CrossMark

SHORT REPORT

A response time cost when different representations of an attended object are integrated

Nicholas Hon[1]

Abstract When selected, attention is thought to spread across the whole of an object. Such spreading is thought to occur via the integration and mutual enhancement of the different mental representations of said object. Neurophysiological studies have demonstrated that such integration is not instantaneous with selection, but rather occurs after some delay. It is currently unclear whether the time needed for integration to be established has a behavioural consequence. Here, it was found that trials that required integration were responded to more slowly than those that did not, even though correct responses in both could be determined by the same information. These data thus suggest that the time taken for integration between representations to be established has a consequence on behaviour, one that can be observed as a response time cost. Furthermore, these findings enlighten the time-course of integration vis-a-vis the processing of information at different levels of the processing stream.

Keywords Attention: object-based · Attention: neural mechanisms · Attention: selective

Introduction

Attending to an object bestows it with several advantages: attended objects are processed and remembered better than unattended counterparts (Posner, 1980; Rock & Gutman, 1981). Furthermore, studies have also demonstrated that

making decisions about two attributes of an attended object are no worse than making a single decision about a unitary attribute of the object (Duncan, 1984), and that two decisions made about a single object are more accurate than separate single decisions made about two different objects (Baylis & Driver, 1993; Duncan, 1984). The observation of this *same object advantage* led to the idea that, when selected, attention spreads across the whole of an object. This attentional spreading is thought to occur through the integration and mutual enhancement of the mental (and neural) representations of the object and its various attributes (Desimone & Duncan, 1995; Duncan, Humphreys, & Ward, 1997; Roelfsema, 2006). Such integration and mutual enhancement acts to ensure that system-wide processing is dominated by the attended object, allowing it to compete successfully against unselected counterparts. Findings from neurophysiological studies support this proposal: when an object is attended to, enhanced activity is observed in the various brain regions that code for or represent the different properties of that object (O'Craven, Downing, & Kanwisher, 1999; Schoenfeld et al., 2003), even those not determining selection (Schoenfeld, Hopf, Merkel, Heinze, & Hillyard, 2014).

It is worth noting that the representations of an object's various properties can be housed at different levels of the processing stream. For example, object-level representations are housed later in the processing stream than representations of basic visual features like size (Grill-Spector et al., 1999; Nishimura, Scherf, Zachariou, Tarr, & Behrmann, 2015). Nonetheless, integration can occur between distantly held representations. Of relevance here, such integrative activity is not instantaneous with selection. For instance, in early visual areas (i.e. those that code for basic visual properties), evidence of integration is observed only at delayed latencies, well after the initial response triggered by the appearance of the stimulus (Haenny, Maunsell, & Schiller, 1988; Motter, 1994). This is

✉ Nicholas Hon
  psyhonn@nus.edu.sg

[1] Department of Psychology, National University of Singapore, 9 Arts Link, Singapore 117570, Singapore

thought to reflect the action of feedback signals returning from upstream sites (Chelazzi, Miller, Duncan, & Desimone, 2001; Hon, Thompson, Sigala, & Duncan, 2009; Roelfsema, Lamme, & Spekreijse, 1998), allowing for integration between representations housed earlier and later in the processing stream.

To date, it is unclear whether the time needed for integration of representations held earlier and later in the processing stream to be established has any consequence on behaviour. For example, it is uncertain whether such early–late integration will affect the speed with which earlier represented information can be used for behavior. One way of studying this is to compare response times on trials in which correct responses can be made on the basis of earlier coded information under two contrasting conditions - when early–late integration is necessary and when it is not. Following this logic, participants in this experiment performed the following task. They viewed pairs of sequentially-presented images (of objects) with the objective of determining whether the second image was different from the first. When a change between the two occurred, the second image could be a differently sized version of the first (*Size* change), a completely different object (*Object* change) or a differently sized new object (*Combination* change). Combination and Size changes were of central interest here. Because these trials would not require a new object to be attended to, Size changes would merely entail an updating of values in the representation of a currently held object's size[1]. On the other hand, Combination changes would, in addition to requiring processing of new size information, also require the attending to, and processing of, a new object (Kahneman, Treisman, & Gibbs, 1992; Treisman, Kahneman, & Burkell, 1983). Thus, on Combination trials, new size information would have to be integrated with a new object-level representation. Notice, though, that correct responses in both Size and Combination trials can be determined by size information, which, being an early processed feature, would also ensure the speediest responses. As such, a difference in response times (RTs) between these two trial types can be taken as an indication of the effect that the time needed to establish integration has on behaviour.

