Supplemental data from: Backward recall and the word length effect

AIMÉE M. SURPRENANT¹, TAMRA J. BIRETA², MARK A. BROWN¹, ANNIE JALBERT¹, GERALD TEHAN³, and IAN NEATH¹

¹Memorial University of Newfoundland, ²The College of New Jersey, ³The University of Southern Queensland

Surprenant et al. (2011) report two experiments which examine whether word length effects are observable with backward serial recall. Additional data and details are reported here.

The word length effect refers to the finding that lists of short (i.e., 1 syllable) words are recalled better than otherwise comparable lists of longer (i.e., multisyllabic) words (Baddeley, Thomson, & Buchanan, 1975).

Although this effect has been observed with numerous different tests (e.g., forward serial recall, free recall, serial recognition, single item probe recognition, complex span), there are few studies that examine whether the word length effect is observable with backward recall.

Only two models -- the Primacy Model (Page & Norris, 1998, 2003) and the Feature Model (Nairne, 1990; Neath & Nairne, 1995; Neath, 2000) -- address both word length and recall direction, and both predict that the word length effect will remain with backward recall.

Bireta et al. (2010) reported an experiment in which the subjects did not know which recall direction would be tested until after list presentation. This ensured that whatever processing was performed during list presentation, it was likely to be the same for both recall directions. Bireta et al. found that a word length effect was observed with forward recall but recall of short (1 syllable) and long (3 syllables) words was equivalent with backward recall.

Surprenant, Bireta, Brown, Jalbert, Tehan, and Neath (2011) reported two experiments which also examined recall direction and the word length effect. This paper presents additional data and analyses not included due to space issues.

Experiment 1

In Experiment 1, Surprenant et al. (2011) presented seven-item lists of 1- or 3-syllable words for 1 s each on a computer screen. One group of subjects were informed of the recall direction prior to list presentation. The other group were informed after the list had been presented.

Table 1 gives the means plotted in Figures 1 and 2 of Surprenant et al. (2011), as well as the corresponding standard deviations.

Figure 1 shows the proportion of short and long words correctly recalled in forward or backward order at each serial position as a function of whether the subjects knew or did not know the recall direction prior to list presentation.
Table 1: Mean accuracy plotted in Figure 1 and mean output time plotted in Figure 2 of Surprenant et al. (2011), as well as corresponding standard deviations.

<table>
<thead>
<tr>
<th></th>
<th>Direction Known</th>
<th>Direction Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Forward</td>
<td>Backward</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Stdev</td>
</tr>
<tr>
<td>Short</td>
<td>0.609</td>
<td>0.147</td>
</tr>
<tr>
<td>Long</td>
<td>0.519</td>
<td>0.155</td>
</tr>
<tr>
<td>Short</td>
<td>1324</td>
<td>338.8</td>
</tr>
<tr>
<td>Long</td>
<td>1431</td>
<td>338.6</td>
</tr>
</tbody>
</table>

Figure 1: Proportion of short and long words correctly recalled at each serial position as a function of recall direction and whether subjects knew the recall direction prior to list presentation (Experiment 1).

A 2 word length (short or long) × 2 recall direction (forward or backward) × 2 knowl-
edge type (advance knowledge or no advance knowledge) × 7 serial positions analysis of variance (ANOVA) was conducted.

There was a significant main effect of serial position, $F(6,288) = 64.681$, $MSE = 0.037$, $p < .001$. There was a significant interaction between serial position and recall direction, $F(6,288) = 85.699$, $MSE = 0.025$, $p < .001$, as can be seen in Figure 1.

The interaction between serial position and length just failed to reach conventional levels of significance, $F(6,288) = 2.090$, $MSE = 0.022$, $p = 0.054$. This was due to a larger difference between recall of short and long items at end positions than at interior positions. For example, the difference at Position 1 was 0.722 vs. 0.624 compared to a difference at Position 4 of 0.472 vs. 0.418.

None of the remaining interactions were significant: For the interaction between serial position and knowledge, $F(6,288) < 1$; for the interaction between length, position, and direction, $F(6,288) < 1$; for the interaction between length, position, and group, $F(6,288) = 1.181$, $MSE = 0.019$, $p > 0.30$; and for the four-way interaction, $F(6,288) = 1.162$, $MSE = 0.019$, $p > 0.30$.

Figure 2: Mean output time (ms) for short and long words correctly recalled as a function of recall direction and output serial position and whether recall direction was known or unknown prior to list presentation (Experiment 1).
Another way of examining the data is to consider the number of participants who recalled more short than long items (Beaman, Neath, & Surprenant, 2008; Logie, Della Sala, Laiacona, Chalmers, & Wynn, 1996). When recall direction was known, 18 people recalled more short than long words with forward recall, 5 people recalled more long than short words, and 2 ties. This is significant by a sign test, \( p < .05 \). In contrast, with backward recall, only 13 people recalled more short than long words, with 11 showing the reverse and 1 tie. This is not significant by a sign test, \( p > .80 \). When recall direction was unknown, sign tests yielded the same pattern of results.

