Psychological Research (2006) 70: 405–413
DOI 10.1007/s00426-005-0007-5

ORIGINAL ARTICLE

Jun-ichiro Kawahara · James T. Enns · Vincent Di Lollo

The attentional blink is not a unitary phenomenon

Received: 15 April 2004 / Accepted: 25 November 2004 / Published online: 13 December 2005

Abstract Identification of the second of two targets is impaired if it is presented less than about 500 ms after the first. Three models of this second-target deficit, known as attentional blink (AB), were compared:resource-depletion, bottleneck, and temporary loss of control (TLC). Five experiments, in which three sequential targets were inserted in a stream of distractors, showed that identification accuracy for the leading target depended on an attentional switch whose magnitude varied with distractor–target similarity. In contrast, accuracy for the trailing target depended on similarity between the target and the trailing mask. These results strongly suggest that the AB is not a unitary phenomenon. Resource-depletion was ruled out as a viable account. The effect of attentional switching was handled naturally by the TLC model, while bottleneck models offered the best account of the effect of backward masking.

People frequently do two things at the same time: they drive while talking on a cell-phone, or walk while chewing gum. However, there is a cost to distributing attention to more than one task. As a rule, tasks that are done concurrently are performed less accurately and more slowly than when done in isolation. This cost is seen not only when two tasks are performed concurrently, but also when they are performed in close temporal sequence. For example, when two visual targets are presented in rapid succession, correct identification of the first interferes with identification of the second. This second-target deficit, known as attentional blink (AB), is most evident when the temporal lag between the two targets is in the range 100–500 ms.

J. Kawahara (✉)
Hiroshima University, Hiroshima, Japan
E-mail: jkawa@hiroshima-u.ac.jp

J. T. Enns
University of British Columbia, Vancourer, Canada

V. Di Lollo
Simon Fraser University, Burnaby, Canada

Performance deficits such as the AB have been found in a variety of dual-task experiments, and have been ascribed to the depletion of some limited processing resource that is required in common by the two tasks (for a review, see Pashler, 1998). In the present work, we consider whether the AB deficit can be explained fully on the basis of resource-depletion, or whether additional mechanisms play a part.

Our approach is couched in a comparison of three theoretical frameworks that have been proposed to account for the AB. One is based on the depletion of processing resources that are required for handling the trailing stimulus. Another is based on a processing bottleneck that occurs when two targets compete for the same mechanism. The third is based on the disruption of an optimal attentional set due to a temporary loss of endogenous control signals that are required for the maintenance of the attentional set. The experimental work is set within the broad context of conventional AB studies in which target stimuli, such as letters, are inserted in a stream of distractors, such as digits displayed in rapid serial visual presentation (RSVP). The main focus is on distinguishing among resource-depletion, processing bottleneck, and temporary loss of control (TLC) as the main factor in the AB deficit.

Three conceptual frameworks

Theoretical accounts of the AB can be classified in three broad categories, depending on the type of mechanism said to underlie the AB deficit: resource-depletion (e.g., Ward, Duncan, & Shapiro, 1996), bottleneck (e.g., Jolicoeur & Dell'Acqua, 1998), and TLC (Di Lollo, Kawahara, Ghorashi, & Enns, 2005). Resource-depletion models account for the AB in terms of the depletion of some limited resource by the first target with a consequent second-target deficit. For example, the attentional dwell-time hypothesis (Ward et al., 1996) and the interference model (Shapiro, Raymond, & Arnell, 1994) hold that the AB occurs because a limited processing resource is preempted by the first target thereby depriving the second target of needed resources.

Bottleneck models, on the other hand, postulate a processing stage which can handle only one item at any given time. If that stage is busy, processing of new items is delayed until that stage is free. Items so delayed suffer through decay or backward masking by trailing items. For example, in Chun and Potter's (1995) two-stage model, the AB is said to occur when the second target arrives while a high-level processing stage is busy with the first target. Similarly, in Jolicoeur and Dell'Acqua's (1998) PRP model, the AB is ascribed to a late-stage bottleneck which delays the deployment of high-level processing resources to the second target.

Resource-depletion accounts of the AB have been questioned by Di Lollo et al. (2005) in a study in which observers were required to identify three contiguous target letters inserted in an RSVP stream of digit distractors. According to the resource-depletion hypothesis, identification accuracy should be highest for the leading target and should decline progressively for each successive target as resources become more depleted. Two findings reported by Di Lollo et al. are inconsistent with this expectation. First, there was no progressive impairment in performance (i.e., no AB deficit) over successive targets: accuracy for the third target letter was as high as for the first. Second, a trailing-target deficit typically found in AB studies became evident when the middle letter in the three-target string was replaced with a digit. This is all the more remarkable because observers now had to identify only two targets.