Methods

Participants

Twenty-three participants with normal or corrected-to-normal vision took part in this experiment.

---

[1] For all trial types, an initial integration between a size and an object representation would happen as a consequence of attention being paid to the first image. However, on Size trials, because there is no new object to attend to in the second image, the new size information would merely require an already integrated size representation to be updated.

Procedure

In this experiment, participants viewed pairs of sequentially-presented images with the objective of determining whether there was a change between the first and second image. The first image of a pair was shown for 1000 ms, followed by a 200-ms blank frame, and then by a second image, which remained onscreen until a response was made. Following the offset of the second image, a fixation sign appeared on the screen for 500 ms, followed by the first image of the next pair and so on. The images comprised line-drawings of real world objects sampled from the Snodgrass and Vanderwart set (Snodgrass & Vanderwart, 1980). A total of 144 images were drawn from this set for this study. The images were presented in black and centered on a white background. The first image of the pair was constructed to fit as closely as possible to the borders of an invisible bounding box that subtended approximately 4.6° × 4.7° of visual angle. On half the trials, the second image was a repetition of the first (*no-change*). On the remaining trials, the second image differed from the first in one of three ways. When a change occurred, the second image could be an enlarged version of the first image (*Size change*), a completely new object (*Object change*) or a new object that was also larger in size than the first (*Combination change*). When a change in size occurred, as in the Size and Combination trials, the second image was always twice the size of the first. Because of the limited number of suitable images available, each pair of stimuli was presented twice and always in the same condition. There were a total of 108 no-change trials and 108 change trials (36 trials in each change condition). Participants made button-presses, on a custom response box, with their index finger to indicate a change and with their middle finger to indicate that the images were the same. Trial order and stimulus-condition assignment were randomized for each participant. Participants were informed of the different types of changes and shown examples of these before the experiment started. Participants were instructed to respond as quickly as possible, but without sacrificing accuracy.

Results

The accuracy data for all change types are presented in Table 1. All trial types were responded to with a high degree of accuracy, with no difference between them [$F(3,66) = 1.47$, $P = .230$].

Figure 1 presents the RT correct data for the different trial types. A one-way ANOVA conducted on these revealed a clear effect of trial type, $F(3,66) = 31.00$, $P < .001$. Planned comparisons revealed that all change types were responded to more slowly than the no-change controls (all *P* values ≤ .001, paired samples *t*-tests). When the change trials were compared
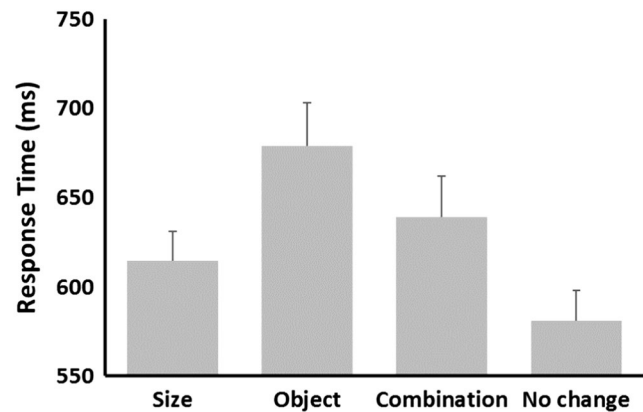
Table 1 Accuracy data for the different change trials presented as percentages

Trial Type	Accuracy (%)
Size	96.9
Object	96.0
Combination	97.2
No Change	97.7

against each other, it was found that Size changes were responded to more quickly than Object changes [$t(22) = 5.86$, $P < .001$], consistent with the idea that size information is processed very early in the processing stream, and, therefore, can be accessed for behaviour more quickly than later-processed object information (Larsen & Bundesen, 1978; Larsen, Bundesen, Kyllingsbaek, Paulson, & Law, 2000; Nishimura et al., 2015). More critically, it was found that Combination changes were responded to more slowly than Size changes [$t(22) = 2.39$, $P = .026$], but more quickly than Object changes [$t(22) = 4.74$, $P < .001$]. This suggests two points. First, responses on Combination changes, although these entailed the presentation of new objects, were unlikely to have been dictated solely by object-level processing; if that were the case, Combination trial RTs would have been identical to those on Object trials. Second, and more importantly, Combination trial responses were not driven solely by size information, which would have guaranteed the fastest responses. Rather, the data suggest that, on Combination trials, some level of integration between new size information and new object information is likely to have occurred before responses were triggered[2].

