# The Midstream Order Deficit

Alex O. Holcombe
Harvard University, Cambridge, MA

Nancy Kanwisher
Massachusetts Institute of Technology, Cambridge, MA

Anne Treisman
Princeton University, Princeton, NJ

Alex O. Holcombe
Department of Psychology
Harvard University
33 Kirkland St
Cambridge, MA 02138
voice: (617) 495-3884
fax: (617) 495-3764
holcombe@wjh.harvard.edu

accepted by Perception & Psychophysics June 1, 2000

Running Head: Midstream Order Deficit

## Abstract

The relative order of an auditory sequence can be more difficult to apprehend when it is presented repeatedly without pause (cycling) than when it is presented once (Warren et al., 1969). We find that this phenomenon, the “midstream order deficit”, can also occur with visual stimuli. The stimuli need not form separate perceptual "streams", and the effect occurs with presentation rates as slow as 5 items/second, even though the identification of individual letters is very accurate at this rate. However, if the first item of the sequence is visually very distinct from the preceding items, relative order reports can be as accurate in the cycling condition as in the single-presentation condition. Our results suggest that the midstream order deficit is not due to masking, an attentional blink, repetition blindness, Reeves & Sperling’s (1986) order illusion, memory limitations, or decision criteria. It may reflect an attentional cost to initiating order encoding, distinct from the allocation of attention required to detect and identify individual items. One’s attention must be set for, or captured by, an initial salient event in order to initiate order encoding successfully.

Warren, Obusek, Farmer, & Warren (1969) reported a striking inability of listeners to perceive the relative order of a sequence of four sounds presented repeatedly, despite fairly accurate performance when the sequence was presented only once. Participants first heard the four sounds separately and learned a name for each. Then they were told that the four sounds would be presented in a particular order repeatedly, and that afterwards they would be asked to report the correct relative order. A correct relative order report could begin with any one of the four sounds. For example, if buzz-beep-hiss-boop was presented over and over, "hiss-boop-buzz-beep" would be one correct response. The sounds were played for 200 milliseconds each, with no interstimulus interval, and the sequence was repeated continuously without a pause between repetions. The participants themselves ended the sequence whenever they were ready to respond. However, they felt that they could not apprehend the order, no matter how long they listened, and in fact their performance was not significantly different from chance. In contrast, another group of participants to whom the sequence was presented just once instead of multiple times performed significantly better than chance.

This paper describes three experiments which first demonstrate a similar effect with visual stimuli and then test several candidate explanations of the effect. Our results suggest that the poor performance with the cycling sequences is not simply due to a greater difficulty in perceiving the individual items; rather, cycling causes a problem specific to apprehending order. In particular, we find that the accuracy of order judgments is correlated with the degree to which an identifiable item is perceptually distinct from the preceding items. We suggest that this is due to a difficulty in rapidly initiating order encoding without a salient perceptual event.

## Experiment 1: Visual letters with single versus multiple presentation

This experiment documents conditions in which accuracy at judging relative order is lower for repeated presentations than for a single presentation of a sequence. We roughly equated forward and backward masking in the two conditions by embedding the single presentation in a longer sequence of meaningless line patterns. In both conditions, the first few items were presented at a very high rate, after which the presentation rate gradually slowed until reaching the final rate by the ninth item. Schematic diagrams of the single and repeated (cycling) conditions are shown in Figure 1 and described in detail below[1].

### Method

[1] A demonstration of the difficulty of apprehending the order of cycling sequences can be viewed at *http://wjh.harvard.edu/~holcombe* or obtained by contacting the first author.

Participants

Eight Harvard University students who were naive to the purpose of the experiment and reported that they had normal or corrected vision were paid to participate.

Apparatus

This and all subsequent experiments were conducted on a MacOS computer running custom software created with Vision Shell C libraries (contact *raco@wjh.harvard.edu*). Participants viewed an Apple 13-inch color monitor from a distance of approximately 76 cm. The participants responded by pressing keys on a keyboard.

Procedure and Design

The sequence began at a very fast presentation rate and then decelerated: the first item was presented for one screen refresh (the monitor had a 60 Hz refresh rate), and each succeeding item was presented slightly longer until the final duration was reached at the ninth item. Four final exposure durations were used: 120, 165, 210, and 255 milliseconds. In each of these conditions, the first five items were presented for the following durations, from first to fifth: 15, 30, 30, 45, and 75 milliseconds. After the fifth item, the duration of each succeeding item was larger than the duration of the previous item by a constant amount, chosen so that the full duration was reached at the ninth item. To fit the refresh rate of the monitor these durations were rounded to the nearest multiple of fifteen milliseconds. All of the items were presented at high contrast in a dimly-lit room. The items subtended approximately 0.48.

In the cycling condition, the four-letter sequence was presented three times after the deceleration. The items were displayed in a thin white rectangle, which was filled before the trial began and became unfilled when the first item of the stream appeared. For each trial, the four-letter sequence was chosen randomly from the set of all consonants other than ‘Q’. In the single-presentation condition, the four-letter sequence was embedded in a stream of line patterns, as shown in Figure 1. The particular line pattern at a given position in each stream was chosen randomly, with the constraint that successive line patterns were always different.

The participants were instructed to fix their eyes on the stimulus the entire time that it was presented, and to report the letters in order, beginning with any of the four. As an example, they were told that if the letter sequence presented was “BFGH”, then the following responses would all be counted as correct: “BFGH”, “FGHB”, “GHBF”, and “HBFG”. They were to enter the letters by typing them into the computer keyboard. The experimenter guided the observer through six practice trials. The first few practice trials used a very slow presentation rate to familiarize the participants with the stimuli and task.

