Recall of Television Versus Print News: Retesting the Semantic Overlap Hypothesis
Juliette H. Walma van der Molen *; Marlies E. Klijn *
* University of Amsterdam.

Online Publication Date: 01 March 2004

To cite this Article van der Molen, Juliette H. Walma and Klijn, Marlies E.(2004)'Recall of Television Versus Print News: Retesting the Semantic Overlap Hypothesis', Journal of Broadcasting & Electronic Media, 48:1, 89 — 107
To link to this Article: DOI: 10.1207/s15506878jobem4801_5
URL: http://dx.doi.org/10.1207/s15506878jobem4801_5
Recall of Television Versus Print News: Retesting the Semantic Overlap Hypothesis

Juliette H. Walma van der Molen and Marlies E. Klijn

The study tested the validity of 2 rival explanations for the relative effectiveness of television and print news: the reading control versus the semantic overlap hypothesis. Participants (N = 100) were either exposed to television stories containing different text-picture formats or to printed versions. The study thus combined the designs of inter- and intra-media comparison studies of news recall. Results of a cued-recall test were in favor of the semantic overlap hypothesis. They suggest that the recall advantage of either television or print depends on the level of overlap between verbal and visual information in the television presentation.

Over the last few decades, public opinion surveys have shown a steady increase in the number of people that rely on television as their main source of news information instead of newspapers or other news media (Roper Reports, 1995). Apart from using television news more often than other news media, many people also believe television news is more credible than other news sources (e.g., Brosius, Donsbach, & Birk, 1996). Given the fact that most people in the Western world now have access to television news around the clock, reliance on it may increase even further in future years. The dominant use of television as a news source, however, has raised concern among communication researchers because the common presumption is that television is too fast-paced to be an effective information medium (e.g., Singer, 1980). Compared with television, print is assumed to be considerably more effective because print offers more opportunities to exercise control over the processing of information than television does. Unlike viewers, readers can digest the news at their own pace, reread passages, and check details—all of which can facilitate memory for news information (e.g., Gunter, 1987).
The relative ineffectiveness of television news compared with print news has been supported by a number of media comparison experiments that compared memory for television news stories with memory for printed versions of the television narratives. In most of these studies, participants indeed remembered more from the printed news than from the television stories (e.g., DeFleur, Davenport, Cronin, & DeFleur, 1992; Facorro & DeFleur, 1993; Furnham & Gunter, 1985, 1987; Gunter & Furnham, 1986; Gunter, Furnham, & Gietson, 1984; Gunter, Furnham, & Leese, 1986; Wicks & Drew, 1991; Wilson, 1974). Unfortunately, however, none of these television-print news comparisons investigated whether the use of control techniques, such as rereading passages, really was responsible for the observed superior recall of print. Therefore, the notion that print is inherently better at conveying news information due to its processing advantages still remains a weakly supported assumption. We will refer to this assumption as the reading control hypothesis.

Another explanation is that television is not necessarily an inferior medium for conveying news information, but that its effectiveness depends on the manner in which the news is visually presented. Walma van der Molen and van der Voort (2000a) proposed that the relative effectiveness of television compared with printed news is determined to a large extent by the degree to which the verbal and visual content of television news shows semantic overlap. Based on limited attentional capacity theories (e.g., Grimes, 1990, 1991), it was argued that when verbal and visual information do not correspond, viewer’s attentional capacity is exceeded and that priority will be given to processing of visual information. This implies that when text and pictures are not semantically related, the main message of a television news story, which is usually presented through the audio channel, will be lost to most viewers. Based on Paivio’s (1986; Clark & Paivio, 1991) dual-coding hypothesis, which posits that recall of verbal information can be enhanced by the extra mnemonic support of related visual information, Walma van der Molen and van der Voort argued that television can facilitate information transfer if relevant pictures were added to the verbal message. In an experiment with adults and children, using regular adult and special children’s news stories, they showed that both adult and child viewers of the children’s news (marked by good semantic overlap) recalled more than readers of printed versions. In addition, they found that the advantage of television over print was due to the relatively large amount of semantically related verbal and visual information in the children’s news items.

