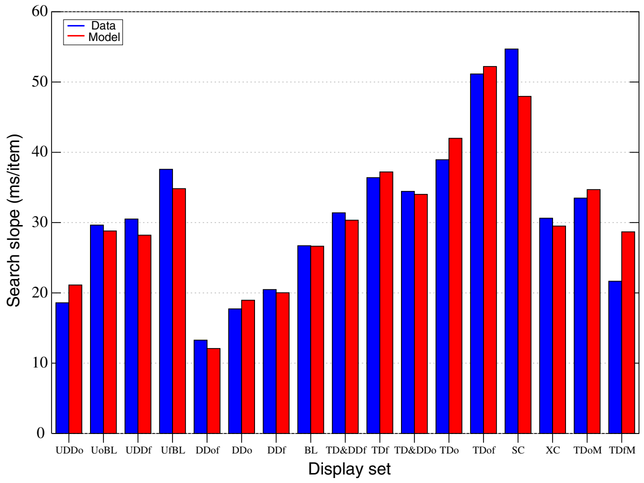


Fig. 10. Participant and model search slopes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

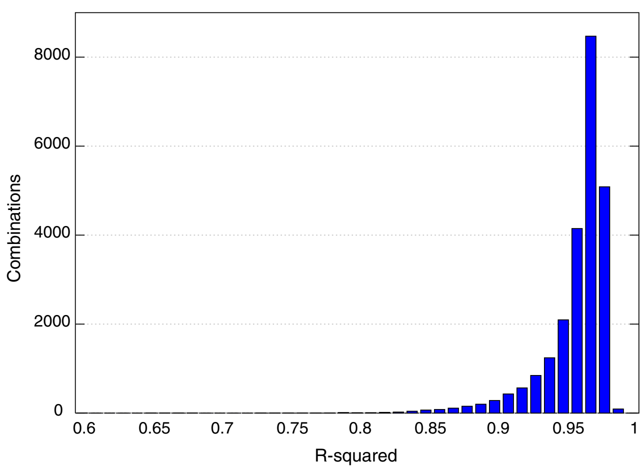


Fig. 11. Predictability of model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

95.9 and the number of combinations yielding an R^2 greater than 0.95 was 19019 (i.e., 79.2%), indicating that the model achieved good predictability to novel display conditions.

10.1. Interpretation

Eq. (3) suggests visual search as a (linear) combination of item similarity and serial search. In its current form, though, it is difficult to see how this model fits within a typical parallel/serial processing framework, because the similarity term does not modulate the search step. Here, we provide two interpretations of the model, where search is primarily a parallel process in one version and serial in the other.

10.1.1. Parallel search

We suppose that the visual search process involves activating a field of units, one per display item, until a critical threshold is reached. Thus, response time depends on the rate of increase in activation and the height of the activation response threshold. So, for example, if the rate of activation change is r and the activation response threshold is I_θ, then the response time is given by:

$$t = \frac{I_\theta}{r} + T, \tag{4}$$

where T is some constant amount of time needed to initiate the process and/or make the appropriate response once the target has been identified. If the increase in activation to the entire field of units representing the display items occurs at a constant rate R, then the rate of activity change for individual units is $r_i = R/n$, where n is number of units representing display items (i.e., the display set size); and the time to activate an individual unit to response threshold I_θ is

$$t = \frac{I_\theta}{r_i} + T = \frac{I_\theta}{R} n + T. \tag{5}$$

We also suppose that the response threshold is modulated by an amount ΔI_θ that is dependent on display conditions. Therefore, response time is given by

$$t = \left(\frac{\Delta I_\theta + I_\theta}{R}\right) n + T. \tag{6}$$

Eq. (6) has the same form as a refactored version of the abstract model:

$$t = (A \cdot s_n + B)n + C, \tag{7}$$

where $s_n = \frac{1}{n} s^{1/2}$ is the *normalized* similarity measure. Hence, the first term in the abstract model, $A \cdot s_n \equiv \frac{\Delta I_\theta}{R}$, is interpreted as the adjustment in activation response threshold due to display item similarity. Greater similarity (feature-sharing) between target and distractors increases the response threshold and therefore the response time. Greater distractor–distractor similarity decreases the response threshold and therefore the response time. The second term, $B \equiv \frac{I_\theta}{R}$, is interpreted as the base response threshold. The more items in the display and therefore the more units that share activation, the slower the rate of activation change to individual units and therefore the longer the response time.

