Igor Utochkin

CONJUNCTIONS CAN GUIDE ATTENTION THROUGH VISUAL SEARCH

Igor S. Utochkin[1]

CONJUNCTIONS CAN GUIDE ATTENTION THROUGH VISUAL SEARCH[2]

Guided search is a mechanism that controls and optimizes the deployment of attention during visual search and allows one to pay attention only to highly relevant items. For instance, when searching for a conjunction of two features, we are able to select a feature-marked subset (e.g., all items sharing same color) prior to focusing attention on particular items. Standard models assume that only separate features can guide attention since they are only available at the preattentive stage of visual analysis and no conjunction information is available at that stage. Here I show that search performance is affected by both the distribution offeatures across the visual field and their conjunctions in particular items. It appears that people are unable to use "pure", unbound features for selecting relevant subsets. This major finding requires reconsidering the standard models of guided search. The concept of distributed attention, which represents multiple items as imperfectly bound objects, seems promising in explaining this finding.

---

[1]National Research University Higher School of Economics.Scientific-Educational Laboratory for Cognitive Research; Head of laboratory. E-mail: iutochkin@hse.ru

Our everyday environment is filled with dozens or even hundreds of various objects that share common properties, such as color, shape, size, etc. When looking for a particular object among other similar objects, one usually has to pay attention to each item serially until a target is found. This pattern describes a typical *conjunction search* task and is explained by the need to attentionally bind otherwise separate properties to see and locate an object properly (Treisman&Gelade, 1980). However, the strategy of simplyserially shiftingone's attention seems rather unreasonable and too slow, especially in large displays or scenes filled with lots of different objects. For example, if you occasionally forget where you have left your blue car,do you indeed inspect every car in theparking lot? Instead, you may want to use a salient feature, for example, blue color, and inspect only blue cars, which makes the search rather efficient. The results of laboratory experiments demonstrate that in large displays, a slow serial search for a conjunction target isreplaced with a fast and near parallel search (e.g., Friedman-Hill & Wolfe, 1995; Wolfe, Cave, &Franzel, 1989). This kind of result suggests that the visual system can use a more smart and economic strategy that allows for theselective allocation of attention to limited subsets of objects while ignoring others. This smart strategy is typically termed *guided search* (Wolfe, 1994; Wolfe et al., 1989).

Wolfe's most elaborated guided search model assumesthat the visual system can split heterogeneousitems to homogenous subsets and inspect subsets separately (Friedman-Hill & Wolfe, 1995; Wolfe, 1994). The model implies that this subset segmentation is based predominantly on early preattentive feature selection. According to Treisman'sstandard framework (Treisman&Gelade, 1980), fast and automatic preattentiveprocessing represents basic features in a set of retinotopic maps, for example, a map forblue, a map forcurvature, a map forvertical, etc. Every map can be subsequently selected to guide visual search. Consider, for example, a visual search task when the observer should find a red X among blue X's and red O's. Withthis template, an observer can select a subset of red items and ignore blue ones since these two colors are represented within separate preattentive maps. Within the map of red, searching for an O is easier because this is the only curve item among straight-lined X's. This is exactly what was demonstrated by Egeth, Virzi, andGarbart(1984) and Friedman-Hill and Wolfe (1995). Therefore, these are basic and separable features thatare assumed to be appropriate for visual guidanceby the classical guidance model.