Contrary to Walma van der Molen and van der Voort’s (2000a) study, the majority of television-print experiments that found print to be a more effective news medium, used regular “adult” television newscasts as stimulus material (Furnham & Gunter, 1985, 1987; Gunter & Furnham, 1986; Gunter, Furnham, & Gietson, 1984; Gunter, Furnham, & Leese, 1986; Wicks & Drew, 1991). Although none of these studies provided information about the correspondence between verbal and visual information in the television items that were used, it may be readily supposed that the semantic overlap between text and pictures was generally low. It has been widely
noted that adult television news is dominated by so-called standard news pictures (e.g., politicians arriving in limousines) and by pictures that are merely used to create an illusion of authenticity and actuality (Brosius et al., 1996). Although these pictures may attract attention, they are usually not semantically related to the verbal message and as a consequence may distract the viewer from the verbal message. In addition, two of the television-print news experiments used television items that depicted only the newsreader ("talking-head only" items) (DeFleur, Davenport, Cronin, & DeFleur, 1992; Facorro & DeFleur, 1993). Although the literature on text-picture relations fails to show consensus on the effectiveness of talking heads (see Lang, 1995), it is clear that talking heads do not provide the viewer with pictures that are semantically related to the verbal message. An alternative explanation for the inferiority of television news compared with printed news may thus be that previous television-print news comparisons used television stories that lacked sufficient semantic overlap between text and pictures. The study by Walma van der Molen and van der Voort was the first to suggest that if television stories contain a greater amount of semantically related audiovisual information, television can be more effective than print. We will refer to this assumption as the semantic overlap hypothesis.

The importance of a high level of semantic overlap, or audiovisual redundancy, is not a new idea in communication research. Several intra-media comparison studies, which compared television news stories marked by indirect text-picture relations with comparable television stories marked by direct text-picture correspondence, have shown that a close semantic correspondence between textual and pictorial information may enhance news recall, whereas indirect text-picture relations or text-picture divergences could impair learning (e.g., Drew & Grimes, 1987; Graber, 1990; Reese, 1984; Son, Reese, & Davie, 1987). Curiously enough, however, the findings from these intra-media comparisons have not been considered in the interpretation of inter-media comparisons of television and print news. To fill this void in research, the present study combined the designs of previous inter- and intra-media comparison experiments. To test the validity of the semantic overlap hypothesis, good and poor semantic-overlap formats of regular television news were varied and memory for these versions was compared with memory for the same printed information. To test the validity of the semantic overlap hypothesis, good and poor semantic-overlap formats of regular television news were varied and memory for these versions was compared with memory for the same printed information. It was this test that was undertaken in the present study.
Hypotheses

In the study presented here, two television-print news comparisons were investigated to test the two opposing hypotheses. As will be discussed in more detail below, it was assumed that if print is inherently a better medium to convey news information than television, readers should remember more than viewers, regardless of whether television stories are supplemented with semantically related or semantically unrelated pictures. On the other hand, if the relative effectiveness of television and print is predominantly determined by the degree of semantic overlap in the television presentation, viewers of television stories containing good semantic overlap should remember more than readers, whereas viewers of television stories containing poor semantic overlap should remember less than readers.

Table 1
Expected Direction of Recall Differences Between Television and Print as a Function of Type of Television Format for the Reading Control and Semantic Overlap Hypothesis

<table>
<thead>
<tr>
<th>Television Format</th>
<th>Poor Semantic Overlap</th>
<th>Good Semantic Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Control Hypothesis</td>
<td>TV &lt; Print (H₁)</td>
<td>TV &lt; Print (H₂)</td>
</tr>
<tr>
<td>Semantic Overlap Hypothesis</td>
<td>TV &lt; Print (H₁)</td>
<td>TV &gt; Print (H₃)</td>
</tr>
</tbody>
</table>

Note. The greater-than or less-than signs mean that there is a statistically significant recall advantage or disadvantage, respectively, of television compared with print.

Table 1 summarizes the expected outcomes of the two television-print comparisons in the case that the reading control or the semantic overlap hypothesis holds true. In either case, we expected to find that:

H₁: Print is remembered better than television if the television presentation is marked by poor semantic overlap between text and pictures.