The observer's initial response was used to calculate the number of letter identities correctly reported. In order to gain supplementary data, if the observer entered a letter that had not been presented, she was provided with the four letters that actually were presented and was told to enter the relative order again using the correct letters. If the participants entered incorrect letters on their first try, this second response was used in analyses of their relative order accuracy. This was to allow for the possibility that participants might encode the correct order but need to be cued with the correct identities to recall the order.

The experimenter guided the participants through six practice trials. The first few used a very slow presentation rate to familiarize the participants with the stimuli and task. After the practice trials, there were 192 experimental trials, with a short break halfway through. The 192 experimental trials were comprised of the single vs. cycling presentation factor crossed with the four final-item exposure durations, presented in pseudorandom order.

## Results and Discussion

Participants usually reported the identities of the letters correctly, averaging at least 94% correct in all conditions. Letter identity accuracy was slightly higher in the cycling condition than in the single-presentation condition (98.1% vs 96.6%), $F(1,7) = 5.6$, $p < .05$. The interaction of the presentation condition with the final duration was not significant. All analyses were carried out with a 3-way ANOVA with subject, duration, and cycling vs. single-presentation as factors and all interactions included.

Participants' high accuracy in reporting the letter identities suggests that they perceived and remembered almost all the letters in both the cycling condition and the single condition. This result is consistent with work showing that, at least at the slower rates, these presentation conditions result in minimal low-level masking, attentional blink (Chun & Potter, 1995; Moore, Egeth, Berglan, & Luck, 1996), and repetition blindness effects (Park & Kanwisher, 1994).

However, participants often reported the letters in the wrong order. Moreover, accuracy in report of order was significantly lower in the cycling condition than in the single-presentation condition, as shown in Figure 2 ($F(1,7) = 5.91$, $p < .05$). Most of the participants also reported that the cycling condition felt more difficult.

The interaction of presentation condition with final duration was marginally significant, $F(3,21) = 2.91$, $p = .058$. However, we suspect that this trend of a smaller difference between the conditions at faster rates is due to random variation, because other similar (unpublished) experiments yielded the opposite interaction. Furthermore, our informal observations clearly show that as the final item duration is increased further, the difference between the two conditions rapidly diminishes to zero or reverses; that is, at slow rates the repetition of the cycling condition becomes a benefit rather than a cost.

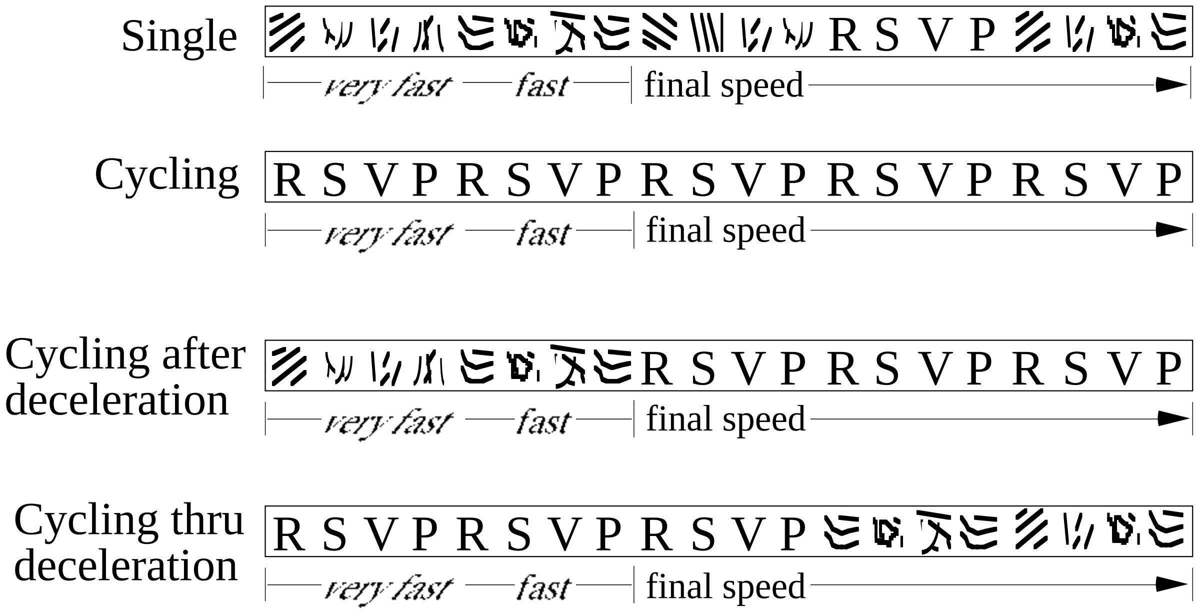


Figure 1, Holcombe et al.