To better probe this possibility, we examined the relationship between the extra time needed to respond to Object changes over and above that needed to respond to Size changes (i.e., $RT_{Object} - RT_{Size}$) and Combination trial RTs (Fig. 2). The logic of this analysis is that, for integration between object and size representations to occur, some level of object processing must be performed before feedback signals can be sent back to earlier coded size representations. Thus, it would be

Fig. 1 Reaction time (RT) data for the different change trials. Error bars indicate 1 SEM

expected that participants who take longer to process object information (as indexed by how long they take on Object trials), relative to the processing of size information, would also take longer on Combination trials because these would require integration between new size and object information. Consistent with this, Object–Size difference was found to correlate positively with RTs on Combination trials [$r = .72$, $P < .001$][3].

Discussion

Integration between the different representations of an object and its attributes is typically associated with behavioural benefits like the *same object advantage*. Neurophysiological studies, however, have demonstrated that such integration takes time to occur. The current data provide evidence of a behavioural consequence of this: trials that required integration (Combination) were responded to more slowly than trials that did not (Size), even though the same information (in this case, size information) would have produced the correct and quickest responses in both cases.

These results also cast light on the time-course of integration. If integration between object and size representations can occur only after object processing is completed, then RTs on Combination trials would have been similar to those on Object trials. This was not what was found. Combination RTs were between Size and Object RTs, suggesting that integration is established before object-level processing is completed. At the same time, the data also suggest that object-level processing is initiated before size processing is done: if object processing can start only after size processing has been completed, then there ought to be no impediment to responding purely on the basis of size, with Combination trial RTs being as swift as

---

[2] This overall pattern of results was replicated in another experiment (same task and presentation parameters as in the main experiment) that utilized word stimuli instead of line drawings. Thus, there were Word changes, Size changes and Combination changes. In general, words contain two types of distinctive information, pertaining to visual word form and semantics, both of which are known to be processed later in the processing stream than size information. In this experiment, a one-way ANOVA revealed a significant effect of change type, F(2, 98) = 9.41, P < .001. Paired t-tests revealed the same pattern as reported in the main experiment above: word changes took longer than both size changes (P < .001) and combination changes (P = .013). Critically, Size changes were faster than Combination changes (P = .051). Also similar to the main analysis, Word-Size differences correlated positively with Combination RTs (r = .614, P < .001).

[3] As recommended by an anonymous reviewer, the correlation between Combination-Size differences and Object RTs was also run. There was also a positive relationship between these [r = .61, P < .001].
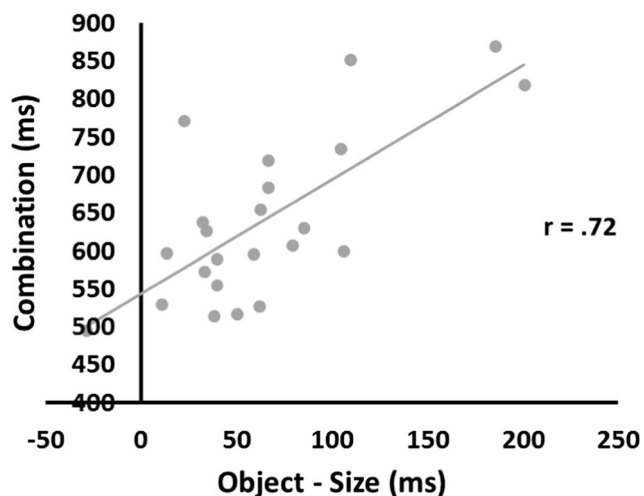
Fig. 2 Relationship between the difference in RTs on Object and Size trials (i.e. RT_{Object}–RT_{Size}) and Combination trial RTs. Each *dot* represents the data from a single participant