Hypothesis 1 is consistent with both the reading control and the semantic overlap hypothesis because the observed superiority of print may be due both to the opportunities print offers for efficient information processing and to the relatively low level of semantic overlap in the television presentation, which might divert viewers' attention away from the verbal message.

The decision determining the plausibility of either the reading control or the semantic overlap hypothesis, therefore, depends on the outcomes obtained for the television-print comparison that compares memory for printed news with memory for television news containing good semantic overlap between text and pictures. It should be noted, however, that an explicit decision about the validity of either
hypothesis is possible only if the complete pattern of results for the two television-print comparisons is in accordance with one or the other explanation.

If the superior recall of print news is mainly due to the opportunities print offers to control the processing of information, readers should remember more than viewers, regardless of the type of television presentation. If the superiority of printed news that was observed in previous television-print comparisons is indeed best explained by the reading control hypothesis, it should therefore be expected that:

\[ H_2: \text{Print is remembered better than television if the television presentation is marked by good semantic overlap between text and pictures.} \]

However, if the relative effectiveness of television compared with printed news is best explained by the degree to which the verbal and visual television content shows semantic overlap, then viewers should remember more than readers if the television presentation is marked by good semantic overlap. If the semantic overlap hypothesis holds true, it should therefore be expected that:

\[ H_3: \text{Television is remembered better than print if the television presentation is marked by good semantic overlap between text and pictures.} \]

Method

Participants

The study was conducted with a sample of 100 university students (34 males and 66 females, mean age = 20.3 years) from the Department of Communication, University of Amsterdam, the Netherlands. All students were required to participate in the experiment as part of a broader project on research methodology. Based on their answers on a self-report questionnaire, it was concluded that none of the participants had visual, speech, or hearing disorders.

Design

Students were randomly assigned to one of two experimental conditions: (a) a television condition in which participants watched three television news stories that had been labeled as containing poor semantic overlap between text and pictures and three television stories that had been labeled as containing good semantic overlap, and (b) a print condition in which participants read literal transcripts of the television narratives of each of the six news stories. The study thus consisted of a 2 (television vs. print) x 2 (poor vs. good semantic overlap) design with medium condition as a between-subjects factor and degree of semantic overlap as a within-subjects factor. In each of the two experimental medium conditions, 50 students participated.
Stimulus Materials

In each medium condition, participants were presented with six news stories. The television news stories had been chosen from the 8 O’clock News, the main evening news program that is broadcast every evening by the Dutch Broadcasting Corporation. From a broad collection of news stories broadcast between 1996 and 1999 that is available at the Department of Communication, an initial selection of news stories was made, based on four criteria. First, in order to minimize the likelihood that participants had previous knowledge of the news stories, only stories were selected that had been broadcast at least one and a half years prior to the experiment. Second, for the same reason, the selected stories involved isolated news events that had not attracted repeated media attention. Third, the news stories did not contain overt violent or otherwise emotionally arousing pictures because thus far it is not clear how arousal would influence the interaction between medium and degree of semantic overlap that was investigated in the present study. Fourth, the television narratives were comprehensible without the accompanying pictures, so that literal transcripts of the television narratives could be used in the print conditions, without additions or deletions.

After the initial selection, stories were content analyzed to establish their degree of semantic overlap between text and pictures. This was done using a procedure that is described by Walma van der Molen (2001). First, news items were divided into shots. For each shot, the concomitant pictorial and verbal information was established and it was coded whether the verbal information was supplemented with (a) directly related pictures, (b) indirectly related (or “standard”) pictures, (c) divergent pictures, or (d) talking head pictures. Subsequently, within each news item, the total amount of time devoted to each type of text-picture correspondence was calculated by summing up the duration of shots belonging to each correspondence category and dividing that by the total duration of the particular news item. Although it proved difficult to find stories that contained a fair amount of redundant audiovisual information, analyzing news stories was continued until three stories were found that could be labeled as containing good semantic overlap and three stories that could be labeled as containing poor semantic overlap between text and pictures. Based on the results of Walma van der Molen and van der Voort’s (2000a) study, which showed that supplementing about 40% of the verbal information with semantically related pictures may be enough to bring about an overall recall advantage of television over print, stories were labeled as containing good semantic overlap when they assigned at least 40% of their time to verbal information that was supplemented with directly related visuals. Stories that were labeled as containing poor semantic overlap predominantly used indirectly related and talking head pictures.