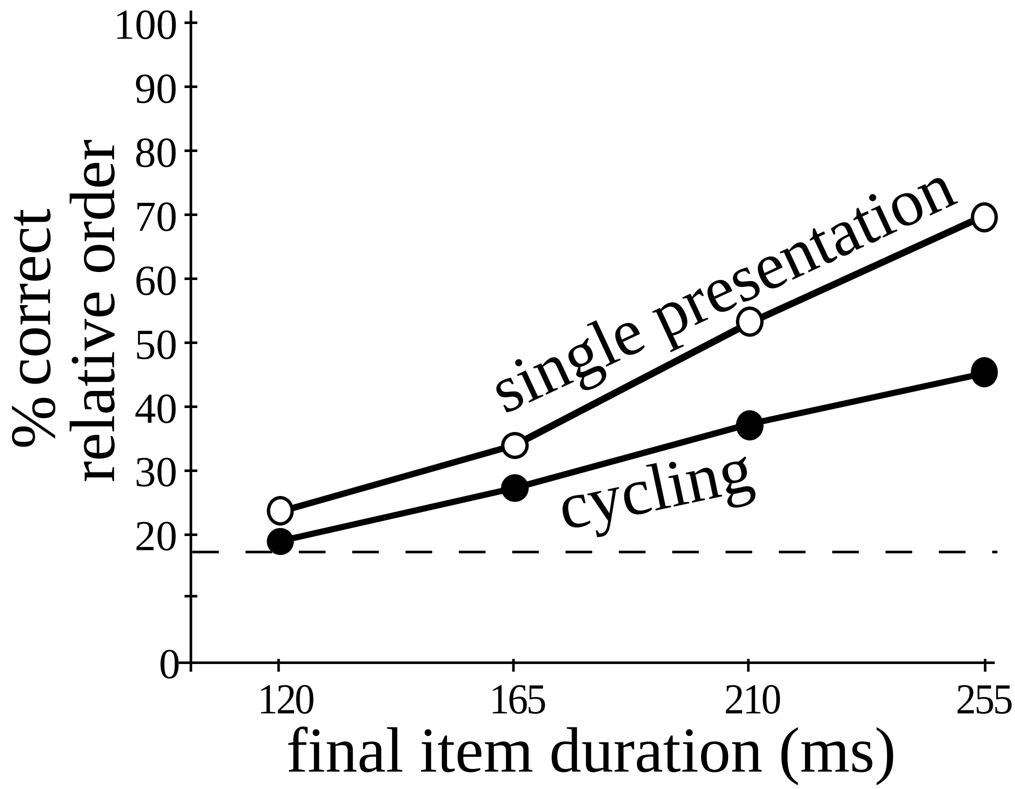


Figure 2, Holcombe et al.

During the trials of the single-presentation condition in which the letter sequence was presented near the beginning of the stream, the observers had to remember the sequence for a few seconds before reporting it. To test whether this led to more forgetting of the sequence, thereby decreasing the accuracy in the single-presentation condition, we examined the effect on accuracy of the letter sequence's position in the stream. The effect of position did not reach significance, either as a categorical variable added to the previous ANOVA, $F(8,64) = .69$, $p = .7$, or added as a regressor (in an analysis of covariance), $t(8) = 2.01$, $p = .08$. Although it nearly reached significance in the second analysis, the regression coefficient was only .01, suggesting that order accuracy was 8% better when the sequence was presented in the last possible position than in the earliest possible position (8 positions earlier). This effect is too small to explain the difference between the cycling and single-presentation conditions.

Within a letter sequence in the single-presentation condition, the first letter was reported in 96% of trials, the second in 98%, the third in 97%, and the fourth in 96%. This pattern provides evidence against the possibility that an attentional blink (Raymond, Shapiro, & Arnell, 1992) contributed to the difficulty of temporal order reports. If attention to the first letter had produced a blink for its successors, the successors should have been reported less often. But there was no indication that letters after the first were identified less accurately. This result and the fact that the attentional blink has not been observed with unmasked streams presented as slow as the current rates suggests that the attentional blink was not a factor in Experiment 1.

Overall, the results of Experiment 1 show that, as with auditory stimuli (Warren et al. 1969), repeated, cycling presentation of visual stimuli can lead to lower accuracy on relative order judgments than is found for single presentation. Our use of slow rates and unmasked presentations and the fact that participants had no difficulty in identifying the items makes it unlikely that the poorer relative order accuracy in the cycling condition was due to a greater masking, attentional blink, or repetition blindness effect.

## Experiment 2: Detection

In the previous experiment full report of the sequence was used to assess the perception of the temporal order. The deficit in accuracy with cycling sequences could therefore reflect differences in memory load or in decision criteria rather than in perceptual identification or attentional processes. In order to determine more precisely the source of the difficulty in the cycling condition, we used an order detection task in Experiment 2.

Participants viewed a specific cue sequence at the beginning of each trial and were asked to determine whether the subsequent stimulus sequence had the same relative order. In this detection task the demands on perceptual identification and memory should be significantly lower than in the full-report tasks. The participants know which letters will be presented, making it easier to identify the individual letters during the presentation. Memory load for the stimulus sequence should be lower because instead of remembering the entire stimulus sequence until the response, the participants need only remember whether or not the stimulus sequence matched the cue. Although the participants must remember the cue sequence, this should be less demanding than remembering the stimulus sequence, because the participants could take their time in committing it to memory before each trial. The detection paradigm also has the advantage that it yields separate measures of the stimulus information available and of the participants' decision criteria.

Another goal of Experiment 2 was to investigate the possibility that the difficulty in apprehending order is not specific to cycling sequences, but rather reflects a more general difficulty in apprehending order from the middle of a relatively undifferentiated stream. To test this we embedded the single cycle of letters in a stream of digits and compared this condition to embedding it in the stream of line patterns previously tested in Experiment 1. Although the line patterns should have been about as effective as the digits in masking, the digits are more similar to the letters at a conceptual level.

Experiment 2 also offers a test of another account of the difficulty with cycling sequences. Perhaps the relative order of the four letters is encoded equally well in the cycling and single-presentation conditions, but the subsequently presented letters of the cycling condition disrupt this memory. If the additional letters do harm retention of the stimulus sequence, then they may also harm retention of the cue sequence, resulting in poorer memory for its order in the cycling condition than in the single-presentation condition. To test this hypothesis, observers were asked to report the cue sequence after each trial.

## Method

### Participants

Thirty Harvard University students who did not know the purpose of the experiment and who reported that they had normal or corrected vision participated for payment or course credit.