At the same time, guided search does notconsider conjunctions of features to be a basis for attentional guidance, even when these conjunctions are formed within the same dimension (Friedman-Hill & Wolfe, 1995). In other words, it is assumed that we can select a relevant subset defined by either feature, but never by both.Again, the explanation is based on the nature of

preattentive processing. As features are unbound at the preattentive level, the visual system"knows" how features are distributed across a display, but has no "knowledge"ofthe variety of their conjunctions (or objects),prior to focused attention. This was discoveredby Treisman(2006), who found that observers were good atestimating proportions of features (either color or shape) in displays at a glance, but failed to estimate proportions of conjunctions.However, there is a potential problem with this experiment in that Treisman used an alphabet of three features within each dimension that could result in up to nine conjunction subsets. So, while estimating proportions of features, observers had to store the information about only three subsets that does notexceed working memory capacity (Luck & Vogel, 1997) and can be easily enumerated (Halberda, Sires, &Feigenson, 2006).But it appears much harder to store information about nineconjunctions or fewer. So, the results of this experiment can hardly serve as final evidence that the visual system has no knowledge about conjunctions.

I demonstrate that some information about conjunctions is actually available prior to focused attention, and can be used for guiding visual search along with information about features. Four experimentswere conducted in order to approve this statement.A simple notion underlies all experiments: If the same features differently co-occur in space, then visual search efficiency would differ since the number of conjunctions is different as well. In Experiments 1 and 2, observers performed a typical search for orientation×color conjunctions. The number of concurrently presented conjunctions was manipulated,while the alphabet of features remained constant. The results will show that the visual system somehow segments items into subsets based on conjunctions, rather than features. In Experiments 3 and 4, observers searched for the same orientation×color conjunctions and an irrelevant dimension (size) was manipulated in a way that provided segmentation to various numbers of conjunction-based groups, but not feature-based groups. Results will show that this conjunctive segmentation also affectsthe subsequent allocation of focused attention to objects.

Experiments 1 and 2: Search among two or three conjunctions

In order to provide a conjunction search, it is sufficient to use two dimensions (e.g., color and orientation), each having two values (e.g., black-white and vertical-horizontal). Then different values are to be combined in non-overlapping conjunctions and distributed among distractors (e.g. half distractors are black and vertical, while the other half is white and horizontal). Finally, a target should be defined as a hybrid conjunction sharing one feature with either distractor type (e.g., white vertical). Thus, up to three feature combinations are typically involved in such a conjunction search. However, four combinations, rather than three, are

actually available using such an alphabet of features. So, using the same features it is possible to vary the number of conjunctions presented. If only separate features guide one's attention, then visual search efficiency would depend on features distribution, not on the variety of conjunctions formed by these features. In contrast, if the visual system were able to select conjunctively defined subsets,then the number of conjunction would have an effect on search efficiency.

Methods

Participants. 18 (Exp. 1) and 16 (Exp. 2) undergraduate students of the Higher School of Economicstook part in these experiments for extra credit in their general psychology course. All reportedhaving normal or corrected-to-normal vision and no neurological problems.

Stimuli. White and black lines oriented vertically or horizontally presented on a gray field. Set sizes were 7, 13, or 37. In Experiment 1, each of four possible color×orientationconjunctions could be a target with equal frequency. In Experiment 2, only the white vertical lines could be a target.