The stories that were marked by good semantic overlap covered the following events: (a) “Dutch marines found to smuggle drugs from Aruba,” (b) “Festivities around the re-opening of the Berlin Reichstag building,” and (c) “Elections for new
parliaments in Scotland and Wales." The following stories contained poor semantic overlap: (d) "Political problems in France," (e) "Dutch nurse accused of killing several people in a nursery home," and (f) "Chinese army practices near Taiwan." The stories thus covered different topic categories (both national and international and both political and nonpolitical) and an attempt was made to balance topic categories across the good and poor semantic overlap conditions. Furthermore, to prevent order effects, stories were presented in two different sequences. Half of the participants in the television and print conditions were presented with the news items in the order: a, d, b, e, c, f. The other half watched or read the stories in the order: e, c, d, b, f, a.

For the two television sequences, news items were combined into the format of a regular newscast, including the news program's leader and credit titles. Each television news story began with a short summary delivered by a newsreader and continued with film footage accompanied by spoken commentary. The total duration of the resulting television newscast was 13.4 minutes. As was done for topic categories, length of news items was balanced across the good and poor semantic overlap conditions. Stories (a) and (f) were shorter news items, taking up 1.5 and 1.4 minutes, respectively, while the other news items ranged in length between 2.0 and 2.3 minutes.

For the print condition, the literal transcripts of the news stories were presented in a newspaper format. Television titles that served to announce each television news story were used as headlines for the printed stories. Introductory commentary, delivered by a newsreader in the early part of each of the television stories, was transformed into a bold printed lead in the print versions. The remaining text of each television story was printed in two columns. Thus, an attempt was made to approximate the natural format of newspapers as closely as possible, while keeping the information conveyed by the printed transcripts identical to the verbal information conveyed by the television stories.

**Procedure**

Participation in the experiment took place in vacant rooms at the university building in groups of about 12 students. Students were tested in groups, rather than individually, to improve the comparability between the present study and previous television-print news comparison studies, most of which also tested participants in groups. The experiment was carried out in 8 test sessions spread over 4 days. In each session, one experimental medium condition was executed. To reduce the possibility of time-of-day effects, television and print conditions were varied across morning and afternoon sessions. Two experimenters conducted the experiment. To control for possible experimenter effects, participation of the experimenters was varied across experimental conditions.

All participants were informed in advance that a test would follow their exposure to the news stories. Based on studies that showed that test expectation
increases mental effort, which in turn may increase learning (e.g., Eveland, 1997; Salomon, 1984; Salomon & Leigh, 1984), we assumed that this procedure would motivate our participants to exercise more control over their processing of the news information than they would in a situation without test expectation and that this procedure would stimulate learning. In previous media comparison experiments investigating adults' recall of television and print news (e.g., Facorro & DeFleur, 1993; Furnham & Gunter, 1985), participants also knew in advance that their knowledge of the news would be tested. We therefore decided not to use an instruction that prevents participants from expecting a test, although such an instruction could increase the ecological validity of the experiment. All participants were told: "As you know, today you are participating in an experiment." In the television condition the experimenter proceeded to say: "I have here a videotape of the 8 O'clock News that we are going to watch. Watch the tape quietly; when it is finished, I want you to complete a test." In the print conditions the experimenter said: "I have here some newspaper stories that I want you to read. Read the stories quietly at your own tempo; when you are finished reading, I want you to complete a test."

The participants who watched the news stories on television were exposed to the news for 13.4 minutes. The participants in the print condition were allowed to read the news stories at their own pace. Based on a series of experiments that showed that invested mental effort is positively associated with performance time (Tyler, Hertel, McCallum, & Ellis, 1979), we assumed that the expectation of a memory test would stimulate at least some readers to take more than 13.4 minutes to process the stories. Therefore, we did not maintain a constant exposure time across the television and print conditions because such a procedure could put some readers at a disadvantage and could prevent a meaningful media comparison. Reading time for the news stories ranged between 6 and 17 minutes, with an average of 10.2 minutes. Given the fact that the participants in our experiment were university students who could all be expected to be fairly proficient readers, reading time was used in the analyses as an indicator of invested mental effort. For purposes of analyses, a distinction was made between slower and faster readers, using the median reading time as cutoff point. Eighteen students were categorized as slower readers (Mreading time = 12.83, SD = 1.95), while 32 students were categorized as faster readers (Mreading time = 8.75, SD = 1.19).