### Procedure and Design

At the beginning of each trial, a 4-letter cue sequence was displayed (simultaneously in one row) until the observer made a keypress. The participants' task was to determine whether the stimulus sequence was in the same relative order as the cue sequence. The four-letter stimulus sequence was drawn randomly from the set of all consonants other than 'Q'. In half of the trials, the cue sequence was identical to the stimulus sequence. For the other half, the cue sequence was formed by scrambling the stimulus sequence- transposing a random pair of the letters, excluding those pairs which included the first letter. (Transpositions involving the first letter were avoided to prevent observers from simply checking the first letter, a strategy which would benefit the single-presentation condition.) The first letter of the cue sequence was always the same as the first letter of the stimulus sequence (although the very first letter of the stimulus was unidentifiable in the cycling condition because it was presented very briefly).

When the participants had memorized the cue sequence, they pressed a key and a filled white rectangle was presented until another key was pressed. The stimulus sequence was then presented, in the same fashion as Experiment 1 except that the four-letter sequence of the single-presentation condition could begin only at the ninth, thirteenth, or seventeenth position of the twenty-item stream. The choice of position was distributed randomly across trials and counterbalanced across conditions. Participants were encouraged to wait for the first letter of the cue and then check whether the rest of the letters appeared in the same order as in the cue. After the presentation of the stimulus sequence, participants pressed a key to indicate whether the stimulus was the same as the cue or different. Finally, the participants were prompted to enter the cue sequence into the keyboard.

The two fully crossed factors of this experiment were final duration of the letters (165 or 180 milliseconds) and cycling versus single presentation. In addition, the single-presentation condition contained either digits or line pattern masks. One block consisted of cycling presentation trials and single-presentation trials with digit masks, and the other block consisted of cycling presentation trials and single-presentation trials with line pattern masks. The order of blocks was counterbalanced across observers.

After the participants were guided through eight practice trials by the experimenter, they each participated in a 96-trial block. After the first block, each participant took a short break, after which the experimenter explained that the masks would change (from digits to line patterns or vice versa). The participant then participated in eight more practice trials, followed by the second 96-trial block.

## Results and Discussion

The pattern of results was equivalent for the two presentation durations used, so the data were collapsed across duration. A signal detection analysis was conducted to obtain estimates of performance independent of guessing or response biases. A high-threshold alpha model was used to estimate sensitivity, as its assumptions were found to fit the data of previous RSVP detection experiments better than the Gaussian $d'$ model (Kanwisher, Kim, & Wickens 1996). In addition, the alpha model is more in accord with the subjective reports of most subjects that on some trials they knew the order but on most of the other trials they did not know the order and had to guess. All of the analyses reported below were also done with uncorrected percent correct as the dependent measure, and the same patterns of results were found, with the same comparisons statistically significant.

Table 1 shows the mean percent correct in each of the three conditions (cycling, single-presentation with line-pattern masks, and single-presentation with digit masks) as well as the means of the sensitivity measure $a$ and the estimated guessing rate $g$. These measures were figured separately for each observer before the means were calculated. The percent correct numbers were higher than in Experiment 2, because chance in this detection experiment was 50%, whereas chance in Experiment 2 was no higher than 17% (17% is an upper bound on chance relative order accuracy because it assumes that the observers knew which four letters were in the sequence on every trial).

---Insert Table 1 about here---

Sign tests revealed that all the pairwise a differences shown in Table 1 were significant. The observers performed less accurately in the cycling condition than in both the single-presentation line pattern masks condition ($p < .001$) and the single-presentation digit masks condition ($p < .01$). All observers also reported that the cycling condition felt most difficult.

Thus even though observers knew exactly what they were looking for and knew exactly which letters would be presented on each trial, their accuracy was still significantly lower in the cycling condition than in the single-presentation condition.

The digit mask condition was more difficult than the line-pattern condition ($p < .05$), yet was easier than the cycling condition ($p < .05$). This suggests that the low accuracy in the cycling condition may be due not to repetition of the sequence *per se*, but rather to a more general difficulty in apprehending order in an undifferentiated stream.

Experiment 2 was also designed to test the hypothesis that the subsequent repetitions of the sequence disrupt memory retention. If this were true, then one would expect observers to forget the cue four-letter sequence more often in the cycling condition. To the contrary, the observers' accuracy in reporting the cue was equal in both conditions, at 94% correct. However, it remains possible that the additional letters of the cycling condition disrupted short-term retention of the stimulus sequence without disrupting the retention of the cue sequence because the cue sequence was more securely committed to memory.

## Experiment 3: Effect of initial deceleration

In the preceding experiments, the cycling condition differed from the single-presentation condition in two important respects. First, the letter sequence was presented during the deceleration phase in the cycling condition, whereas it was not presented during the deceleration phase in the single-presentation condition. Second, in the cycling condition the letter sequence was presented multiple times after the deceleration phase instead of just once. In Experiment 3, these two factors were varied independently to determine which accounts for the low order accuracy in the cycling condition. To preview the results, performance with cycling presentation is as good as performance with single presentation if the sequence is not presented during the deceleration phase, suggesting that cycling does not impair accuracy if the beginning of the sequence is salient.

### Methods

Participants

Eight Harvard University students who had normal vision and were naïve to the purpose of the experiment were paid to participate.

Procedure and Design

The procedure was the same as Experiment 2 except as noted. Instead of a detection procedure, observers reported the four-letter sequence as in Experiment 1. The four display conditions of this experiment are schematized in Figure 1.

The cycling and single-presentation displays were identical to those of Experiment 1. But in the new "cycling during deceleration" condition, the letter sequence was presented throughout the deceleration phase and then presented once immediately after the deceleration phase. Line pattern masks were presented thereafter. In the "cycling after deceleration" condition, line pattern masks were presented during the deceleration phase and the letter sequence was presented thereafter. If the lower accuracy in the cycling condition is due simply to the cycling of the sequence, regardless of whether this cycling occurs during the deceleration phase, then accuracy in both of the new conditions should be lower than accuracy in the single presentation condition. If, alternatively, it is critical that the cycling begin during the deceleration phase, then lower accuracy should only occur in the cycling during deceleration condition.