Why was a trailing-target deficit in evidence when two targets were separated by a distractor but not when three targets were presented contiguously? Di Lollo et al. (2005) suggested TLC as the critical factor. Their proposal is based on an endogenously established attentional set aimed at rejecting distractors and accepting targets. Such a set involves many brain regions and needs to be maintained by endogenous signals from higher levels, such as the prefrontal cortex. This set can be maintained easily during the period leading to the first target, thus permitting efficient exclusion of leading distractors. As soon as identification of the first target begins, however, the endogenous maintenance signals are discontinued, and the attentional set becomes vulnerable to alteration by intervening distractors. This occurs when the leading target is followed by an item from a different category, such as a distractor or a mask. Perception of the ensuing target is then impaired because the initial attentional set has been disrupted exogenously by the intervening distractor. No such disruption occurs when the item following the first target is another target because it fits the current attentional set.

At a most general level, the TLC account could be regarded as yet another form of resource sharing or as another type of bottleneck. However, the nature of the sharing postulated in the TLC account differs fundamentally from that postulated in other models of the AB. Other models assume that the shared mechanisms are concerned with the processing of individual targets. To the extent that those mechanisms are engaged in processing the first target, they are unavailable—or are less available—for the trailing target. In contrast, according to the TLC hypothesis, the conflict arises from the need to carry out two distinct classes of processes: one is the process of issuing endogenous signals aimed at maintaining an optimal attentional set, the other is the process of target identification. In a very general sense, this is still a "sharing problem"; but the competition is for control of the central executive, not for control over target identification mechanisms. The limit, in this case, is not the number of items that can be processed at the same time, but the fact that the central executive cannot continue to maintain an optimal input configuration while at the same time oversee the processing of a target. To be clear about this, what is hypothesized in the resource-depletion and bottleneck accounts is a resource that can handle only one or, at best, two targets before depleting. In contrast, on the TLC hypothesis, the system can handle many more items without loss, provided they fit the current input configuration.

Resource-depletion beyond the AB

Given the pattern of results obtained by Di Lollo et al. (2005), it becomes pertinent to ask whether the resource-depletion hypothesis may be inadequate not only as an account of the AB but also as a general account of dual-task deficits. That account is questioned most specifically by an unexpected finding in the study of Di Lollo et al. (2005). In the condition in which three letter-targets were presented sequentially, identification accuracy was relatively low for the first target, increased significantly for the second target and then decreased for the third target, back to about the level of the first target. This inverted-V pattern is inconsistent with a resource-depletion hypothesis which, as noted above, would predict a progressive decrement in accuracy from the first to the third target.

Before dismissing resource-depletion as a viable account, however, we need to consider whether the RSVP methodology used in the study of Di Lollo et al. (2005) permitted an unbiased test of the resource-depletion hypothesis. One reason for questioning the suitability of the RSVP paradigm is that the arrival of the leading target entails a switch in attentional set. When the first target arrives, the observer's set must change from the one aimed at rejecting distractors to the one aimed at processing targets. Such attentional shifts are known to interfere with target identification (Kawahara, Zuvic, Enns, & Di Lollo, 2003; Sperling & Weichselgartner, 1995). Thus, in the study of Di Lollo et al., identification accuracy for the leading target might have been artificially depressed by the concomitant switch in attentional set. In this scenario, the advantage conferred to the first target by the untapped reservoir of processing resources was obscured by the harmful effect of the attentional switch. It could plausibly be claimed that, had performance not been depressed by the attentional switch, accuracy for the first target would have been higher than the accuracy for the second target, thereby producing the progressive decrement in performance from the first to third target predicted by the resource-depletion hypothesis.

This issue can be resolved with a simple experiment in which identification accuracy for three contiguous targets embedded in an RSVP stream, as in the study of Di Lollo et al. (2005), is compared with the identification accuracy for the same three targets presented by themselves, without any leading or trailing items. Based on the findings of Di Lollo et al., we expected an inverted-V function when the three targets were embedded in an RSVP stream of distractors. By the same token, based on the resource-depletion hypothesis, we expected a monotonic decreasing function when the three targets were presented by themselves. This comparison was carried out in Experiment 1.

Experiment 1

Method

Observers Thirty-two experimentally naive undergraduate students at the University of British Columbia participated in the experiment. All reported normal or corrected-to-normal vision. The 32 observers were assigned randomly to one of two groups of 16 observers each: the Stream group and the No Stream group.

Apparatus and stimuli Stimuli were displayed as black characters on the white screen of a NEC AccuSync computer monitor operating in a Microsoft Windows environment. The stimuli consisted of digits and uppercase letters subtending approximately 1° of the visual angle at a viewing distance of 57 cm.