Immediately after presentation of the six news items, participants in the television condition were presented with the memory test. Participants in the print condition had been asked to raise their hand when they were finished reading, at which time they were presented with the memory test. We did not use a distracter task because previous television-print news comparison studies also administered recall measures immediately after news exposure. A cued recall test was used to measure participants' memory for the news stories. The test was preceded by written instructions.
Recall Measure

Memory for the news stories was measured by means of a paper-and-pencil test containing 50 open-ended cued recall questions. The memory test contained only questions about information that was present in the television narratives—and thus in the printed texts—of the news stories. The test did not, therefore, include questions on visual information that was not conveyed verbally. For each of the six news topics, questions were generated about the five most common components of news stories: event, place, principal(s), cause, and consequence. In addition, questions about story details were generated. Two versions of the memory test were used: one for each sequence in which the news items were presented. Cronbach’s alpha (K-R 20) for the memory test was .84. Examples of questions are: “One of the news stories was about marines smuggling drugs. What type of army equipment did they use to smuggle their drugs?” “How long has this clandestine operation been going on?”

To enable an analysis of the contribution of different types of text-picture relations to story recall, two independent judges classified the questions in the memory test into four types: (a) questions about verbal information that was supplemented with semantically redundant visual information on television (“direct overlap”), (b) questions about verbal information that was supplemented with only partially related visuals on television (“indirect overlap”), (c) questions about verbal information that was supplemented with conflicting visuals on television (“divergent overlap”), and (d) questions about verbal information that was supplemented with talking head pictures on television (“talking head”). Inter-rater agreement, as measured by Cohen’s kappa, was .96 for all news stories. Table 2 provides a summary of the numbers of questions in the memory test pertaining to direct overlap, indirect overlap, and talking head information for the news stories that were labeled as containing good and poor semantic overlap, respectively. None of the questions in the memory test was categorized as pertaining to “divergent overlap information.”

Results

Initial data checks showed that the distribution of recall scores satisfied the assumptions underlying analysis of variance. There were no statistically significant differences in recall scores between males and females, between participants who were tested in the morning versus the afternoon sessions, or between participants who had received different presentation orders. Therefore, in subsequent analyses these factors were disregarded.

Main Analysis

A 2 (television vs. print) x 2 (poor vs. good semantic overlap) analysis of variance was performed on the total scores on the memory test, with medium condition as a
Table 2
Numbers of Questions in the Memory Test Pertaining to Verbal Information that was Supplemented with Direct, Indirect, or Talking Head Pictures in Different Television Presentations

<table>
<thead>
<tr>
<th>Type of Text-Picture Correspondence</th>
<th>Direct</th>
<th>Indirect</th>
<th>Talking Head</th>
</tr>
</thead>
<tbody>
<tr>
<td>Television Format</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good Semantic Overlap</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. &quot;Marines Smuggle Drugs&quot;</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>b. &quot;Berlin Reichstag&quot;</td>
<td>6</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>c. &quot;Scotland and Wales&quot;</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Poor Semantic Overlap</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. &quot;Political Problems in France&quot;</td>
<td>—</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>e. &quot;Scandal in Nursery Home&quot;</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>f. &quot;Chinese Army&quot;</td>
<td>—</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>8</td>
<td>13</td>
</tr>
</tbody>
</table>

between-subjects factor and degree of semantic overlap as a within-subjects factor. No statistically significant main effect was found for medium condition. Overall, readers \( M = .53, SD = .15 \) remembered just about as much as viewers \( M = .54, SD = .15 \). A statistically significant main effect was found for degree of semantic overlap: Overall, the news items that were marked by good semantic overlap \( M = .56, SD = .18 \) were remembered better than the news items that were marked by poor semantic overlap \( M = .50, SD = .18 \), \( F(1, 98) = 15.92, MSE = 0.01, p < .001, \eta^2 = .14 \). However, no simple overall interpretation can be given to these results because there was a statistically significant interaction between medium and degree of semantic overlap, \( F(1, 98) = 69.89, MSE = 0.01, p < .001, \eta^2 = .42 \), that explained most of the variance. In accordance with the semantic overlap hypothesis, viewers of television stories containing poor semantic overlap remembered less than readers, whereas viewers of television stories containing good semantic overlap remembered considerably more than readers (see Table 3).

