<u>3. Modulability does not undermine the stimulus-driven nature of attentional capture</u>

by Giovanni Galfano^{*1} and Massimo Turatto²

¹Università di Padova, Padua, Italy ²Università di Trento, Trento, Italy

In the following paragraphs, we will argue that although many studies reviewed by Ruz and Lupiáñez clearly show that a top-down modulation plays an important role in determining the location to which attention will be deployed, attentional capture is stimulus-driven in nature. We agree with the authors' claim that the endogenous component is much more relevant than was hypothesised by early works in the field, but we will discuss data showing that, under some critical conditions, it is possible to observe a purely stimulusdriven attentional deployment, that is attentional capture by an item characterised as being a salient singleton in a task-irrelevant feature dimension.

Criteria for defining a purely stimulus-driven attentional capture.

In our view (Turatto & Galfano, 2001; also see Yantis, 1993), the <u>conditio sine qua non</u> for defining a purely stimulus-driven attentional capture is that the investigated feature must be completely irrelevant to the task at hand. In order to rule out any top-down effects, there are two main criteria to be met:

1) The target defining attribute(s) must be clearly and totally independent of and dissociated from the feature dimension of the irrelevant singleton.

2) In a visual search paradigm with multiple-items displays, participants must perform the task at hand by means of a Feature Search Mode (FSM); that is, their attentional set must be restricted to the target defining attribute(s). As reported by Ruz and Lupiáñez, Bacon and Egeth (1994; also see Lamy & Tsal, 1999) proposed a distinction that has proved to be very useful in clarifying some apparently contradictory results emerged in the literature. Specifically, they showed that results from previous studies taken as evidence for stimulus-driven capture, were obtained under conditions where participants

^{*} Acknoledgements: We are grateful to Kyle Cave for very helpful comments. The writing of this commentary was supported in part by grants from Ministero dell'Università e della Ricerca Scientifica e Tecnologica (MURST). Address correspondence to: Giovanni Galfano. Dipartimento di Psicologia dello Sviluppo e della Socializzazione, Via Venezia, 8. 35131 Padova, Italy. E-mail: ggalfano@unipd.it Web: www.psy.unipd.it/~galfano Telf: +39.049.8276672. Fax: +39.049.8276600

were allowed (e.g., Theeuwes, 1992) or even coerced (Pashler, 1988) to perform the experimental task by means of a Singleton Detection Mode (SDM), a 'search for the discrepant element, whatever it is' strategy. This strategy implies that participants have a broad attentional set that not only includes the target defining features, but also the singleton features. Under such conditions, any result showing that the presence of the singleton item affected performance cannot be said to show a purely stimulus-driven phenomenon, since the adoption of an SDM makes the singleton relevant somehow. One possibility to make an SDM ineffective and to force participants to adopt an FSM is to use a very inefficient visual search task (Wolfe, 1998) in which the target does not pop-out from the background, which means that all attentional resources must be focused on the target defining attributes. Reformulated in the perceptual load hypothesis (Lavie, 1995) framework, this criterion would state that a real stimulus-driven attentional capture might be said to occur under conditions of high perceptual load only.

In the next paragraph we report evidence from studies whose experimental conditions were as such as to fully conform to both the criteria discussed above.

The peril of relying on null results: Evidence that attentional capture is not always contingent.

Evidence have been discussed by Ruz and Lupiáñez that would confirm the view that not all visual properties are equally effective in capturing attention. We agree with this point as long as attentional capture is not considered as an 'all-or-none' phenomenon. We believe that the general distinction of visual features in static and dynamic (see, e.g., Folk, Remington, & Johnston, 1992) has both a theoretical and an empirical validity. In fact, many studies have shown that a task-irrelevant onset singleton is able to attract attention to its location (Yantis, 1998). As Ruz and Lupiáñez have pointed out, the fact that abrupt visual onsets receive a high attentional priority might rely on two different sub-properties characterising this kind of stimuli: an abrupt luminance change and/or a new object status (Yantis & Hillstrom, 1994). The relevant fact, in our opinion, is that whereas static discontinuities (stimuli that vary along space) such as those produced by the manipulation of colour, form, or luminance did not prove to capture attention automatically, dynamic discontinuities (stimuli that vary along both space and time) such as abrupt onsets do show a clear stimulus-driven attentional capture (e.g., Jonides & Yantis, 1988). This result by itself allows one to argue that employing the classic display-size method to assess the phenomenon, a stimulus-driven capture emerges only if the task irrelevant singletons are abrupt onsets. The observed lack of evidence for stimulus-driven attentional capture by static discontinuities, however, does not necessarily speak against the possibility that salient task-irrelevant singletons in the colour or shape feature dimensions grab attention involuntarily. What Jonides and Yantis' results say is simply that abrupt onsets receive a stronger bottom-up attentional priority (Yantis &