Procedure At the beginning of each trial, a fixation cross appeared in the centre of the screen. Each trial began 500 ms after the observer pressed the spacebar. The display sequence for the Stream group consisted of an initial stream of 5–10 digits selected randomly with replacement from the digits 0–9, with the constraint that the selected digit was not one of the two immediately preceding items. The display stream continued with three sequential letters chosen randomly without replacement from the English alphabet excepting I, O, Q, and Z, and terminated with a single digit that acted as a mask after the last letter. The display for the No Stream group was the same as for the Stream group except that the leading stream of distractors and the trailing mask were omitted. That is, the RSVP stream consisted of only three sequential letters. Each item was displayed for 30 ms with an inter-stimulus interval (ISI) of 70 ms. Observers reported the identity of the three letters, in any order, by pressing the corresponding keys on the keyboard.

Results

The results are illustrated in Fig 1. An analysis of variance (ANOVA) was performed on the accuracy scores for the three targets in the Stream and No Stream groups. The ANOVA comprised one between-subjects factor, Group (Stream *vs.* No Stream), and one within-subject factor, Position in the Target String (First, Second, or Third). The analysis revealed a significant effect of Group, $F(1,30) = 17.20$, $P < 0.001$, $MSE = 2420.04$, a significant effect of Position, $F(2,60) = 10.80$, $P < 0.001$, $MSE = 641.79$, and a significant interaction effect, $F(2,60) = 31.81$, $P < 0.001$, $MSE = 1890.67$.

Discussion

The results in Fig. 1 do not exhibit the progressive decrement in accuracy predicted by the resource-depletion hypothesis. If anything, they show the opposite trend. In both conditions, accuracy increased markedly from the first to the third target, rather than decreasing as predicted by the resource-depletion hypothesis. This is especially notable in the No Stream condition, which was designed as an explicit test of that hypothesis.

Instead of the progressive decrement predicted by the resource-depletion hypothesis, Fig. 1 reveals a complementary pattern of results in the Stream and No Stream

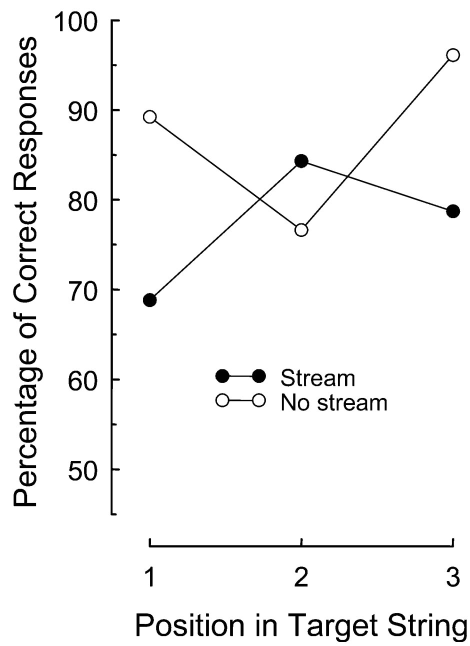

Fig. 1 Mean percentage of correct responses in Experiment 1

conditions. The No Stream condition yielded a V-shaped function of accuracy over targets. In contrast, the Stream condition yielded an inverted V-shaped function, replicating that obtained by Di Lollo et al. (2005) under similar conditions. The two functions differ most markedly in respect to the first and the third targets, with accuracy in the Stream condition being substantially lower in both cases.

This pattern of results has important implications for all three classes of models of the AB: resource-depletion, bottleneck, and TLC. It is clear from Fig. 1 that resource-depletion, as such, does not provide an adequate account. According to resource-depletion models of the AB, identification of the second of two targets is said to be impaired because processing resources are depleted by the first target. In direct contradiction of this hypothesis, the results of the No Stream condition show that identification of the third of the three targets was unimpaired by the requirement to identify two previous targets. This means that the system is capable of handling far more than the two targets used in conventional AB studies without depleting its resources.

Similarly, bottleneck models cannot explain why first-target accuracy was so much lower in the Stream than in the No Stream condition (Fig. 1). At the time the first target arrived, the bottleneck stage must be presumed to have been free in both the Stream and the No Stream conditions. This is because distractors are said to be excluded from the bottleneck stage.

In contrast, the lower first-target accuracy in the Stream condition is explained naturally by the TLC model. In the No Stream condition, the system was configured to process letter-targets from the start. So, when the first target arrived, the system was ready and the target was processed efficiently and without delay. In the Stream condition, however, when the first target arrived the system was set for rejecting distractors and had to be reconfigured to process the first target. The reconfiguration caused processing of the first target to be delayed for a brief period during which the target was vulnerable to masking by trailing items. Reduced accuracy of identification then followed. This delay hypothesis is tested directly in Experiment 2.