Three distracting conditions were used. The standard 2-conjunction condition included two distracting conjunctions sharing either color or orientation with a target and the distractor ratio was 1/1 (Fig. 1A). In the 3-conjunctioncondition, distractors included three possible conjunctions,except the one reserved for a target (if any). Here the distractor ratio was 1/1/1 (Fig. 1B). Thus, the alphabet of features remained constant while the number of conjunctions varied. However, balancing the conjunctions ratio in a 3-conjunction condition made the feature ratio imbalanced (1/2 both in orientation and color) that could be used as an efficient *feature-based cue* to guide attention (Bacon &Egeth, 1997; Zohary& Hochstein, 1989). To dissociate between conjunction- and feature-based strategies, the third condition used two standard distracting conjunctions with a feature ratio of 1/2 (Fig. 1C).
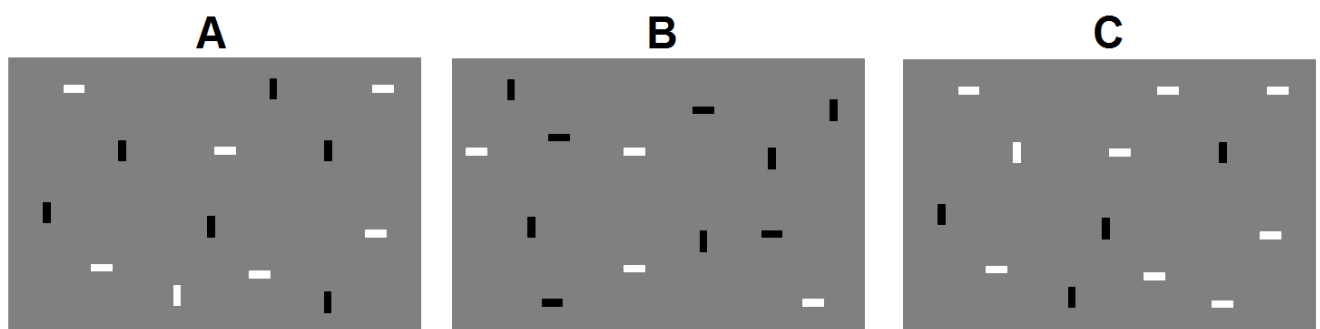

Figure 1. Three distracting conditions in Experiments 1 and 2. (A) 2 conjunctions, ratio 1/1; (B) 3 conjunctions, ratio 1/1/1; (C) 2 conjunctions, ratio 1/2.

Procedure. Participants were seatedat about 70 cm from the monitor. They were instructedto search for a unique color×orientation conjunction. They had to press one of two keys

on the response keypad as soon as they established whether a particular conjunction was present. In Experiment 1, observers looked for an *odd* conjunction, that is, any unique combination of color and orientation (*variable mapping* task). In Experiment 2, they looked for a known conjunction – a white vertical line (*consistent mapping* task).

24 trials were used for each condition of both experiments. The overall number of trials was 3 set sizes × 3 distracting conjunctions × 2 types of trials (target present/target absent) × 24 trials = 432 trials per observer. A 1-minute rest pause followed every 108 trials.

Results and discussion

Trials with incorrect or missed responses were excluded from the analysis. A within-subjects ANOVA was conducted on the rest trials. The model included the identity of participants as a random factor to handle individual differences between observers.

The results of Experiment 1 are summarized in Figure 2A. The main effect of the set size was significant (F (2, 34) = 59.93, p < .001, η^2_p = .776). The effect is provided by the reaction time (RT) increment with a set size that corresponds to the serial mode of visual search typical for conjunction search. Distracting conditions also had a significant effect on RT (F (2, 34) = 87.93, p < .001, η^2_p = .838). Here, the fastest search was found in 2-conjunction conditions with a feature ratio of 1/2 and the slowest search was found in 3-conjunction condition. Finally, different distracting conditions yielded different slopes for the set size-RT functions (set size × distracting condition interaction, F (4, 68) = 28.92, p < .001, η^2_p = .628). Remarkably, removing 3-conjunction conditions from the model eliminated interaction effects while removing any of the other conditions did not. Therefore, it appears that the number of conjunctions affects slopes, rather than the feature ratio.
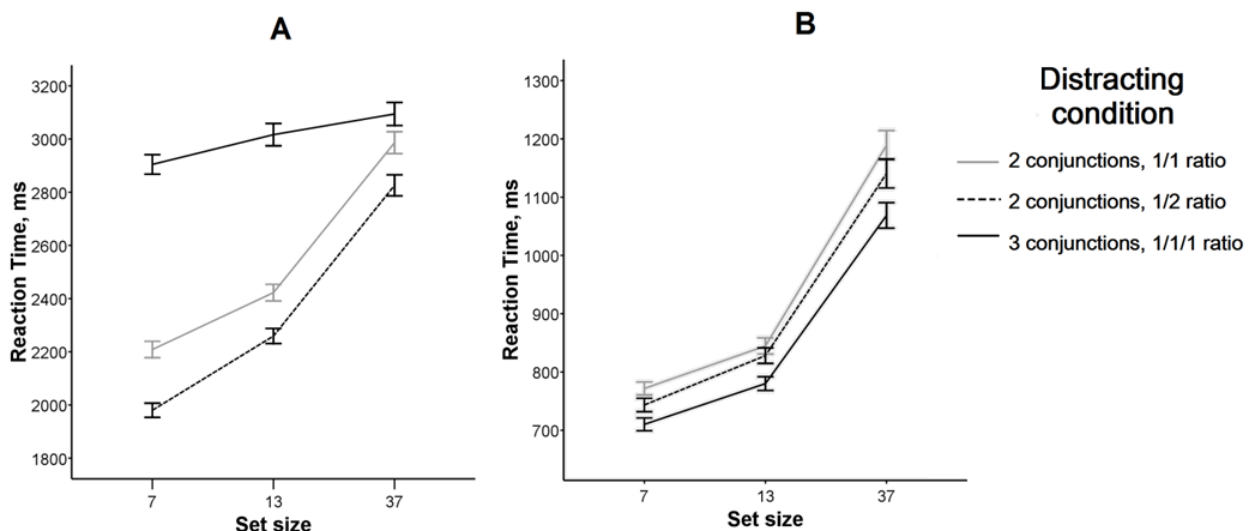