10.1.2. Serial search

A serial implementation of the model is obtained by rewriting Eq. (3) as

$$t = B\left(\frac{A}{B} \cdot s^{1/2} + n\right) + C, \tag{8}$$

where $\frac{A}{B} \cdot s^{1/2} + n$ is the mean number of items checked for a match to the target. A random walk through the display set without maintaining a memory for the locations of items checked requires on average n search steps, where n is the size of the display set. Greater target–distractor similarity means that search is more likely to return to distractors that have already been checked—apparent set size will be larger than the actual display set size and search time will take longer. Greater distractor–distractor similarity means that search is more likely to skip distractors not yet checked—apparent set size will be smaller and search time will be shorter. However, the extent to which memory influences search is still a controversial issue. Several reports have indicated that relatively few items are reexamined during search (Muller & Muhlenen, 2000; Takeda & Yagi, 2000; Takeda, 2004). Yet, eye movement data have shown significant returns to previously fixated locations (Gilchrist & Harvey, 2000; Hooge, Over, van Wezel, & Frens, 2005).

11. General discussion

The results reported here indicate that inter-item similarity is central to search efficiency over a wide range of display conditions; and tend to support Attention Engagement Theory over Feature Integration Theory. Although Experiments 1, 3 and 5 provided support for both theories, Experiments 2, 4 and the combined target–distractor/distractor–distractor condition in Experiment 5 favored AET over FIT. Experiment 6 highlighted the relative preference for the orientation dimension that while not explicitly part of AET, or FIT could be accommodated by either theory. However, Experiment 7 indicated that the relative influence of target–distractor and distractor–distractor similarity is not explained by AET. Both theories only partially accounted for the data in this experiment. And, when the seven experiments were taken together, AET did not provide a good overall fit to the data.

Clearly, search time as a ratio of target–distractor to distractor–distractor similarity is too simple, and did not account for the subtle variations in performance observed here. Our abstract model refines that relationship in a way that addresses these problems. The model we proposed independently weights the dimension-specific contributions of target–distractor and distractor–distractor similarity. It accounted for most of the variance in search time over display conditions ranging from univariate (feature search) to bivariate (conjunction search) with independent manipulation of target–distractor and distractor–distractor similarity on both orientation and frequency feature dimensions. Thus, it provides strong support for the idea that a single mechanism underlies the apparent differences between feature and conjunction search.

What, then, should be the form of such a unifying mechanism? Even within our more refined model, there is still variation in possible mechanisms. As already discussed, the model permits both a parallel implementation, where all items enter into a race toward a decision threshold activation level; and a serial implementation, where only one item is checked at a time. An example of a fully parallel model that implements the original ideas of inter-item similarity proposed in AET is the Search via Recursive Rejection model of Humphreys and Muller (1993), although they did not provide specific predictions for conjunction search.

Another possibility is a combination of parallel/serial processing such as proposed by Wolfe and colleagues in their Guided Search model. In the Guided Search model, each item is activated in parallel according to a normalized exponential function of the difference between it and the other items, minus a weighted linear function of the features it shares with the target. Search is then a serial process of checking each item for a match to the target in rank order of activation. When target–distractor similarity is increased, the activation of the target item is lowered relative to dissimilar distractors, pushing the target to the tail of the order and increasing search time. When distractor–distractor similarity is increased, the activation of distractor items will be lowered relative to the target, pushing the target item toward the head of the order and so reducing search time. Thus, the Guided Search model makes similar predictions to AET with respect to inter-item similarity.