The experimenter diagrammed and verbally described the deceleration and the four display conditions, then guided observers through eight practice trials which exposed them to each of the conditions twice. Each trial began by informing the observer of the condition that was about to be presented. The observer participated in 192 experimental trials, with a short break halfway through. The four display conditions were fully crossed with the two final exposure durations, 165 and 180 milliseconds. This resulted in 24 trials for each of the resulting condition x duration combinations, which were presented in pseudorandom order.

## Results and Discussion

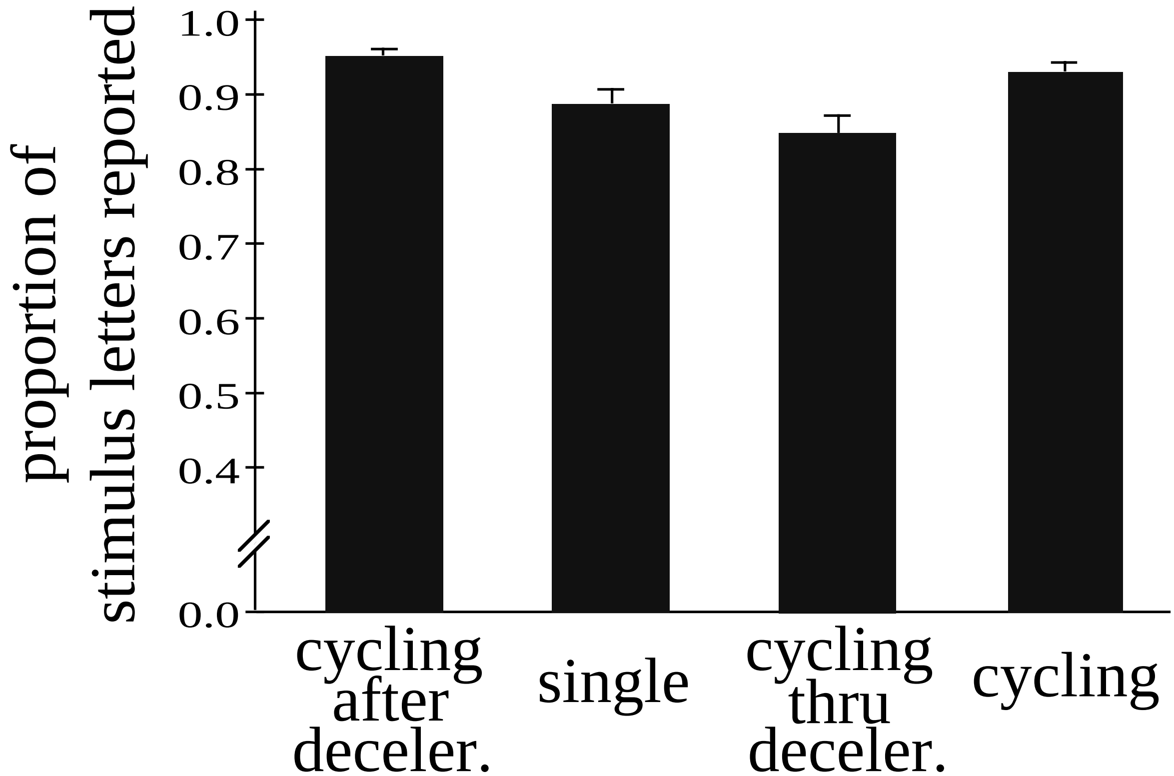


Figure 3, Holcombe et al.

As in the previous experiments, accuracy in reporting the letter identities was generally high, as shown in Figure 3. For the statistical tests, least-significant differences tests were used in the context of an ANOVA (with duration, stimulus type, subject, and all interactions as factors), which is equivalent to testing contrasts on the pairs of means. Interactions with presentation duration were not significant, and the pattern of results for the two presentation durations was the same, so the data were collapsed across duration. Letter identity accuracy was higher in the cycling condition than in the single-presentation condition, $p < .01$. It was slightly lower in the cycling through deceleration condition than in each of the other conditions, $ps < .05$. This may be explained by the possibility that in the cycling through deceleration condition, the participants tried to apprehend the sequence during the deceleration phase, and the extremely rapid presentation of the letters during the deceleration occasionally led to misidentification of the letters.