those on Size trials. Again, this was not found. Taken together, the current results suggest the following account regarding integration: feedforward signals from early visual representations are sent to object-level ones before early visual processing is completed. This allows object-level processing to begin before earlier processes are done. Of greater relevance here, feedback signals from object representations that enable integration with earlier visual representations are sent on the basis of partially completed object processing. Once integration has been established, information housed in early visual representations are free to guide responding. A central point here, then, is that integration can be initiated on the basis of partial processing. That is, feedforward signals can be sent on the basis of partially processed early visual information, and feedback signals can be sent back on the basis of partially processed higher-order visual information. It is worth noting that exchanges between different levels of the processing stream on the basis of partially processed information is not unknown in cognition. For example, this is a mainstay feature of interactive activation models (e.g., McClelland & Rumelhart, 1981).
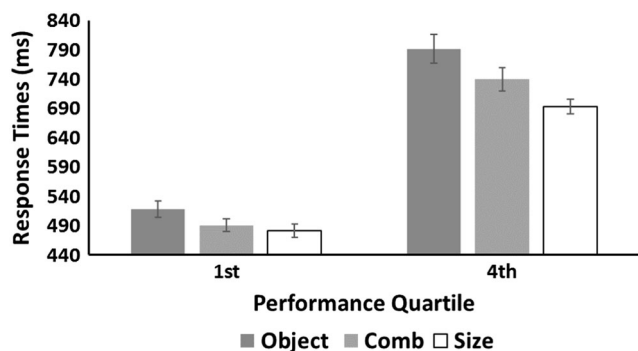

Fig. 3 RTs from the first (fastest trials) and fourth (slowest trials) performance quartiles. Error bars indicate 1 SEM

The data from this study suggest that such a mechanism also operates within the context of object-based attention.

A further point suggested by these data is that, on Combination trials, information held in individual representations was not accessed for behaviour before integration was established. Rather, information in individual representations was used for behaviour only once some level of cross-representational integration occurred. (Information in size representations could have been used fruitfully for correct responding in both Size and Combination trials, but the latter, requiring object-size integration, took longer to respond to.) This suggests the broader idea that the establishment of integration may supercede the use of information in individual representations, at least when one type of information is not emphasized over others, as was the case here.

An intuitively appealing alternative interpretation is that Combination trial performance stems from interference between two different types of information (i.e. size and object-level). This is unlikely for two reasons. First, the two types of information would not have been in conflict with each other, since both point to the same, correct, solution. Second, the data are, in fact, inconsistent with this idea. If the interference idea were correct, then one would expect interference to be greatest when object-level information is processed fastest (since more speedily-processed information should provide greater interference). The exact opposite was observed here—Combination RTs were longest for participants who took longer to process new object information (relative to new size information).

At the same time, these data are inconsistent with strategy-based accounts. To begin with, it is unlikely that participants, although not instructed to, strategically opted to emphasise one type of information over the other. If half the participants emphasized size information and half emphasized object information, the pattern of results reported earlier would, in fact, be obtained, but only in a numerical sense. Critically, neither of the comparisons between Combination trials and the other two change types would be significant. Similarly, no relationship between Object-Size differences and Combination RTs would have been found. As such, we can rule out this alternative.

A variation of the preceding idea is that, on Combination trials, participants might have vacillated between emphasizing size or object information. If this were the case, then, when RTs from the different conditions are ordered from fastest to slowest, one would predict that (1) Combination and Size RTs would be similar when considering the fastest trials, and (2) Combination and Object RTs would be similar when considering the slowest. To test this, the trials from each condition were ordered from fastest to slowest and then partitioned into performance quartiles. These recast data were entered into a 2 (quartile: 1st vs 4th) x 3 (change type: size, object,

combination) ANOVA (Fig. 3)[4]. This analysis revealed significant effects of quartile [F(1, 7) = 762.7, $P < .001$] and change type [F(2, 14) = 90.99, $P < .001$], along with an ordinal interaction between them, F(2, 14) = 30.13, $P < .001$. This suggests that the same pattern of results was observed at both ends of the performance distribution. Critically, paired samples t-tests revealed differences between the fastest quartile Combination and Size trials ($P < .001$) and between the slowest quartile Combination and Object trials ($P < .001$). This is contrary to the predictions of the vacillation hypothesis. Extending this, a significant difference was found between the fastest Combination quartile and the slowest Size quartile ($P < .001$), suggesting that fast Combination RTs were not simply slow responses to size information. Similarly, there was a significant difference between the slowest Combination quartile and the fastest Object quartile ($P < .001$), suggesting that slow Combination RTs were not merely fast responses to object information. Taken together, these findings suggest that the vacillation hypothesis is an unlikely account for the main findings.