Like AET, Guided Search proposes an initial parallel phase for filtering/weighting the display items followed by a serial phase of matching items to the target. Guided Search appears to be analogous to AET in that the linear function of items sharing features with the target corresponds to the target–distractor similarity component and the exponential difference function corresponds to distractor–distractor similarity component. However, there is a subtle difference between Guided Search and AET, and by extension our model. Distractor–distractor similarity is computed over non-target items, whereas the difference function is computed over all items. So, while increasing target–distractor similarity lowers target activation relative to dissimilar distractors, it may also lower the activation of distractors sharing a feature with the target relative to distractors not sharing a target feature. In this case, the target will be pushed towards the head of the order, decreasing search time. For example, the target–distractor orientation condition ({00 01 12 23}) has greater target–distractor similarity than the no similarity baseline condition ({00, 11, 22, 33}), but the two conditions have the same distractor–distractor similarity. The activation of the 01 distractor will also be lowered relative to the 12 and 23 distractors because it shares a feature with the target. For AET and our model, the increased search slope is due only to the increase in target–distractor similarity. For Guided Search, the difference in search slope depends on relative changes to target and distractor item activations. Just how these components balance out will depend on the precise details and weighting of the activation function. Although beyond the scope of the current paper, it may be possible to tease apart these differences experimentally to determine which supports/conflicts with the data.

Our model can also be extended to triple conjunction search and continuous features. Wolfe et al. (1989) reported steeper search slope for trivariate display sets when the target shares two features with the distractors than when only one feature is shared. This effect was modeled by Guided Search. In our case, the first condition ({000, 001, 010, 100}) has greater target–distractor similarity than the second ({000, 011, 101, 110}). Likewise, then, our model also predicts a shallower slope for triple conjunction.

We assumed a discrete (categorical) representation for features. Continuous feature representations can be implemented by replacing the equality term ($x = y$) in Eqs. (1 and 2) with a weighted nonlinearity term, such as $\frac{2}{1+e^{w|x-y|}}$, which becomes equivalent to the discrete model as weight w becomes large. However, the influence of intra-dimensional feature relationships is likely to involve computing more than a *distance measure*, at least for orientation. Search is faster when distractors are symmetric about the vertical axis than when they are not (Wolfe & Friedman-Hill, 1992), indicating that other computational mechanisms may be involved.

Although our results admit a parallel processing model for inefficient search, this does not imply that all search is necessarily parallel. A number of studies have provided clear evidence for item-to-item serial deployment of attention (Bricolo, Gianesini, Fanini, Bundesen, & Chelazzi, 2002; Woodman & Luck, 1999; Woodman & Luck, 2003). These cases involved very difficult search conditions resulting in slopes of 100 ms/item or more. In contrast, the search conditions used here were relatively easy with slopes between 13 and 56 ms/item.

Further work is needed to determine the extent to which this model also explains efficient search, and the broader behavioral aspects of feature/conjunction search, such as response accuracy. Efficient search is often observed for unidimensional arrays with only one type of distractor. In our terms, this would include set types such as {00, 10, 10, 10}. For this set, our model predicts a search slope of 8.7 ms/item, without changing any of its parameters. Yet, it also predicts a negative slope (−10.7 ms/item) for the set type {00, 11, 11, 11}. A negative slope was observed as part of a U-shaped effect for feature search (Sagi & Julesz, 1987)—when the distance between items is sufficiently small, search becomes faster with increased set size. Santhi and Reeves (2004) explained this effect by their model, where the target–distractor contrast term eventually dominates the noise term due to distractors as more distractors are included in the search set. In our case, though, the predicted slope is monotonic and implies that the similarity-based threshold adjustment term is greater than the base threshold term, in which case a subprocess is taking negative time. A reasonable normalizing assumption is to bound the magnitude of the reduction to be no greater than the base threshold, which prevents slopes becoming negative.

We have not addressed response accuracy in this work, because participants were given sufficient time to make a response so that accuracy was high. There is a family of models based on signal detection theory that account for a range of feature/conjunction conditions using a speeded response paradigm that elicits significant differences in accuracy (Eckstein, Thomas, Palmer, & Shimozaki, 2000). These models are termed *capacity unlimited* in that their internal representations are instantiated independently of set size, in contrast to *capacity limited* models that force a serial processing stage. Palmer (1998) provides a review of these two classes of models and a critical accuracy test to distinguish them. McElree and Carrasco (1999) have used the speed-accuracy tradeoff paradigm to show that both features and conjunctions are detected in parallel. This work and other response times studies (Starreveld, Theeuwes, & Mortier, 2004) have provided further evidence for a wider role of parallel processes in visual search.