A second notable difference illustrated in Fig. 1 is the markedly higher accuracy of the third-target identification in the No Stream than in the Stream condition. This result can be explained most readily by the bottleneck model. When the third target arrived, the bottleneck stage was busy with the second target and, therefore, processing of the third target was delayed. In the Stream condition, this resulted in impaired identification because, while delayed, the third target was replaced by the trailing mask, perhaps through a process of object-substitution (Di Lollo, Enns, & Rensink, 2000). In the No Stream condition, however, there was no trailing mask and, therefore, the third target was processed without impairment at the end of the delay period.

The difference in the third-target accuracy, on the other hand, cannot be explained readily by the TLC model. In both the Stream and the No Stream conditions, the third target was preceded by two other targets. So, the system should have been configured identically in both conditions, and the third target should have been processed to the same level of accuracy. The lower accuracy in the Stream condition can be explained by the TLC model only with the additional consideration of backward masking. The importance of backward masking for both the bottleneck and the TLC accounts is examined further in Experiment 4.

Experiment 2

In Experiment 1 we found that first-target accuracy was lower in the Stream than in the No Stream condition. According to the TLC account, the first-target accuracy suffered in the Stream condition because of the time required for system reconfiguration from distractor-rejection to target-acceptance. Such reconfiguration was said to delay the processing of the first target and thus make it vulnerable to backward masking by trailing items. Experiment 2 was designed to provide a direct estimate of that delay by recording reaction times (RTs) to a letter-target that was preceded by a stream of digits, as in the Stream condition of Experiment 1, or by a corresponding blank interval, as in the No Stream condition of Experiment 1. On the TLC hypothesis, longer RTs were expected in the Stream condition.

Method

A new group of 12 observers served in Experiment 2. On any given trial, the target was an upper-case letter: C or G. Observers were instructed to identify the target, and to respond as quickly and accurately as possible by pressing the left- or the right-arrow key on the keyboard. The observer's RT was recorded, along with the accuracy of the response. There were two within-subject conditions, determined by whether or not a distractor-stream preceded the target. In the Stream condition, the target letter was preceded by a stream of digits as in the Stream Condition of Experiment 1. In the No Stream condition, the distractors were replaced by a corresponding number of blank frames. The Stream and No Stream conditions were presented in separate blocks of trials. On any given trial, the stream ended with the target. Namely, there was no trailing mask in Experiment 2. In every other respect, the procedures in Experiment 2 were the same as in Experiment 1.

Results and discussion

The results are illustrated in Fig. 2. A *t*-test for related samples was performed on the RT scores in the two conditions. The *t*-test revealed a significant difference

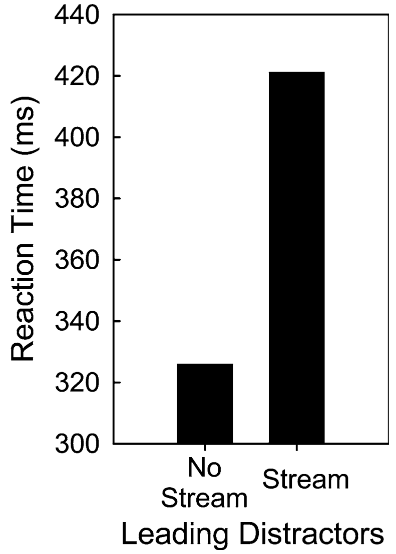

Fig. 2 Mean reaction times in the Stream and No Stream conditions in Experiment 2

between the Stream and the No Stream conditions, $t(11) = 13.53$, $P < 0.001$.

The results in Fig. 2 clearly support the idea, based on the TLC model, that the transition from distractors to the target in the RSVP stream causes a delay in target processing. The longer RT in the Stream condition is consistent with the TLC account of why first-target accuracy was impaired in the corresponding condition in Experiment 1: the processing delay rendered the first target vulnerable to masking by the trailing item.

Experiment 3

The results of Experiment 2 were consistent with the TLC account of Experiment 1 in which accuracy for the first target was lower in the Stream than in the No Stream condition. According to the TLC account, the lower accuracy arose from a processing delay due to the system's reconfiguration from distractor-rejection to target-acceptance. However, an alternative account of these findings is possible, not involving system reconfiguration. It is possible that the delay in responding to the first target in the Stream conditions in Experiments 1 and 2 was due not to the switch from distractors to the target but to a distracting effect produced by the transient response triggered by the onset of each leading item in the RSVP stream.