Johnson, 1990), clearly detectable by the display-size method. This, by itself, does not rule out the possibility that by using a more sensitive method, a stimulus-driven attentional capture might emerge even for stimuli characterised as being irrelevant static singletons. In fact, any claim based on a null result can say very little about the nature of the attentional capture phenomenon.

This idea has motivated many subsequent studies in pursuing stimulusdriven attentional capture by static discontinuities by employing different stimulus arrays from those used by Jonides and Yantis (e.g., Folk & Annett, 1994; Todd & Kramer, 1994; Yantis & Egeth, 1999), and, more importantly, different methodologies for assessing attentional capture (e.g., Gibson & Jiang, 1998). The null result relying on the lack of stimulus-driven attentional capture by static discontinuities was replicated in all the above cited studies. However, recent evidence has been reported that jeopardises the claim that static discontinuities are not able to grab attention in a purely bottom-up manner. Theeuwes and Burger (1998; also see Theeuwes, Atchley, & Kramer, 2000), using the method termed 'Identity Intrusion' based on the 'Additional singleton' paradigm (Simons, 2000), showed that, in a letter search task where participants were to decide which of two possible targets was present in a background of distractor letters, an irrelevant singleton colour distractor that was never the target affected performance producing a significant decrease in RTs when the singleton distractor identity was compatible with the target letter compared to when the singleton distractor identity was incompatible with it. This result was taken as evidence that the irrelevant colour singleton captured attention in a stimulus-driven manner. This interpretation accounts for the compatibility effect and is supported by two considerations. First, the colour singleton was never the target. Second, the letter search task proved to engage participants in an inefficient search (i.e., steep slopes in RTs plotted as a function of display size). Thus, both criteria indicated above for defining a stimulus-driven attentional capture are met. Moreover, the presence of a compatibility effect rules out the possibility that the results are accountable in terms of a filtering cost (Folk & Remington, 1998) rather than in terms of a spatial shift of attention.

In a more recent paper, we (Turatto & Galfano, 2000) have shown that a processing advantage for when the target appears at the singleton (an irrelevant unique item in colour, form, or luminance) location can be observed using a procedure we called the 'Distance method'. Our method adopts the logic of the 'Irrelevant feature search' paradigm (Simons, 2000), and has the advantage of requiring a single display size, with a fixed number of elements. In our study, participants were looking for the presence or absence of a vertical line segment (the target) embedded among variously tilted distractor segments. Line segments were presented inside shapes, whose features were manipulated to create the irrelevant singleton (also see Theeuwes, 1992). The target, when present, appeared at the singleton location in 1/n of target present trials (where n is display-size). Participants were more accurate in finding the target when it appeared inside the singleton element than when it appeared in the more accurate of the non singleton locations. As target defining attributes

forming participants' attentional set concerned orientation (and not colour, shape, or luminance), we clearly met the first criterion for defining a stimulusdriven capture. The steep slopes observed in a similar search task with no singleton in which we varied display-size, ensured us that the task was very resource-demanding (i.e., a non-efficient search) thus allowing for reasonably ruling out the possibility that participants performed the task by means of an SDM. In another paper (Turatto & Galfano, 2001) where colour was the only investigated feature, we have replicated this finding with a T-L task (leading to a non efficient search: Wolfe, 1998) and RTs as the main dependent variable. Crucially, we have included a no-singleton condition mixed with singleton trials in the same block, and we have observed that participants were faster in finding the target when it was at the singleton location than the no-singleton condition. This result favours an interpretation in terms of spatial capture and definitely rules out the possibility of explaining the singleton-advantage pattern as a filtering cost (see Turatto & Galfano, 2001).