For forward recall, 18 people showed a word length effect and 7 showed a reverse effect, \( p < .05 \). For backward recall, 15 showed a word length effect, 9 showed a reverse effect and there was 1 tie, \( p > .30 \).

The computer recorded when each response was made. Figure 2 shows the mean output time for each correctly recalled short and long word as a function of recall direction and whether recall direction was known or unknown prior to list presentation. Figure 3 shows the mean cumulative output time and the appendix shows cumulative output times for each subject.

A 2 word length (short or long) \( \times \) 2 recall direction (forward or backward) \( \times \) 2 knowledge type (advance knowledge or no advance knowledge) \( \times \) 7 serial positions analysis of variance (ANOVA) was conducted on mean output time (as opposed to cumulative output time).

The main effect of length was not significant, \( F(1,48) = 2.818, MSE = 771988, p = 0.10 \). The main effect of recall direction was significant, \( F(1,48) = 61.642, MSE = 1955939, p < .001 \). The main effect of knowledge type was significant, \( F(1,48) = 6.218, MSE = 3479498, p < .05 \). There was a significant main effect of serial position, \( F(6,288) = 125.217, MSE = 1396391, p < .001 \).

There was a significant interaction between serial position and recall direction, \( F(6,288) = 15.752, MSE = 511233, p < .001 \). As can be seen in Figure 2, this was due primarily to the almost linear function seen with backward recall compared to a large decrease in output time from Position 1 to Position 2 with forward recall.
No other interactions were significant. The interaction between serial position and group was not significant, $F(6,288) < 1$. The interaction between serial position and length was not significant, $F(6,288) < 1$. The interaction between length, position, and group was not significant, $F(6,288) = 1.709, MSE = 754116, p > 0.10$. The interaction between length, position, and direction was not significant, $F(6,288) = 1.364, MSE = 840768, p > 0.40$. The four-way interaction was also not significant, $F(6,288) = 1.020, MSE = 824557, p > 0.40$.

<table>
<thead>
<tr>
<th></th>
<th>Direction Known</th>
<th>Direction Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Forward</td>
<td>Backward</td>
</tr>
<tr>
<td></td>
<td>Mean  Stdev</td>
<td>Mean  Stdev</td>
</tr>
<tr>
<td>Short</td>
<td>0.559 0.098</td>
<td>0.559 0.085</td>
</tr>
<tr>
<td>Long</td>
<td>0.493 0.110</td>
<td>0.543 0.081</td>
</tr>
<tr>
<td>Short</td>
<td>1404 555.1</td>
<td>1661 482.0</td>
</tr>
<tr>
<td>Long</td>
<td>1451 310.6</td>
<td>1675 529.0</td>
</tr>
</tbody>
</table>

Table 2: Mean accuracy plotted in Figure 3 and mean output time plotted in Figure 4 of Surprenant et al. (2011), as well as corresponding standard deviations.

**Experiment 2**

In Experiment 2, Surprenant et al. (2011) again presented seven-item lists of 1- or 3-syllable words for 1 s each on a computer screen. One group of subjects, Forward/Backward, received all forward recall trials in the first block and then all backward recall trials in the second block. A second group, Backward/Forward, received all backward recall trials in the first block and then all forward recall trials in the second block. Table 2 gives the means plotted in Figures 3 and 4 of Surprenant et al. (2011), as well as the corresponding standard deviations.

Figure 4 shows the proportion of short and long words correctly recalled in forward or backward order at each serial position as a function of whether the subjects knew or did not know the recall direction prior to list presentation.

A 2 word length (short or long) × 2 recall direction (forward or backward) × 2 recall direction order (forward then backward or no backward then forward) × 7 serial positions analysis of variance (ANOVA) was conducted.

There was a significant effect of serial position, $F(6,288) = 67.639, MSE = 0.078, p < 0.001$. The interaction between position and group was significant, $F(6,288) = 2.234, MSE = 0.035, p < 0.05$. As can be seen in Figure 3, the curves are slightly steeper and have more “recency” with backward recall than forward recall. The interaction between position and direction was significant, $F(6,288) = 141.983, MSE = 0.019, p < 0.001$, as they were mirror images of each other.
None of the remaining interactions were significant: For the interaction between position and length, $F(6,288) = 1.258$, $MSE = 0.016$, $p > 0.25$; for the interaction between position, length, and group, $F(6,288) = 1.335$, $MSE = 0.016$, $p > 0.20$; for the interaction between direction, length, and position, $F(6,288) = 1.658$, $MSE = 0.018$, $p > 0.10$; and for the four-way interaction, $F(6,288) < 1$.

As in Experiment 1, we looked at the number of participants who recalled more short than long items. In the Forward First condition, 18 people recalled more short than long words with forward recall, with 7 recalling more long than short items. This is significant by a sign test, $p < .05$. In contrast, with backward recall, 15 people recalled more short than long words with forward recall, and 10 showed the reverse. This is not significant by a sign test, $p > .40$. The pattern is the same in the Backward First condition. For forward recall, 20 people showed a word length effect and 4 showed a reverse effect, with 1 tie, $p < .05$. For backward recall, 16 showed a word length effect, 8 showed a reverse effect and there was 1 tie, $p > .15$. 