This possibility was checked in Experiment 3 by replicating the Stream condition of Experiment 1 with different types of leading distractors (random-noise dots, vertical grids, digits, and letters in script font). The intent was to vary the degree of similarity between the distractors and the letter-targets while preserving the transient responses triggered by the distractors in the RSVP stream. We reasoned that the magnitude of the transient responses would be approximately equal across distractor types. Therefore, if the impairment in the first-target accuracy was due to transient-based distraction, its magnitude would be approximately equal across distractor types.

If, on the other hand, the lower first-target accuracy in the Stream condition of Experiment 1 was due to a distractor-to-target switch, the impairment should be largest when the distractor–target similarity is high, and smallest when the similarity is low. This is because the task of discriminating the transition from distractors to targets in the RSVP stream should be harder and, therefore, more time-consuming when distractor–target similarity is high than when it is low.

Method

Procedures in Experiment 3 were the same as in Experiment 1, with the following exceptions. A new group of 12 observers participated in Experiment 3. The targets were three upper-case letters drawn in Courier New font, presented sequentially, as in the Stream condition of Experiment 1. The three targets were preceded by a stream of distractors drawn from one of four sets of stimuli. Depending on the condition, each frame in the distractor-stream consisted either of a digit (Digit condition), as in the Stream condition in Experiment 1, or a grid of black vertical lines (Grid condition), or an aggregate of random-noise dots (Noise condition), or a letter in script font (Script condition). Examples of the leading streams of distractors are illustrated in Fig. 3. The average number of pixels was approximately the same (about 142) for each item in the four distractor types. This was done in an attempt to equalize the strength of the transient response across the four types of distractors.

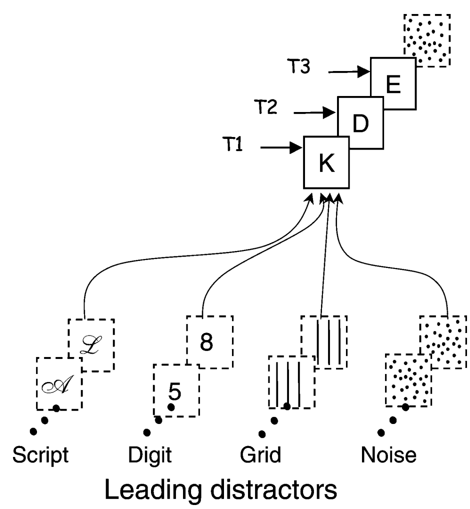

Fig. 3 Schematic representation of the display sequence of stimuli in Experiment 3

In the Grid condition, any given frame contained three vertical lines, each 2 pixels wide and 24 pixels high. Viewed at a distance of approximately 57 cm, each vertical line subtended a visual angle of approximately 1°. The horizontal separation between adjacent lines was 6 pixels. On any given frame, the grid was shifted horizontally by ±8 pixels at random, within an imaginary square of 1° side. In the Noise condition, each frame consisted of 142 black dots positioned randomly within an imaginary square of 1° side. A new random configuration of the dots was drawn for each frame. In the Script condition, the letters were drawn in "English 111 Vivace BT" (E111viva.ttf) font. The three target letters were never used as distractors. Thus, the target letters differed from the distractors both in identity and in font. In each of the four distractor conditions, the RSVP stream ended with a random-dot mask similar to the frames in the Noise condition. The display sequence in Experiment 3 is illustrated in Fig. 3.

The four distractor conditions were presented in separate blocks of 100 trials each. The sequence of the blocks was determined as follows. For the first observer, the sequence was: Grid, Noise, Digit, Script. This sequence was rotated by one place for each successive observer, such that the sequence for the second observer was: Noise, Digit, Script, Grid, and so on for each successive observer.

Results and discussion

The results are illustrated in Fig. 4. An ANOVA was performed on the accuracy scores for the three targets in the four conditions defined by the type of leading distractor. The ANOVA comprised two within-subject factors: Type of Leading Distractor (Grid, Noise, Digit, or Script), and Position in the Target String (First, Second, or Third). The analysis revealed a significant effect of Type of Leading Distractor, $F(3,33)=21.98$, $P<0.001$, $MSE=607.23$, a significant effect of Position, $F(2,22)=52.06$, $P<0.001$, $MSE=11067.72$, and a significant interaction effect, $F(6,66)=10.60$, $P<0.001$, $MSE=336.24$. Multiple comparisons in the first-target accuracy following the ANOVA (Ryan, 1960) revealed that the Script condition differed significantly from the remaining three conditions ($P<0.001$ in every case), and that the Digit condition differed significantly from the Grid condition, $t(99)=3.19$, $P<0.005$.