To determine whether the results presented in Table 3 indeed fitted the complete pattern of outcomes predicted by the semantic overlap hypothesis, two \( t \) tests were performed, using an alpha level of .025 based on a Bonferroni correction (Pedhazur & Pedhazur Schmelkin, 1991). These follow-up tests showed that the outcomes of the two television-print comparisons for poor and good semantic overlap were consistent with the semantic overlap hypothesis. As predicted by H1, there was a statistically significant recall disadvantage of television \( M = .44, SD = .19 \) compared with print \( M = .57, SD = .15 \) if the television presentation was marked
Table 3
Mean Recall Scores for Television and Print as a Function of Type of Television Format

<table>
<thead>
<tr>
<th>Television Format</th>
<th>Medium</th>
<th>Television</th>
<th>Print</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good Semantic Overlap</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. “Marines Smuggle Drugs”</td>
<td></td>
<td>.56 (.26)</td>
<td>.42 (.23)</td>
</tr>
<tr>
<td>b. “Berlin Reichstag”</td>
<td></td>
<td>.68 (.17)</td>
<td>.57 (.25)</td>
</tr>
<tr>
<td>c. “Scotland and Wales”</td>
<td></td>
<td>.63 (.23)</td>
<td>.50 (.20)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>.63 (.16)</td>
<td>.50 (.17)</td>
</tr>
<tr>
<td>Poor Semantic Overlap</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. “Political Problems in France”</td>
<td>.36 (.24)</td>
<td>.50 (.23)</td>
<td></td>
</tr>
<tr>
<td>e. “Scandal in Nursery Home”</td>
<td></td>
<td>.50 (.26)</td>
<td>.61 (.18)</td>
</tr>
<tr>
<td>f. “Chinese Army”</td>
<td></td>
<td>.49 (.27)</td>
<td>.62 (.20)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>.44 (.19)</td>
<td>.57 (.15)</td>
</tr>
</tbody>
</table>

Note: Scores represent mean proportions of correct answers (with standard deviations in parentheses). All recall differences between television and print were statistically significant at least at $p = .01$ level, except for the recall difference in news story “e” which was significant at $p = .02$.

by poor semantic overlap between text and pictures, $t(98) = 3.76, p < .001$. On the other hand, as predicted by H3, there was a statistically significant recall advantage of television ($M = .63, SD = .16$) over print ($M = .50, SD = .17$) if the television presentation was marked by good semantic overlap between text and pictures, $t(98) = 3.77, p < .001$.

In addition, to check whether the results of the two television-print comparisons held true for all of the $2 \times 3$ news stories, separate $t$ tests were performed for each news story. These $t$ tests showed that all six comparisons were in the predicted direction. As presented in Table 3, for all television news items that were labeled as containing poor semantic overlap, a statistically significant recall disadvantage was found compared with print (all $p \leq .02$). At the same time, for all television news items that were labeled as containing good semantic overlap, a statistically significant recall advantage was found compared with print (all $p \leq .01$). Table 3 also shows that there were overall differences in recall across the six news items. Separate contrast analyses showed that two news items—items (a) and (d)—scored lower than the other news items. Lower overall recall of these two stories, however, did not affect the relative effect of the televised and printed versions of these items. As predicted, item (a) led to a recall advantage, while item (d) led to a recall disadvantage of television compared to print. Although the above described results were clearly in favor of the semantic overlap hypothesis, additional analyses were performed to check whether our results were true both for readers who took less time...
processing the news stories and for readers who took more time digesting the news. A univariate analysis of variance on readers' recall with slow and fast reading as a between-subjects factor revealed that, overall, slow readers ($M = .56, SD = .18$) recalled somewhat more of the news information than fast readers ($M = .51, SD = .13$), but that this difference was not statistically significant, $F(1, 48) = 1.77, MSE = 0.039, p > .18$. Nevertheless, because slow readers did recall somewhat more of the news than fast readers and because we had used reading time as an indicator of invested cognitive effort, separate $t$ tests were conducted to investigate the interaction between medium and semantic overlap both for slow and fast readers. The results were consistent with those obtained for the complete group of readers: As predicted by H1, there was a statistically significant recall disadvantage of television ($M = .44, SD = .19$) compared with both the slow readers ($M = .60, SD = .16$) and the fast readers ($M = .55, SD = .13$) if the television presentation was marked by poor semantic overlap between text and pictures, $t(66) = 3.26, p < .003$, and $t(80) = 2.76, p < .008$, for the two comparisons respectively. While at the same time, as predicted by H3, there was a statistically significant recall advantage of television ($M = .63, SD = .16$) compared with the slow readers ($M = .53, SD = .21$) and the fast readers ($M = .49, SD = .15$) if the television presentation was marked by good semantic overlap between text and pictures, $t(66) = 2.01, p < .05$, and $t(80) = 4.18, p < .001$, respectively.