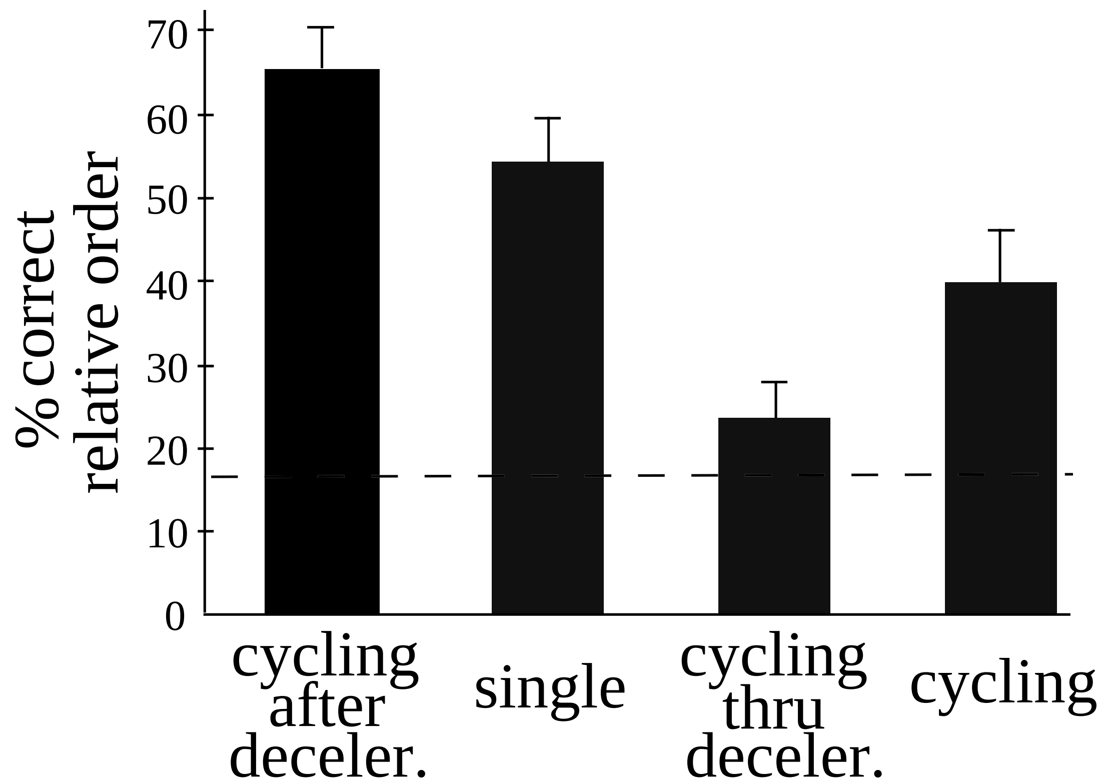


Figure 4, Holcombe et al.

The percentage of correct relative order responses in each condition is plotted in Figure 4. Statistical comparisons were performed using the same statistical analysis as in the previous paragraph, but with relative order accuracy substituted for identity accuracy as the dependent variable. Interactions with presentation duration were not significant so the data were collapsed across duration. Performance differed significantly ($p < .01$) in each of the conditions from each of the other conditions. Relative order reports were less accurate in the cycling condition than in the single-presentation condition, replicating the pattern of the previous experiments. In the new cycling through deceleration condition, accuracy was also lower than in the single-presentation condition.

The differences in order accuracy between the conditions were much larger than the differences in identity accuracy. In particular, the 5% difference between identity accuracy in the cycling through deceleration condition and the single-presentation condition can not explain the 30% difference in relative order accuracy. Comparison of Figures 3 and 4 shows that in general, the differences in number of stimulus letters reported were small compared to the differences in correct order reports.

Relative order accuracy in the cycling after deceleration condition was higher than in the single-presentation condition. However, if only trials in which participants reported all four stimulus letters correctly are considered, the difference between relative order accuracy in the cycling after deceleration condition and the single-presentation condition diminishes to insignificance. This suggests that the higher order accuracy in the cycling after deceleration condition was due to better memory for the items in the cycling after deceleration condition, perhaps due to repeated exposure. Nevertheless, relative order accuracy is clearly at least as good in the cycling after deceleration condition as in the single-presentation condition. In other words, when the letters were presented only after the fade-in period (the cycling after deceleration condition), performance was as good as after single presentation. That is probably the most revealing finding of this experiment.

This effect can be explained in terms of the degree to which an identifiable letter in the stream differs perceptually from the preceding items. In the cycling condition the first several items of the stream are presented too rapidly to be identified. Gradually the presentation rate decreases so that the sequence seems to fade in without a differentiated initial letter. In contrast, in the single-presentation and cycling after deceleration condition, the first identifiable letter is preceded only by line patterns, which are very different from the letters. This results in high order accuracy. We term the lower relative order accuracy in the cycling condition the "midstream order deficit" (MOD), because it occurs when observers are forced to apprehend the order of a rapid sequence from midstream: when there is little distinction between the beginning of the sequence and preceding items.

In contrast to the large role of the items preceding the sequence, the presence or absence of an end to the sequence different from the subsequent items seems to have had little if any effect. Compare the single-presentation condition, in which line patterns immediately followed the presentation of the sequence, to the cycling after deceleration condition, for which letters followed the sequence. Accuracy was comparably high in both conditions.

## General Discussion

Our results show that accuracy in perceiving the relative order of rapidly-presented items in a cycling sequence is much lower than if the sequence is presented just once. This effect is quite general, as it occurs with sounds (Warren et al., 1969), visual letters, shapes, and colors (O'Brien & Treisman, unpublished data). The accuracy of order reports in our experiments was correlated with the degree to which an identifiable letter is visually distinct from the preceding items. Accuracy was highest when the first identifiable letter constituted the onset of the stimulus (as shown in an unpublished experiment), and became progressively less accurate when the letter sequence was preceded by line patterns (Experiments 1 and 2), by digits (Experiment 2), and by other letters (the cycling conditions).

The effect occurred at relatively slow presentation rates for which observers' performance in reporting the letter identities is high, so it cannot be due to a problem in identifying the individual items. Furthermore, masking was approximately equated in the cycling and single-presentation conditions by the line-pattern masks, so this "midstream order deficit" (MOD) is unlikely to be due to masking. In addition, the MOD occurred even when a detection task was used, so it is probably not due to a memory limitation or to differences in decision criteria.

Previous theories and other attentional phenomena do not explain the MOD

Warren (1974) and Teranishi (1977) provided evidence that the threshold duration for reporting the order of cycling sound sequences is determined by the duration required to name the individual items as they are presented. Warren (1982) explained the lower threshold duration for singly-presented sequences by suggesting that after single presentation, persistence of the sequence in short-term memory allows naming at a slower rate. Although this account may be partly right, it does not explain why the difficulty with cycling sequences disappeared when the sequence had a clear beginning, i.e. when it was not presented during the deceleration phase (Experiment 3).