Ruz and Lupiáñez argue that our results may be interpreted as evidence of an attentional misguidance effect (see Todd & Kramer, 1994) rather than as evidence of stimulus-driven attentional capture. The attentional misguidance hypothesis applies to the 'Irrelevant feature search' paradigm and claims that participants may decide to start the search from the singleton location, which would act as a sort of landmark. In this view, any benefit shown by the singleton would be the consequence of the adoption of a specific search strategy, and therefore could be accounted for by invoking the involvement of a top-down component. We do not agree with this interpretation of our data on the basis of the fact that our experimental procedure fully met the criteria for defining a purely stimulus-driven attentional capture. Moreover, in a more recent paper (Turatto, Galfano, Gardini, & Mascetti, 2001), we have provided evidence that rules out this alternative explanation of our previous data. In fact, by comparing the sensitivity of the classic display-size method against that of our distance method, we obtained evidence for stimulus-driven attentional capture by colour in the distance method only, thus replicating the findings of Jonides and Yantis (1988). This result not only shows that the display-size method is less powerful than the distance method in detecting attentional capture, but also rules out any alternative account of our previous findings (Turatto & Galfano, 2000; 2001) in terms of attentional misguidance. In fact, as this hypothesis assumes that our participants strategically decided to begin the search from the salient item, one would have expected the attentional capture pattern to emerge in both the display-size method and the distance method. Because only the distance method was able to show evidence for such phenomenon, we can safely assume that the singleton was effectively not treated by participants as a landmark from which to start the visual search task. This means that no relevant top-down component affected the results, and that stimulus-driven control was therefore present in isolation.

In a rather different perspective from that of the typical paradigms used in the attentional capture literature, Stolz (1996) has tested the widely accepted implicit view that exogenous orienting represents an encapsulated system impervious to high-level cognitive influence. According to Posner (e.g.,

Posner & Petersen, 1990), there are three operations involved in orienting of attention: engage, move and disengage. In the terminology used by the spotlight metaphor, the engage operation would correspond to attentional capture. Stolz (1996, Experiments 2 and 5) used a variant of the classical spatial cueing paradigm where she had a word at the fixation location and a word acting as the non-informative cue for the target location. Participants were required to perform a discrimination task on the target. The crucial manipulation was that the cue was semantically related to the fixation word on half the trials and unrelated to it in the other half. Besides the classic validity effect, the results showed that the semantic relationship between the fixation word and the cue impaired performance in invalid trials only. This implies that the semantic status of the cue influenced its ability to hold attention at the cued location. The interference effect has been interpreted to take place at either the move or disengage stages and, crucially, not at the engage stage, as no effect of semantic relatedness was detected in valid trials. In sum, this pattern of results is consistent with the notion that, on valid trials, the cue captured attention and that attentional capture is impervious to high-level influences (see Pylyshyn, 1999 for a detailed discussion concerning the impenetrability of perceptual processes determining where attention will be deployed).