Figure 2. Results of (A) Experiment 1 and (B) Experiment 2. Error bars denote ± 1 *S.E.M.*

The results of Experiment 1 demonstrate that the distributions of both features and conjunctions have effects on visual search. However, the effects are different. Deviation from a 1/1 featureratio improved performance when the number of concurrently present distracting conjunctions was constant. This is consistent with previous findingson *distractor ratio effects* (Bacon &Egeth, 1997; Zohary& Hochstein, 1989). It appears that people can flexibly switch their attention between subsets, selecting the smallest one and thus reducing the number of items to be searched. Any dimension – be it color or orientation – can be used to guide attentionin this case. However, when there are 3 conjunctions among distractors, search generally becomesslower despite the "efficient" 1/2 feature ratio.Presumably this is due to the need to attend a larger numberof conjunctive-defined subsets. At the same time, in the 3-conjunction condition, the slope of the set size-RT function decreases substantially, indicating a rather efficient search. In accordance with the guided search framework (Wolfe, 1994), this may serve as a criterion of increasing guidance tendencies. It is unclear for a while which exact mechanism provides this kind of guidance. However, here we can see that the guidance is based on the variety of conjunctions, rather than on features.

More evidence for the guiding role of conjunctions was found in Experiment 2 (Fig. 2B). As in Experiment 1, the effect of set size was found to be significant ($F (2, 30) = 52.02$, p < .001, $\eta^2_p = .776$). The effect of distracting condition proved to be significant as well ($F (2, 30) = 20.12$, p < .001, $\eta^2_p = .573$). However, the fastest search was found in the 3-conjunction conditionthis time. Interaction between these two factors failed to reach significance, meaning that their effects are additive.

Unlike in Experiment 1, observers in Experiment 2 had a definite template that could be used for top-down guidance of attention (Wolfe, 1994). Observers were able to select a relevant subset by similarity with a template. Theoretically, this kind of selection can be guided by any single feature – such as searching among either all white or all vertical items – rather than conjunction.This feature-based strategy improves performance in the 2-conjunction condition with a 1/2 ratio in the same way as in Experiment 1. However, 3-conjunction displays include an irrelevant subset that shares no features with the target. A plausible explanation may be that this irrelevantsubset can probably be voluntarily inhibited (Treisman, 1988), thereby providing additional search advantages. The point is that this inhibition can be provided only when conjunctions are being somehow distinguished before focusing attention. On the other hand, our finding is rather consistentwith what Duncan and Humphreys (1989, 1992) thought about how template-dissimilar items are being rejected on the grounds of multidimensional difference measures. It appears, therefore, that conjunction-based inhibition contributes attentional guidance along with feature-based inhibition.