Indeed, in our experiments the deceleration or fade-in period was critical to obtaining lower relative order accuracy in the cycling condition than in the single-presentation condition. Yet in Warren et al.'s (1969) experiments with auditory stimuli, there was no fade-in or deceleration period, and still observers were much less accurate in the cycling condition than the single-presentation condition. This discrepancy may be due to a difference in observer strategy rather than the difference of auditory versus visual stimuli. In Warren et al.'s study each observer only participated in one trial. Thus the observers did not realize the importance of preparing to encode the order starting with the very first item. They probably waited for the stimulus to begin and then tried to encode the order, which left them in the position of trying to apprehend the order midstream, without an item distinct from the preceding items.

Audition researchers have sometimes attributed the difficulty in perceiving the order of cycling stimuli to "auditory streaming" (Bregman & Campbell, 1971). Auditory streaming (Miller and Heise, 1950) refers to the tendency of alternating acoustic stimuli to be experienced not as alternating, but rather as two continuous "streams" of sound, particularly when the sounds are cycled (Bregman, 1978). Bregman & Campbell (1971) found that perceiving order across streams was very difficult, since streaming items are not perceived to occur successively, and streaming may explain part of the difficulty of perceiving the order in Warren et al.'s (1969) experiment.

In the case of our experiments, however, subjective reports strongly suggest that streaming did not occur. Participants reported that they did not experience the letters as segregated into multiple streams. Rather, the letters were experienced as successively presented, even at the fastest rates. Furthermore, in unpublished experiments conducted by O'Brien & Treisman using visual stimuli, there was no effect of the hetereogeneity of the items, which would be expected to modulate the stream segregation if it had occurred (Bregman & Campbell, 1971). Accuracy with three colors or three shapes cycling was identical to accuracy with two colors and one shape or two shapes and one color.

The attentional blink (Raymond et al, 1992) is unlikely to play a part in the MOD. First, the attentional blink is not typically found at presentation rates as slow as those of the present experiments, and extrapolation from published data suggests that it would not occur at these rates (e.g., Moore, Egeth, Berglan, & Luck, 1996). Furthermore, the lack of a serial position effect in the single-presentation condition is strong evidence that a blink did not occur. The temporal order illusion documented by Reeves & Sperling (1986) also is unlikely to be a factor in the MOD, since Reeves & Sperling's errors only occurred at faster rates than those used in the present experiments, and they were accompanied by an attentional blink. Similarly, the rate of repetition is too slow for repetition blindness to play a role (Park & Kanwisher, 1994).

In sum, we have rejected accounts based on masking, memory limitations, decision criteria, naming time, perceptual streaming, the attentional blink, the order illusion discovered by Reeves & Sperling (1986), and repetition blindness.

## Theory and conclusions

In our data, the accuracy of order reports is correlated with the degree to which an identifiable letter is visually distinct from the preceding items. The presence of a salient end to the sequence is not important, and neither is the distinctiveness of the items after the first. This last point is evident from the single-presentation condition: the transitions from letter to letter were no more distinct than those in the cycling condition, yet order was accurately reported for the subsequent letters in the single-presentation condition. The overall lesson of the MOD appears to be that observers cannot easily begin their encoding with an arbitrarily selected item in a cycling stream. But why should this be?

Short-term memory is capacity-limited and encoding into explicit short-term memory requires attention (Sperling, 1960; Rensink, 2000). Hence there is a need to actively initiate encoding of order. In the case of our letter streams, this initiation act is distinct from the act of allocating spatial attention, as participants presumably allocate their spatial attention to the letter stream from its onset. When an item in a sequence is marked by a perceptual or categorical discontinuity, an automatic or exogenous segmentation occurs, just as a visual onset or change can pull exogenous spatial attention to a new location. We believe that such a discontinuity allows rapid initiation of order encoding. However, in the cycling condition, engaging attention on an arbitrary item requires an act of endogenous segmentation which appears to have a low temporal resolution (relative to exogenous attention and other visual processes), just as endogenous attention has a coarse *spatial* resolution (He, Cavanagh, & Intriligator, 1996). If engaging endogenous attention on an item within a stream takes 100 to 200 ms from the onset of the target item, as it does with a spatially distinct item (Nakayama & Mackeben, 1989), the target must last long enough to allow it to be processed once attention has reached it. If instead it is succeeded by another item in the same location, the whole attempt to initiate the encoding of order will have to restart, and is likely to fail again for the same reasons. Experiments on sequential task-switching certainly suggest long delays in resetting attention (Rogers & Monsell, 1994; Allport, Styles, & Hsieh, 1994), although these tasks are not directly comparable.

It is possible that midstream deficits are not confined to order judgments. In an experiment by Holcombe & Cavanagh (1999) which also presented stimuli in a rapid visual stream, two conditions were compared: in one a leftward-tilted Gabor adjacent to a red patch rapidly alternated with a rightward-tilt Gabor adjacent to a green patch. In the other condition, the pairing was reversed: leftward-tilt and green alternated with rightward-tilt and red. The fastest rate at which observers could correctly discriminate between these conditions in 75% of trials was approximately 180 msec/stimulus. After correcting for the difference in chance level, this threshold is similar to the temporal order threshold in the cycling condition, and may also be explained by the long latency of endogenous attentional engagement. In this case, a voluntary attention shift between the color patch and the Gabor stimulus may have been necessary within each frame of the sequence to encode both of the adjacent items before the next stimulus was presented. If the stimuli were replaced before the switch was completed, participants would be unable to determine the color-orientation pairing.