Figure 4: Proportion of short and long words correctly recalled at each serial position as a function of recall direction and whether subjects received all the forward recall trials first and backward recall trials second or the reverse (Experiment 2).

---

**Figure 4:** Proportion of short and long words correctly recalled at each serial position as a function of recall direction and whether subjects received all the forward recall trials first and backward recall trials second or the reverse (Experiment 2).
As in Experiment 1, the computer recorded when each response was made. Figure 5 shows the mean output time for each correctly recall short and long word as a function of whether the first block was forward or backward. Figure 6 shows the cumulative output time and the appendix shows individual cumulative output times.

A 2 word length (short or long) × 2 recall direction (forward or backward) × 2 block order (forward first or backward first) × 7 serial positions analysis of variance (ANOVA) was conducted on mean output time (as opposed to cumulative output time).

The main effect of length was not significant, \( F(1,48) = 1.511, \text{ MSE } = 735016, \ p > 0.20 \). The main effect of recall direction was significant, \( F(1,48) = 57.945, \text{ MSE } = 1.660 \times 10^6 \). The main effect of block order was not significant, \( F(1,48) < 1 \). The main effect of position was significant, \( F(6,288) = 92.229, \text{ MSE } = 9.205 \times 10^7, \ p < 0.001 \).
Figure 6: Mean cumulative output time from Experiment 2 as a function of whether forward recall was in the first block and backward was in the second block (left panel) or backward recall was in the first block and forward was in the second block (right panel). Note: SF = short forward; LF = long forward; SB = short backward; LB = long backward.

The position by direction interaction was significant, $F(6,288) = 19.461$, $MSE = 879982$, $p < 0.001$. The position by group interaction was also significant, $F(6,288) = 2.838$, $MSE = 998002$, $p < .05$, and both of these interactions are apparent in Figure 4.

The position by length interaction was not significant, $F(6,288) = 1.016$, $MSE = 733960$, $p > 0.40$. None of the other interactions were significant; indeed, for all remaining interactions, $F(6,288) < 1$.

**GENERAL DISCUSSION**

The Primacy Model (Page & Norris, 1998) assumes backward recall is done by performing a series of forward recalls. That is, the subject recalls each item in the list, beginning with the 1st item, until the most active item is reached, in which case it is output as the response. Then, the subject again recalls the list from the 1st item but now stops one item earlier and outputs that item. The output time data here clearly rule out such an account.

We can take the data from the forward recall first condition of Experiment 2 as an indicator of what typical forward recall is like. The best fitting straight line for short words had slope = 1.203 and intercept = 1.675; this resulted in $r^2 = 0.995$. The best fitting straight line for long words had slope = 1.254 and intercept = 1.817; this resulted in $r^2 = 0.987$. Thus, mean output time for forward recall is linear. These are shown in Figure 7.

We can take the data from the backward recall first condition of Experiment 2 as an indicator of what typical backward recall is like. The best fitting straight line for short words had slope = 0.987 and intercept = 1.543; this resulted in $r^2 = 0.989$. The best fitting straight line for long words had slope = 1.039 and intercept = 1.831; this resulted in $r^2 = 0.985$. Thus, mean output time for backward recall is linear. However, it is no slower than forward recall to initiate: 2.804 vs. 2.381 for short words, $t(48) < 1$, and 2.908 vs. 2.675 for long words, $t(48) < 1$. According to the Primacy Model, backward recall must take longer than forward recall, due to the multiple run throughs.

Although it is always possible that mean values, such as used here, differ from individual subject patterns, it is unlikely that
individual subjects are showing the pattern predicted by the Primacy Model because of the numerically faster RTs for backward recall. In addition, inspection of the individual subject data suggests very few subjects used the “repeatedly run through” strategy.

![Figure 7: Cumulative output time (in seconds) when short and long words when subjects received all the forward recall trials first (forward) or received all the backward recall trials first (backward) and the best fitting straight lines.](image)

**REFERENCES**


Appendix A

Individual cumulative output times for Experiment 1 of Surprenant et al. (2011) when the recall direction was known. The y-axis is mean cumulative output time in seconds, and the x-axis is output position.
Individual cumulative output times for Experiment 1 of Surprenant et al. (2011) when the recall direction was unknown. The y-axis is mean cumulative output time in seconds, and the x-axis is output position.
Individual cumulative output times for Experiment 2 of Surprenant et al. (2011) when forward recall was in the first block and backward recall was in the second block. The y-axis is mean cumulative output time in seconds, and the x-axis is output position.
Individual cumulative output times for Experiment 2 of Surprenant et al. (2011) when backward recall was in the first block and forward recall was in the second block. Some subjects did not recall any items correctly in a specific position, so those and subsequent data points are missing. The y-axis is mean cumulative output time in seconds, and the x-axis is output position.