Evidence supporting this position also comes from the results of a study by Theeuwes et al. (2000; also see Theeuwes & Godijn, 2001), who were interested in investigating the time course of stimulus-driven and goaldirected processes in controlling visual attention. By varying the SOA between the onset of an irrelevant singleton and the onset of the target, these researchers showed that top-down control is able to suppress stimulus-driven capture only if participants are given enough time before target presentation (also see Kim & Cave, 1999; 2001). Specifically, their findings are consistent with the notion that top-down control does not operate at the level of early visual processing and does not literally override bottom-up capture, but works speeding up the operation of disengaging attention from the irrelevant singleton location instead. This interpretation would also disentangle the reason for why Folk et al. (1992) could not find evidence for stimulus-driven capture either by a colour or an abrupt onset singleton using the spatial cueing paradigm. In fact, in those experiments, the SOA between the onset of the irrelevant cue and the onset of the search display was 150 ms. As noted by Theeuwes et al. (2000), by the time the search array was presented in the Folk et al. study, participants may have been able to disengage spatial attention from the location of the cue and to move the attentional focus over the location of the target. This interpretation is also consistent with the finding of a stimulus-driven attentional capture in our previous studies (Turatto & Galfano, 2000; 2001; Turatto et al., 2001), where a 0-ms SOA between the onset of the irrelevant singleton and the onset of the target was employed. It is important to note that the study of Theeuwes et al. (2000; also see Kim & Cave, 1999) is susceptible of an SDM criticism. In fact, the procedure employed did not ensure that participants used an FSM to find the target, since the target always appeared inside a relevant shape singleton (as in Theeuwes, 1992). However, the modulation of attentional capture shown along

the different SOAs, allows one to assume that the bottom-up activation produced by the 'irrelevant' colour singleton was stronger than the activation of the top-down component produced by the template matching process.

The importance of saliency in visual behaviour.

Further contributions attesting the relevance of saliency and bottom-up processes in human behaviour have been provided from other research fields, such as neurophysiology or psychophysics. For example, Constantinidis and Steinmetz (2001) recorded activity from single neurons in area 7a of the Posterior Parietal Cortex, known to play a crucial role in orienting of spatial attention and observed that, in a spatial multiple-stimulus display version of the match-to-sample task, responses to salient odd-coloured targets were enhanced. Clearly, this result does not represent evidence for a stimulusdriven capture by colour, as colour was the target defining attribute. However, the pattern emerged testifies the importance of saliency in controlling visual processing. The relevance of saliency has also emerged in a series of psychophysical studies conducted by Nothdurft (e.g., 1993; 2000a; 2000b). For example, Nothdurft (2000a) showed that performance in responding to a singleton target was significantly enhanced when the target was a singleton in two feature dimensions compared to when it was a singleton in only one feature. As in the case of the Constantinidis and Steinmetz (2001) study, this result is no way evidence for stimulus-driven capture, but shows that saliency per se relevantly affects visual performance. Evidence that bottom-up processes are important in visual processing has also been provided both in conjunction searches (Sobel & Cave, in press) and in the context of a recently discovered phenomenon strongly related to the attentional capture research field, that is Change Blindness (Simons, 2000). This latter effect refers to the observation that repeated changes between scenes often go unnoticed for a surprisingly long time (see, e.g., Turatto, Angrilli, Mazza, Umiltà, & Driver, 2002). In particular, Scholl (2000) has reported evidence that salient items, even when they are uncorrelated with the loci of change, show a dramatic decrease in change blindness. This pattern of results has been interpreted as evidence that the salient elements (late-onset items and colour singletons) captured attention in a stimulus-driven fashion therefore producing an attenuated change blindness effect.

Finally, many computational models of visual attention make the clear prediction that a salient, although task-irrelevant singleton, would draw attention on its location (e.g., Cave, 1999; Cave & Wolfe, 1990; Itti & Koch, 2000; Koch & Ullman, 1985; Li, 2002; Niebur & Koch, 1998; Parkhurst, Law, & Niebur, 2002; Wolfe, 1994). The recent evidence showing purely stimulus-driven attentional capture give a critical empirical support to these models.

Conclusion

In sum, we do not put into question the fact that endogenous factors play a crucial role in determining attentional control. Our view is rather that, saliency is also relevant and if early studies were unable to show a purely stimulus-driven capture of attention by static discontinuities (i.e., discontinuities in feature dimensions whose ecological priority is less prominent than that of abrupt visual onsets, Yantis & Johnson, 1990), this does not mean that the bottom-up component plays no role in controlling the deployment of visual attention. In this commentary, based on previous proposals (e.g., Yantis, 1993; Bacon & Egeth, 1994), we have suggested some criteria to investigate stimulus-driven attentional control in isolation. As a matter of fact, evidence from our and other labs (e.g., Turatto & Galfano, 2000; 2001; Turatto et al., 2001; Kim & Cave, 2001; Scholl, 2000; Theeuwes & Burger, 1998) has appeared in the literature that shows that stimulus-driven attentional capture by static singletons does indeed take place. We emphasise that this growing evidence has been collected by carefully conforming to both the criteria discussed above.