Outside the laboratory, the MOD may occur during the viewing of modern television programs or commercials which utilize rapid successive cuts. The present work suggests that at rates faster than about three per second, observers will find it difficult to accurately apprehend the order of the scenes, even though they may momentarily comprehend each individual scene (Potter, 1976; 1993). The well-studied phenomena of the attentional blink and change blindness (Rensink, O’Regan, & Clark, 1997) have revealed severe limitations on our ability to remember what we see. The midstream order deficit (as well as possible variants of it) merits further study, since it illustrates yet another limitation on our ability to retain what we see.

References

Allport, D. A., Styles, E . A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), Attention and Performance XV, 421-452. Hillsdale, NJ: Erlbaum.

Bregman, A. S. (1978). Auditory streaming is cumulative. Journal of Experimental Psychology: Human Perception and Performance, 4, 380-387.

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. Journal of Experimental Psychology, 89, 244-249.

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(1), 109-127.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. Nature, 383, 334-337.

Holcombe, A. O. & Cavanagh, P. (1999). The temporal resolution of visual feature binding. Paper presented at the 7th Annual Workshop on Object Perception and Memory, Los Angeles, CA.

Kanwisher, N. G., Kim, J. W., & Wickens, T. D. (1996). Signal detection analyses of repetition blindness. Journal of Experimental Psychology: Human Perception and Performance, 22(5), 1249-1260.

Miller, G. A., & Heise, G. A. (1950). The trill threshold. Journal of the Acoustical Society of America, 22, 637-638.

Moore, C. M., Egeth, H., Berglan, L. R., & Luck, S. J. (1996). Are attentional dwell times inconsistent with serial visual search? Psychonomic Bulletin & Review, 3(3), 360-365.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. Vision Research, 29, 1631-1647.

Park, J., & Kanwisher, N. (1994). Determinants of repetition blindness. Journal of Experimental Psychology: Human Perception and Performance, 20(3), 500-519.

Potter, M. C. (1976). Short-term conceptual memory for pictures. Journal of Experimental Psychology: Human Learning and Memory, 2(5), 509-522.

Potter, M. C. (1993). Very short-term conceptual memory. Memory & Cognition, 21(2), 156-161.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? Journal of Experimental Psychology: Human Perception and Performance, 18(3), 849-860.

Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. Psychological Review, 93(2), 180-206.

Rensink, R. A. (2000). The dynamic representation of scenes. Visual Cognition, 7, 17-42.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. Psychological Science, 8, 368-373.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. Journal of Experimental Psychology: General, 124, 207-231.

Sperling, G. (1960). The information available in brief presentations. Psychological Monographs, 74, entire issue.

Teranishi, R. (1977). Critical rate for identification and information capacity in hearing system. Journal of the Acoustical Society of Japan, 33, 136-143.

Warren, R. M. (1974). Auditory temporal discrimination by trained listeners. Cognitive Psychology, 6, 237-256.

Warren, R. M. (1982). Auditory perception: A new synthesis. New York: Pergamon Press.

Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. (1969). Auditory sequence: Confusion of patterns other than speech or music. Science, 164, 586-587.

Author Notes

This research was supported by NIH grant MH45245 to NK, NSF Grant SBR-9511633 to AT, an NIMH NRSA graduate fellowship to AOH, and NIH grant EY092582 to Patrick Cavanagh. We are indebted to Rik Henson for sharing his knowledge of theories of short-term order memory. We thank Etienne Benson, Jared Bush, and Damian Stanley for their help conducting the experiments, and Charles Stromeyer, Patrick Cavanagh, and two anonymous reviewers for useful comments on earlier manuscripts. Portions of this research were reported at the 1996 annual meeting of the Association for Research in Vision and Opthalmology, Ft. Lauderdale, Florida, the 1998 meeting of the Eastern Psychological Association, Boston, Massachusetts, and the 1998 International Conference on Short-term Memory, Quebec, PQ, Canada.

Send correspondence to A. O. Holcombe, 2nd floor boxes, 33 Kirkland St, Cambridge, MA 02138, or send e-mail to *holcombe@wjh.harvard.edu*. NK's e-mail address is *ngk@psyche.mit.edu* and AT's is *treisman@phoenix.princeton.edu*.

Table 1

Mean estimated alpha (sensitivity), *g* (percent of trials that observer guessed that the target was presented when he did not apprehend it) and percent correct for the three conditions of Experiment 4.

| | single- line pattern masked | single- digit masked | cycling |
|---|---|---|---|
| alpha | 0.85 | 0.76 | 0.71 |
| *g* | 0.2 | 0.26 | 0.32 |
| % correct | 84 | 78 | 74 |

Figure Captions

Figure 1. A schematic of the conditions of the experiments.  The first few items were presented very quickly and the rate gradually slowed until the final duration was reached at the ninth item.  All items were presented in the same spatial location.  In the single presentation condition the letters appeared after the deceleration in a randomly chosen position.  Experiment 1 used only the single and cycling conditions.  Experiment 2 used the single and cycling conditions as well as a variant of the single condition in which the line pattern masks were replaced by digits.  Experiment 3 used all of the pictured conditions.

Figure 2.  Mean probability of reporting the correct relative order at each duration for the cycling and single presentation conditions of Experiment 1.  If observers knew which four letters were presented, but knew nothing about their order, their relative order accuracy would be 16.6%, as there are six possible relative orders for four items.  This is represented by the dotted line.

Figure 3. Proportion of stimulus letters reported for each condition of Experiment 3.  “deceler.” = deceleration.  Standard error bars are based on subject means (which is different from the ANOVA error term used for the statistical tests).

Figure 4.  Proportion of correct relative order reports for each of the conditions of Experiment 3. “deceler.” = deceleration.  Standard error bars are based on subject means (which is different from the ANOVA error term used for the statistical tests).  The dotted line represents accuracy if the letters presented were known but the order was guessed.

.