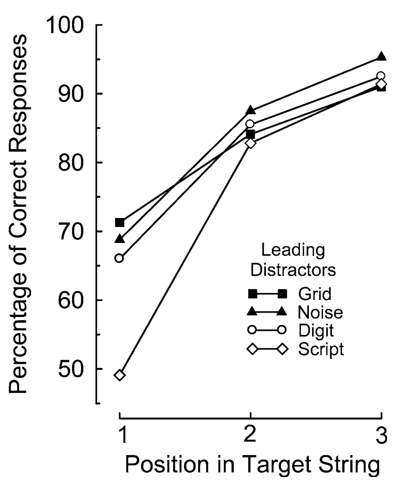

Fig. 4 Mean percentage of correct responses in Experiment 3

Identification accuracy for the first target varied as a function of distractor type. Accuracy was lowest in the Script condition, in which the similarity between distractors and targets was the highest, and increased as similarity decreased, with the Digit condition at an intermediate level. Similar results have been reported by Maki, Bussard, Lopez, and Digby (2003) and by Visser, Bischof, and Di Lollo (2005). These results strongly suggest that the first-target deficit seen in Experiment 1 was due not to the distracting effects of transient responses triggered by the distractors but to a delay in the first-target processing arising from the switch from distractors to targets in the RSVP stream. According to the TLC hypothesis, first-target accuracy was lowest in the Script condition because the high similarity between distractors and targets led to a relatively long delay in switching from distractor-rejection to target-acceptance modes. In turn, the longer delay caused the first target in the Script condition to be vulnerable to masking over a relatively longer period, leading to the impaired accuracy of identification.

This account hinges on the assumption that the latency of processing the first target was directly related to the similarity between distractors and targets, as implemented by the four distractor types in Experiment 3. Experiment 4 was designed to examine this assumption by recording RTs to a letter-target that was preceded by the same types of distractors as were used in Experiment 3. On the TLC hypothesis, the RTs for the four distractor types should be ordered from longest to shortest in the same sequence as the first-target accuracy in the corresponding condition in Experiment 3 (Fig. 4).

Experiment 4

Method

A new group of 12 observers served in Experiment 4. Procedures were the same as in Experiment 2 except that there were four within-subject conditions, determined by the type of distractor-stream that preceded the target. The four types of distractors were the same as in Experiment 3 (Noise, Grid, Digit, or Script), and were presented in separate blocks of trials.

Results and discussion

The results are illustrated in Fig. 5. An ANOVA was performed on the RT scores in each of the four condi-

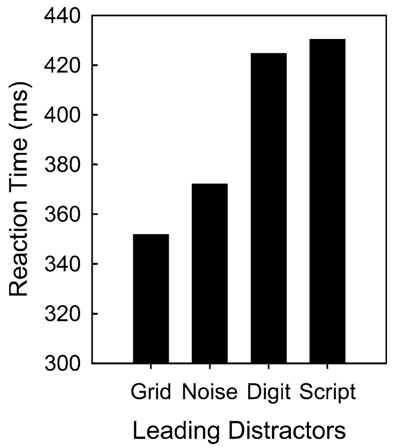

Fig. 5 Mean reaction times in each condition in Experiment 4

tions defined by the type of leading distractor. The ANOVA comprised a single within-subject factor: Type of Leading Distractor (Noise, Grid, Digit, or Script). The analysis revealed a significant effect of the Type of Leading Distractor: $F(3,33) = 55.46$, $P < 0.001$, $MSE = 18062.41$.

The results in Fig. 5 show that the order of RTs in Experiment 4 matched the order of first-target identification accuracy in the corresponding condition of Experiment 3. This supports the idea, based on the TLC model, that the impairment in the first-target identification accuracy in Experiment 3 was mediated by a delay caused by the switch from distractor-rejection to target-acceptance modes.

Experiment 5

Experiment 1 showed that the identification accuracy for the first and the last targets was substantially lower in the Stream than in the No Stream condition. This does not mean, however, that the two instances of impairment were mediated by the same mechanism. In Experiments 2–4 we found support for the hypothesis that the first-target impairment in Experiment 1 was mediated by a processing delay triggered by the switch from distractors to targets. The same task-switching principle, however, cannot apply to the impairment of the third target because it was preceded by another target. Rather, we have hypothesized that the third-target impairment in Experiment 1 might have been brought about by backward masking in the Stream, as distinct from the No Stream, condition.

This backward-masking account was pursued in Experiment 5 by building on Hellige, Walsh, Lawrence, and Prasse's (1979) finding that the strength of backward masking is modulated by the similarity between the target and the mask. We reasoned that if the third-target impairment in Experiment 1 had been brought about by backward masking, then the magnitude of that impairment should be modulated by the similarity between the target and the trailing mask. In Experiment 5 the target-mask similarity was manipulated in the same manner as the distractor–target similarity was manipulated in Experiment 3. Namely, in Experiment 5 there were four types of masks, corresponding to the four types of distractors used in Experiment 3.