Experiments 3 and 4: Search for a double conjunction with an irrelevant size

In these experiments, observers performed the same conjunction search task as they did previously. On each trial, distractors shared either color or orientation with a target. An irrelevant dimension (size) varied between objects, providing an additional source of conjunction variability. If solely independent features guide visual search, then observers can easily select either color- or orientation-defined subsets regardless of size variation. Otherwise, if conjunctions are being analyzed as well, size can hardly be unbound from the relevant features and would be taken into account during search.

Methods

Participants. 15 (Exp. 3) and 22 (Exp. 4) undergraduate students of the Higher School of Economics took part in the experiments for extra credit in their general psychology course. All reported having normal or corrected to normal vision and no neurological problems. None of them participated in Experiments 1 and 2.

Stimuli. The stimulation was rather similar to that used in Experiments 1 and 2 regarding colors, orientations, and set sizes. There were always only two conjunction types among distractors, each sharing a feature with a potential target. An important addition was the size variation between items that could be either as large or twice as large as in Experiments 1 and 2 (Fig. 3). The character of size variation within a display defined three distracting conditions. In the *neutral* condition, all items had the same size − either small or large (Fig. 3A). In the *congruent* condition, all distractors of one type were large and all other distractors were small (Fig. 3B). In the *incongruent* condition, half of the distractors of both types werelarge and the other half was small (Fig. 3C). So, in both congruent and incongruent conditions, the proportion of different sizes remained constant, but they gave different conjunctions with colors and orientations.
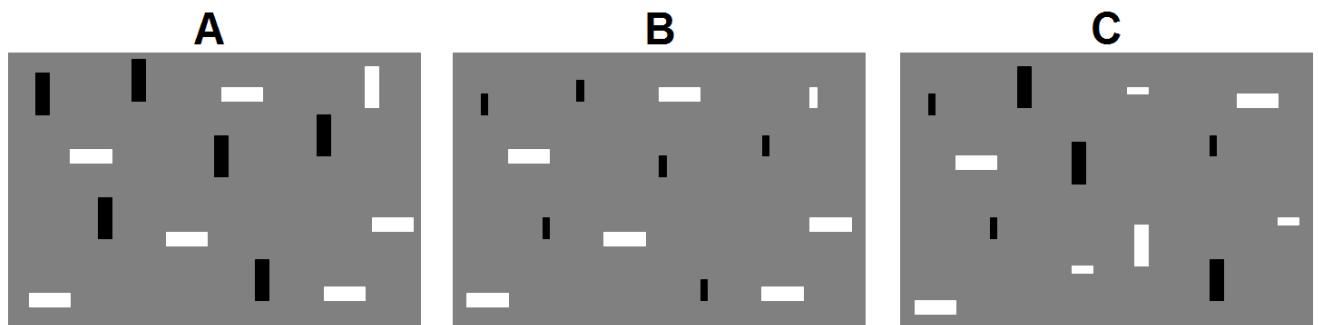

Figure 3. Three distracting conditions in Experiments 3 and 4: (A) neutral; (B) congruent; (C) incongruent.

Procedure. As in the previous pair of experiments, observers had to search for either an odd color × orientation conjunction (Experiment 3) or a white vertical line (Experiment 4). The size of the target varied in unpredictable ways from trial to trial. The design was quite similar to that used in previous experiments, although distracting conditions have now been defined by size distribution among items, rather than the number of color × orientation conjunctions.