A challenge for our model is to implement a version whereby response time can be traded for accuracy in a way that is consistent with this sort of data and the underlying neurodynamics. Several researchers have proposed detailed parallel processing models for inefficient search (Deco & Zihl, 2001; Herd & O'Reilly, 2002). These models were primarily motivated by neurocomputational principles, rather than detailed fitting of behavioral data. Nonetheless, we expect that incorporating these principles to explain the results presented here will be mutually informative for both the psychophysics and neuroscience of vision.

Acknowledgments

We thank Adam Reeves and two anonymous reviewers for their thoughtful comments and suggestions on this work.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.visres.2006.06.016.

References

Bricolo, E., Gianesini, T., Fanini, A., Bundesen, C., & Chelazzi, L. (2002). Serial attention mechanisms in visual search: a direct behavioral demonstration. *Journal of Cognitive Neuroscience, 14*(7), 980–993.

Cave, K. R., & Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology, 22*(2), 225–271.

Deco, G., & Zihl, J. (2001). Top-down visual attention: a neurodynamical approach. *Visual Cognition, 8*(1), 119–140.

Duncan, J., & Humphreys, G. (1992). Beyond the search surface: visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 578–588.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*(433-469), 381–410.

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & Psychophysics, 62*(3), 425–451.

Gilchrist, I. D., & Harvey, M. (2000). Refixation frequency and memory mechanisms in visual search. *Current Biology, 10*(19), 1209–1212.

Herd, S. A., & O'Reilly, R. C. (2002). Serial visual search from a parallel model. *Vision Research, 45*(24), 2987–2992.

Hooge, I. T., Over, E. A., van Wezel, R. J., & Frens, M. A. (2005). Inhibition of return is not a foraging facilitator in saccadic search and free viewing. *Vision Research, 40*(14), 1901–1908.

Humphreys, G. W., & Muller, H. J. (1993). SEarch via recursive rejection (SERR): a connectionist model of visual search. *Cognitive Psychology, 25*(1), 43–110.

McElree, B., & Carrasco, M. (1999). The temporal dynamics of visual search: evidence for parallel processing in feature and conjunction searches. *Journal of Experimental Psychology: Human Perception and Performance, 25*(6), 1517–1539.

Muller, H. J., & Muhlenen, A. von. (2000). Probing distractor inhibition in visual search: Inhibition of return. *Journal of Experimental Psychology: Human Perception and Performance, 26*(5), 1591–1605.

Palmer, J. (1998). Atentional effects in visual search: relating search accuracy and search time. In R. D. Wright (Ed.), *Visual Attention* (pp. 348–388). New York, NY: Oxford University Press.

Quinlan, P. T. (2003). Visual feature integration theory: past, present, and future. *Psychological Bulletin, 129*(5), 643–673.

Sagi, D., & Julesz, B. (1987). Short-range limitation on detection of feature differences. *Spatial Vision, 2*(1), 39–49.

Santhi, N., & Reeves, A. (2004). The roles of distractor noise and target certainty in search: a signal detection model. *Vision Research, 44*(12), 1235–1256.

Selst, M. V., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology, 47A*(3), 631–650.

Starreveld, P. A., Theeuwes, J., & Mortier, K. (2004). Response selection in visual search: the influence of response compatibility of nontargets. *Journal of Experimental Psychology: Human Perception and Performance, 30*(1), 56–78.

Takeda, Y. (2004). Search for multiple targets: Evidence for memory-based control of attention. *Psychonomic Bulletin and Review, 11*(1), 71–76.

Takeda, Y., & Yagi, A. (2000). Inhibitory tagging in visual search can be found if search stimuli remain visible. *Perception and Psychophysics, 62*(5), 927–934.

Treisman, A. M. (1991). Search, similarity, and integration of features between and within dimensions. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 194–214.

Treisman, A. M. (1992). Spreading activation or feature integration? A reply to Duncan and Humphreys (1992). *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 589–593.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–113.

Wolfe, J. M. (1994). Guided search 2.0: a revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202–238.

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences, 7*(2), 70–76.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 419–433.

Wolfe, J. M., & Friedman-Hill, S. R. (1992). On the role of symmetry in visual search. *Psychological Science, 3*(3), 194–198.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews. Neuroscience, 5*(6), 495–501.

Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature, 400*(6747), 867–869.

Woodman, G. F., & Luck, S. J. (2003). Serial deployment of attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance, 29*(1), 121–138.