Method

Procedures in Experiment 5 were the same as in Experiment 3, with the following exceptions. The targets were three upper-case letters drawn in Courier New font, presented sequentially as in the Stream condition in Experiment 1. The three targets were preceded by a stream of distractors consisting of aggregates of random-noise dots, similar to those used in the Noise condition in Experiment 3. The stream ended with a mask which, depending on the condition, consisted either of a digit or a grid or an aggregate of random-noise dots or a letter in script font. The type of mask and the display sequence in Experiment 5 are illustrated in Fig. 6.

Results and discussion

The results are illustrated in Fig. 7. An ANOVA was performed on the scores for the three targets in the four conditions defined by the type of mask. The ANOVA comprised two within-subject factors: Type of Mask (Noise, Grid, Digit, or Script), and Position in the Target String (First, Second, or Third). The analysis revealed a significant effect of the Type of Mask, $F(3,33) = 14.13$, $P < 0.001$, $MSE = 1015.26$, a significant effect of Position, $F(2,22) = 4.11$, $P < 0.05$, $MSE = 1249.71$, and a significant interaction effect, $F(6,66) = 27.68$, $P < 0.001$, $MSE = 679.98$.

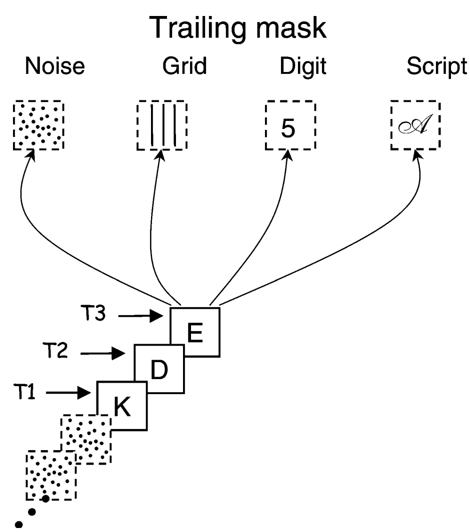

Fig. 6 Schematic representation of the display sequence of stimuli in Experiment 5

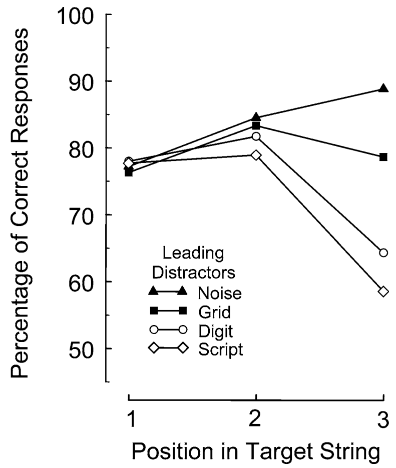

Fig. 7 Mean percentage of correct responses in Experiment 5

It is clear from Fig. 7 that the degree of the third-target impairment depended on the similarity between the target and the trailing mask. This finding implicates backward masking as the primary source of third-target impairment in the Stream condition in Experiment 1. Earlier findings by Hellige et al. (1979) showed that the strength of masking increases with increasing structural similarity between the target and the mask. The results of the present Script condition extend Hellige et al.'s findings to conceptual similarity.

None of the three theoretical frameworks examined in the present work can account fully for the results in Fig. 7. The decrement in accuracy from the first to the third target in the Script condition could be taken as evidence for resource-depletion. However, then it should be asked why the opposite trend is evident in the Noise condition. Assuming a unitary resource, resource-depletion models (e.g., Ward et al., 1996) do not seem capable of explaining why different types of masks affect third-target accuracy in different ways.

The TLC model does not fare any better. The third target was presented directly after two other targets, hence the system should have been configured appropriately. The trailing mask may well have changed the system's configuration, but this should not have affected the processing of items presented earlier, notably the third target.

Bottleneck models (e.g., Chun & Potter, 1995) are somewhat more successful but require additional assumptions about the mechanisms of masking. According to this model, masking occurs only in Stage 1 while Stage 2 is busy with an earlier target. It can be assumed, therefore, that the third target was masked while the second target was being processed in Stage 2. What needs to be explained, however, is why the strength of masking varied with the type of mask.

In view of the difficulties encountered by all three conceptual frameworks in accounting for the results in Fig. 7, the option must be considered that backward masking might be a source of impairment orthogonal to that attributable to multiple tasking. On this option, masking would exert an overriding influence on performance, thus overwhelming any interference due to multiple tasking.