Perhaps prompted by the knowledge that Combination trials entailed both new size and object information, participants adopted a strategy of *always* waiting for some level of object information (whether this indicates a new object or not) before making their responses, regardless of trial type. However, such a strategy would inflate Size and Combination RTs (since both these trial types can be responded to very quickly on the basis of size information alone), without affecting Object trial RTs (object processing is a requisite for correct responding on such trials anyway). In such a case, one would predict a negative relationship between Object-Size differences and Combination RTs (i.e. smaller Object-Size differences associated with longer Combination RTs)[5]. Here, the opposite pattern was found. Overall, the current data are consistent with the integration-related account.

In conclusion, these data suggest that, while producing behavioural benefits like the same object advantage and ensuring that attended objects dominate overall processing, integration

is not a cost-free enterprise. Given its time-consuming nature, integration, when it is necessary, can affect the speed of responses to attended objects.

References

Baylis, G. C., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance, 19*(3), 451–470.

Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (2001). Responses of neurons in macaque area V4 during memory-guided visual search. *Cerebral Cortex, 11*(8), 761–772.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Revue of Neuroscience, 18,* 193–222.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*(4), 501–517.

Duncan, J., Humphreys, G., & Ward, R. (1997). Competitive brain activity in visual attention. *Current Opinion in Neurobiology, 7*(2), 255–261.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron, 24*(1), 187–203.

Haenny, P. E., Maunsell, J. H., & Schiller, P. H. (1988). State dependent activity in monkey visual cortex. II. Retinal and extraretinal factors in V4. *Experimental Brain Research, 69*(2), 245–259.

Hon, N., Thompson, R., Sigala, N., & Duncan, J. (2009). Evidence for long-range feedback in target detection: Detection of semantic targets modulates activity in early visual areas. *Neuropsychologia, 47*(7), 1721–1727.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*(2), 175–219.

Larsen, A., & Bundesen, C. (1978). Size scaling in visual pattern recognition. *Journal of Experimental Psychology: Human Perception and Performance, 4*(1), 1–20.

Larsen, A., Bundesen, C., Kyllingsbaek, S., Paulson, O. B., & Law, I. (2000). Brain activation during mental transformation of size. *Journal of Cognitive Neuroscience, 12*(5), 763–774.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: 1. An account of basic findings. *Psychological Review, 88,* 375–407.

Motter, B. C. (1994). Neural correlates of attentive selection for color or luminance in extrastriate area V4. *Journal of Neuroscience, 14*(4), 2178–2189.

Nishimura, M., Scherf, K. S., Zachariou, V., Tarr, M. J., & Behrmann, M. (2015). Size precedes view: Developmental emergence of invariant object representations in lateral occipital complex. *Journal of Cognitive Neuroscience, 27*(3), 474–491.

O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature, 401*(6753), 584–587.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Rock, I., & Gutman, D. (1981). The effect of inattention on form perception. *Journal of Experimental Psychology: Human Perception and Performance, 7*(2), 275–285.

---

[4] For this analysis, we utilised only the first 32 trials (when ordered from fastest to slowest) from each condition. The reason for this is that all participants responded correctly to at least 32 (out of a total of 36) trials in all conditions. This ensured that each performance quartile comprised the same number of trials.

[5] Assuming the use of this strategy, there are two broad scenarios. First, a participant can wait for very little object information. Such a participant would have fairly short Size and Combination RTs. This would result in the participant having a large Object-Size difference and short Combination RTs. Conversely, a participant who waits for a lot of object information would have long Size and Combination RTs, resulting in a small Object-Size difference coupled with long Combination RTs. As one can see, both cases would result in a negative relationship between Object-Size differences and Combination RTs, which is not the pattern obtained.

Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience, 29,* 203–227.

Roelfsema, P. R., Lamme, V. A., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature, 395*(6700), 376–381.

Schoenfeld, M. A., Hopf, J. M., Merkel, C., Heinze, H. J., & Hillyard, S. A. (2014). Object-based attention involves the sequential activation of feature-specific cortical modules. *Nature Neuroscience, 17*(4), 619–624.

Schoenfeld, M. A., Tempelmann, C., Martinez, A., Hopf, J. M., Sattler, C., Heinze, H. J., & Hillyard, S. A. (2003). Dynamics of feature binding during object-selective attention. *Proceedings of the National Academy of Sciences of the United States of America, 100*(20), 11806–11811.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 174–215.

Treisman, A., Kahneman, D., & Burkell, J. (1983). Perceptual objects and the cost of filtering. *Perception & Psychophysics, 33*(6), 527–532.