The null effect reported by previous studies is attributable either to the low sensitivity of the employed experimental paradigms (e.g., Jonides & Yantis, 1988) or to the adoption of the inappropriate temporal parameters (e.g., Folk et al., 1992), whereas the debate concerning the earlier evidence for the phenomenon (see e.g., Bacon & Egeth, 1994; Lamy & Tsal, 1999; Pashler, 1988; Theeuwes, 1992) raised from the difficulty in studying the bottom-up component in isolation.

REFERENCES

- Bacon, W. M., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. Perception and Psychophysics, 55, 485-496.
- Cave, K. R. (1999). The FeatureGate model of visual selection. <u>Psychological Research</u>, <u>62</u>, 182-194.
- Cave, K. R., & Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search. <u>Cognitive Psychology</u>, 22, 225-271.
- Constantinidis, C., & Steinmetz, M. A. (2001). Neuronal responses in area 7a to multiplestimulus displays: 1. Neurons encode the location of the salient stimulus. <u>Cerebral</u> <u>Cortex</u>, 11, 581-591.
- Folk, C. L., & Annett, S. (1994). Do locally defined feature discontinuities capture attention? <u>Perception and Psychophysics</u>, 56, 277-287.
- Folk, C. L., & Remington, R. W. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. <u>Journal of Experimental</u> <u>Psychology: Human Perception and Performance</u>, 24, 847-858.
- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. <u>Journal of Experimental Psychology:</u> <u>Human Perception and Performance</u>, <u>18</u>, 1030-1044.
- Gibson, B. S., & Jiang, Y. (1998). Surprise! An unexpected color singleton does not capture attention in visual search. <u>Psychological Science</u>, <u>9</u>, 176-182.

- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. <u>Vision Research</u>, <u>40</u>, 1489-1506.
- Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. <u>Perception and Psychophysics</u>, 43, 346-354.
- Kim, M.-S., & Cave, K. R. (1999). Top-down and bottom-up attentional control: On the nature of interference from a salient distractor. <u>Perception and Psychophysics</u>, <u>61</u>, 1009-1023.
- Kim, M.-S., & Cave, K. R. (2001). Perceptual grouping via spatial selection in a focusedattention task. <u>Vision Research</u>, <u>41</u>, 611-624.
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. <u>Human Neurobiology</u>, <u>4</u>, 219-227.
- Lamy, D., & Tsal, Y. (1999). A salient distractor does not disrupt conjunction search. <u>Psychonomic Bulletin and Review</u>, <u>6</u>, 93-98.
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. <u>Journal of</u> <u>Experimental Psychology: Human Perception and Performance</u>, <u>21</u>, 451-468.
- Li, Z. P. (2002). A saliency map in primary visual cortex. <u>Trends in Cognitive Sciences</u>, <u>6</u>, 9-16.
- Niebur, E., & Koch, C. (1998). Computational architectures for attention. In R. Parasuraman (Ed.) <u>The attentive brain</u> (pp. 163-186). Cambridge, MA: MIT Press.
- Nothdurft, H.-C. (1993). Saliency effects across dimensions in visual search. <u>Vision</u> <u>Research</u>, <u>33</u>, 839-844.
- Nothdurft, H.-C. (2000a). Salience from feature contrast: Additivity across dimensions. Vision Research, 40, 1183-1201.
- Nothdurft, H.-C. (2000b). Salience from feature contrast: Variations with texture density. <u>Vision Research</u>, <u>40</u>, 3181-3200.
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. <u>Vision Research</u>, <u>42</u>, 107-123.
- Pashler, H. (1988). Cross-dimensional interaction and texture segregation. <u>Perception and</u> <u>Psychophysics</u>, <u>43</u>, 307-318.
- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. <u>Annual</u> <u>Review of Neuroscience</u>, <u>13</u>, 25-42.
- Pylyshyn, Z. (1999). Is vision contiguous with cognition? The case for cognitive impenetrability of visual perception. <u>Behavioral and Brain Sciences</u>, 22, 341-365.
- Ruz, M., & Lupiáñez, J. (2002). Attentional capture and exogenous orienting: Upon their automaticity and sensitivity to endogenous control. <u>Psicologica</u>. This issue.
- Scholl, B. J. (2000). Attenuated change blindness for exogenously attended items in a flicker paradigm. <u>Visual Cognition</u>, 7, 377-396.
- Simons, D. J. (2000). Attentional capture and inattentional blindness. <u>Trends in Cognitive</u> <u>Sciences</u>, <u>4</u>, 147-155.
- Sobel, K. V., & Cave, K. R. (2002). The roles of salience and strategy in conjunction search. In press.
- Stolz, J. A. (1996). Exogenous orienting does not reflect an encapsulated set of processes. <u>Journal of Experimental Psychology: Human Perception and Performance</u>, <u>22</u>, 187-201.
- Theeuwes, J. (1992). Perceptual selectivity for color and form. <u>Perception and</u> <u>Psychophysics</u>, <u>51</u>, 599-606.
- Theeuwes, J., & Burger, R. (1998). Attentional control during visual search: The effect of irrelevant singletons. Journal of Experimental Psychology: Human Perception and <u>Performance</u>, 24, 1342-1353.