Results and discussion

Both experiments yielded essentially the same results though, in slightly different scales (Fig. 4). The effects of set size on RT were significant (Experiment 3: $F (2, 28) = 85.08$, $p < .001$, $\eta^2_p = .859$; Experiment 4: $F (2, 42) = 61.41$, $p < .001$, $\eta^2_p = .745$), demonstrating a serial mode of search. The effects of distracting conditions on RT were also significant in both experiments (Experiment 3: $F (2, 28) = 111.08$, $p < .001$, $\eta^2_p = .888$; Experiment 4: $F (2, 42) = 88.19$, $p < .001$, $\eta^2_p = .808$), as well as the effects of set size × distracting condition (Experiment 3: $F (4, 56) = 24.70$, $p < .001$, $\eta^2_p = .636$; Experiment 4: $F (4, 84) = 26.20$, $p < .001$, $\eta^2_p = .555$), demonstrating changes in the slopes of search functions. Noteworthy, removing any of the distracting conditions did not eliminate interaction effects, indicating different slopes for all conditions. It was found that the fastest search with the least slope was being performed in the congruent condition and the slowest search with the largest slope was being performed in the incongruent condition (Fig. 4).
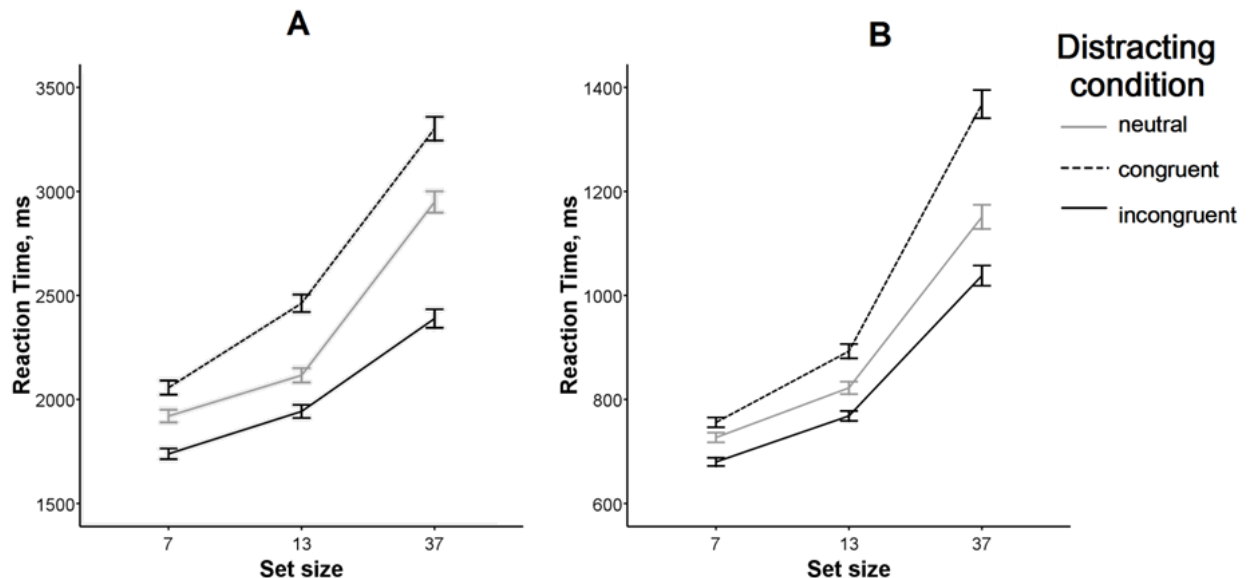

Figure 4. Results of (A) Experiment 3 and (B) Experiment 4. Error bars denote ± 1 *S.E.M.*

The principle result of both experiments is that the slopes of search functions depend on how an irrelevant feature is combined with two other features that are relevant. If the observer

relies on a feature-based guidance strategy alone, it would probably be easy to select and inspect a subset sharingany relevant feature without taking size into consideration. Then it should be equally easy to find a target within the selected subset, as it is the only feature singleton in this subset. However, it looks like the observer spends additional time to find this singleton, either because of alternate switching between sizes or because of serial selection of size-defined subsets. Whatever strategy is used, it appears that a singleton feature can hardly be searched unbound from size.