General discussion

The objective of the present work was to compare three theoretical models of the AB: resource-depletion (e.g., Ward et al., 1996), bottleneck (e.g., Jolicoeur & Dell'Acqua, 1998), and TLC (Di Lollo et al., 2005). In Experiment 1, three letter-targets were presented alone or were inserted in an RSVP stream of digit distractors. Identification accuracy increased from the first to the third target, disconfirming expectations based on resource-depletion models. The leading stream of distractors impaired accuracy for the first target, supporting the TLC account, while the trailing mask impaired accuracy for the third target, supporting a bottleneck account. In Experiments 2–4 we manipulated the distractor–target similarity and found evidence in favour of the TLC account. In Experiment 5 we varied the target-mask similarity and found that the strength of masking varied not only with structural similarity (cf. Hellige et al., 1979) but also with conceptual similarity.

Collectively, the outcomes of all five experiments cannot be fully explained by any of the three theoretical frameworks. This raises the option that the AB deficit may not be a unitary phenomenon mediated by a single mechanism. Rather, the evidence points to at least two major factors: attentional switching and backward masking. Identification of the leading target is impaired if an attentional switch is required within the RSVP stream from distractor-rejection to target-acceptance. In terms of the TLC model, such a switch entails a time-consuming system reconfiguration during which the upcoming target is vulnerable to masking by subsequent items. This account is buttressed by the outcomes of Experiments 2–4 which provided a link between the time required for system reconfiguration and the magnitude of the impairment in target identification.

The second factor, backward masking, affects accuracy for the trailing target, and is explained naturally by bottleneck models. Namely, the deficit occurs when the trailing target arrives while the system is busy processing an earlier target. While so delayed, the representation of the trailing target is vulnerable to overwriting by the ensuing mask, and the AB deficit follows.

The present work strongly suggests that attentional switching and backward masking are distinct factors that affect performance independently of one another. Attentional switching affects only the leading target; backward masking affects only the trailing target. These effects are seen in Experiments 3 and 5. In Experiment 3, different types of leading distractors led to different levels of first-target accuracy but did not affect accuracy for the third target because all conditions shared the same mask (Fig. 4). The converse pattern is evident in

Experiment 5, where different types of trailing masks led to correspondingly different levels of accuracy for the third target but did not affect accuracy for the first target because all conditions shared the same stream of leading distractors (Fig. 7).

In summary, far from being a unitary phenomenon, the AB deficit seems to arise from multiple factors working independently from one another, as also suggested by Kawahara (2003). For this reason, it is unlikely that all the experimental evidence can be encompassed by a single theoretical model. We have seen that resource-depletion models based on a unitary resource are generally disconfirmed by the evidence. In contrast, the TLC and the bottleneck models can each account for separate aspects of the data. The effect of attentional switching can be handled naturally by the TLC model, while bottleneck models offer the best account of the effect of backward masking.

References

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 109–127.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, 129, 481–507.

Di Lollo, V., Kawahara, J., Ghorashi, S. M. S., & Enns, J. T. (2005). The attentional blink: Resource limitation or temporary loss of control? *Psychological Research*, 69, 191–200.

Hellige, J. B., Walsh, D. A., Lawrence, V. W., & Prasse, M. (1979). Figural relationship effects and mechanisms of visual masking.*Journal of Experimental Psychology: Human Perception and Performance*, 5, 88–100.

Jolicoeur, P. & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, 36, 138–202.

Kawahara, J.(2003). Mandatory processing of distractors: Another determining factor for the attentional blink. *Japanese Psychological Research*, 45, 140–151. (in English).

Kawahara, J., Zuvic, S. M., Enns, J. T., & Di Lollo, V. (2003). Task switching mediates the attentional blink even without backward masking. *Perception & Psychophysics*, 65, 339–351.

Maki, W. S., Bussard, G., Lopez, K., & Digby, B. (2003). Sources of interference in the attentional blink: target-distractor similarity revisited. *Perception & Psychophysics*, 65, 188–201.

Pashler, H. E. (1998). The psychology of attention. Cambridge, Massachusetts: MIT Press.

Ryan, T. (1960). Significance test for multiple comparisons of proportions, variances, and other statistics. *Psychology Bulletin*, 57, 318–328.

Shapiro, K. L., Raymond, J. E., & Arnell, K. M. (1994). Attention to visual pattern information produces the attentional blink in RSVP. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 357–371.

Sperling, G.,& Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review*, 102, 503–532.

Visser, T. A. W., Bischof, W. F., & Di Lollo, V. (2005). Rapid serial visual distraction: Task-irrelevant items can produce an attentional blink. *Perception & Psychophysics* (in press).

Ward, R., Duncan, J., & Shapiro, K. (1996). The slow time-course of visual attention. *Cognitive Psychology*, 30, 79–109.