- Theeuwes, J., & Godijn, R. (2001). Attentional and oculomotor capture. In C. L. Folk & B. S. Gibson (Eds.), <u>Attraction, distraction and action: Multiple perspectives on</u> <u>attentional capture</u> (pp. 121-149). Amsterdam: North-Holland.
- Theeuwes, J., Atchley, P., & Kramer, A. F. (2000). On the time course of top-down and bottom-up control of visual attention. In S. Monsell & J. Driver (Eds.), <u>Attention</u> <u>and Performance XVIII</u> (pp. 105-125). Cambridge: MIT Press.
- Todd, S., & Kramer, A. F. (1994). Attentional misguidance in visual search. <u>Perception</u> <u>and Psychophysics</u>, <u>56</u>, 198-210.
- Turatto, M., & Galfano, G. (2000). Color, form, and luminance capture attention in visual search. <u>Vision Research</u>, <u>40</u>, 1639-1643.
- Turatto, M., & Galfano, G. (2001). Attentional capture by color without any relevant attentional set. <u>Perception and Psychophysics</u>, <u>63</u>, 286-297.
- Turatto, M., Galfano, G., Gardini, S., & Mascetti, G. G. (2001). Stimulus-driven attentional capture: An empirical comparison of display-size and distance methods. Submitted
- Turatto, M., Angrilli, A., Mazza, V., Umiltà, C., & Driver, J. (2002). Looking without seeing the background change: Electrophysiological correlates of change detection versus change blindness. <u>Cognition</u>, <u>84</u>, B1-B10.
- Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. <u>Psychonomic Bulletin and Review</u>, 1, 202-238.
- Wolfe, J. M. (1998). Visual Search. In H. Pashler (Ed.), <u>Attention</u> (pp. 13-73). Hove, UK: Psychology Press.
- Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. <u>Journal of Experimental Psychology: Human Perception and Performance</u>, <u>19</u>, 676-681.
- Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), <u>Attention</u> (pp. 223-256). Hove, UK: Psychology Press.
- Yantis, S., & Egeth, H. E. (1999). On the distinction between visual salience and stimulus-driven attentional capture. Journal of Experimental Psychology: Human Perception and Performance, 25, 661-676.
- Yantis, S., & Hillstrom, A. P. (1994). Stimulus-driven attentional capture: Evidence from equiluminant visual objects. <u>Journal of Experimental Psychology: Human</u> <u>Perception and Performance</u>, 20, 95-107.
- Yantis, S., & Johnson, D. N. (1990). Mechanisms of attentional priority. <u>Journal of Experimental Psychology: Human Perception and Performance</u>, <u>16</u>, 812-825.