Another appropriate way of thinking about the results of these experimentsis that in terms of the classical Garnerianframework(Garner &Felfoldy, 1970). In speeded classification tasks, Garner and Felfoldy found facilitation (redundancy gain) when an irrelevant dimension co-varied with a relevant one (congruent condition),and interference when an irrelevant dimension varied orthogonally (incongruent condition).I used a similar manipulation with an irrelevant dimension, so the same framework is applicable here. In accordance with the Garnerian framework, this pattern of results refers to the concept of *integrality*. The most important notion about integrality in the context of this study is that integrality is typically observed with spatially non-separated dimensions (Garner &Felfoldy, 1970, Exp. 2) – the boundedfeatures of the same objects.

General Discussion

The main finding that can be extracted from these experiments is that visual search is sensitive to how different features are conjoined across items. I found this sensitivity for both relevant (Experiments 1 and 2) and irrelevant dimensions (Experiments 3 and 4). So, it appears that information about the variety of conjunctions is available automatically and is used in obligatory fashion, regardless of the observer's set and intentions. Finally, this information somehow affects search efficiency, or has a guiding effect, especially when one is looking for a known target (Experiment 2).

The overall conclusion that the distribution of conjunctions is being taken to account prior to focusing attention on a targetis somewhat in contradiction withstandardguidance modelsmainly relying on a preattentivemethodof subset segmentation (Treisman, 1988; Wolfe, 1994; Wolfe et al., 1989). Binding preattentive features to conjunctions is believed to be the exclusive function ofattention. Perhaps the concept of distributed attention is what could remove the theoretical contradiction. According to Treisman(2006), the broad "attentional window" can rapidly operate over the entire visual field, allowing global representation of summary properties for multiple objects.It appears that this broad attention can act rather rapidly and automatically (Ariely, 2001; Chong &Treisman, 2003), which is consistent with my findings. This broad

attention is capable of binding features of objects to some extent, but this globally bound representation (global object file) appears rather inaccurate, meaning that the observer will fail to locate any particular conjunction and estimate its proportion precisely.However, information obtained from viewing through a broad attentional window can be usefulfor the subsequent guidance of focused attention, which eventually provides a more accurate and local representation of an object.

References

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12,* 157–162.

Bacon, W.F., &Egeth, H.E. (1997). Goal-directed guidance of attention: Evidence from conjunctive visual search. *Journal of Experimental Psychology: Human Perception and Performance, 23 (4),* 948-961.

Chong, S. C., &Treisman, A. M. (2003). Representation of statistical properties. *Vision Research, 43,* 393–404.

Duncan, J., & Humphreys, G. (1989). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance, 18 (2),* 578-588

Egeth, H.E., Virzi, R.A., &Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 32-39.

Friedman-Hill, S. R, & Wolfe, J. M. (1995). Second-order parallel processing: Visual search for the odd item in a subset. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 531-551.

Garner, W.R., &Felfoldy, G.L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology, 1,* 225–241.

Halberda, J., Sires, S. F., &Feigenson, L. (2006). Multiple spatially overlapping sets can be enumerated in parallel. *Psychological Science, 17,*572-576.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279-281.

Treisman, A. M. (1988). Features and objects: Fourteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology, 40A (2),* 201-237.

Treisman, A. M. (2006). How the deployment of attention determines what we see. *Visual Cognition, 14,*411-443.

Treisman, A. M., &Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,*97-136.

Wolfe, J. M. (1994) Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1,* 202-238.

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,* 419-433.

Zohary, E., & Hochstein. S. (1989). How serial is serial processing in vision? *Perception,18,*191–200.

Igor S. Utochkin

National Research University Higher School of Economics (Moscow, Russia). Scientific-Educational Laboratory for Cognitive Research. Head of the laboratory.

E-mail: iutochkin@hse.ru, Tel. +7 (916) 486-27-51