**Contribution of Different Types of Text-Picture Correspondence**

To further establish whether the recall advantage of television over print, which was found in the television stories that were marked by good semantic overlap, was indeed mainly due to the contribution of redundant audiovisual information, an analysis of variance on the recall scores of news items a, b, and c was conducted with medium as a between-subjects factor and recall of direct overlap, indirect overlap, and talking head information as a within-subjects factor. As shown in Figure 1, a statistically significant interaction was found between medium and type of text-picture correspondence, $F(2, 196) = 5.04, MSE = 0.02, p < .008, \eta^2 = .05$. Follow-up tests showed that students' superior recall of the television stories was especially due to their superior recall of redundant audiovisual information. On questions about verbal information that was supplemented with directly related television pictures, viewers ($M = .74, SD = .16$) scored significantly higher than readers ($M = .56, SD = .18$), $t(98) = 5.32, p < .001$. However, when verbal information was supplemented with indirectly related pictures or with talking head pictures, viewers ($M_{\text{indirect overlap}} = .36, SD = .25$), ($M_{\text{talking head}} = .60, SD = .21$), performed slightly but not statistically better than readers ($M_{\text{indirect overlap}} = .28, SD = .24$), ($M_{\text{talking head}} = .55, SD = .22$), $t(98) = 1.60, p > .11$, and $t(98) = 1.19, p > .23$, respectively.

Finally, to further investigate the recall advantage of print, which was observed for the news items marked by poor semantic overlap, an additional analysis of variance
on the recall scores of news items d, e, and f was conducted with medium as a between-subjects factor and recall of indirect overlap and talking head information as a within-subjects factor. Because only one question pertained to verbal information that was supplemented with direct television pictures (see Table 2), the analysis was restricted to recall of indirect overlap and talking head information. For questions about information that was supplemented with indirectly related television pictures, mean recall scores (with standard deviations in parentheses) were .36 (.25) and .47 (.20) for the television and print conditions, respectively. For questions that pertained to talking head information, mean recall scores were .51 (.20) and .66 (.16) for television and print, respectively. In both medium conditions, talking head information was recalled better than indirectly related information, \( F(1, 98) = 67.22, MSE = 0.02, p < .001, \eta^2 = .41 \). However, no interaction effect was found between
medium and text-picture correspondence, $F(1, 98) = .84, MSE = 0.02, p > .36$, indicating that the recall advantage of print over television was statistically significant for both categories of text-picture correspondence.

**Discussion**

The results of the present study support and extend the findings from an earlier study by Walma van der Molen and van der Voort (2000a), which indicated that the relative efficacy of televised and printed news is determined to a large extent by the amount of semantic overlap between verbal and visual television content. The results of the current study were in line with the hypotheses we had formulated on the basis of the semantic overlap hypothesis. Television stories that contained poor semantic overlap were remembered less well than printed stories (H1), while at the same time, memory for television stories containing good semantic overlap was considerably better than recall of the same printed stories (H3). In addition, it was found that the superior recall of the “good” television stories was indeed predominantly due to the fairly large amount of directly related audiovisual information in these items.

By combining the designs of previous inter- and intra-media comparison studies of memory for news information, we believe that our study successfully challenged the generally accepted notion that print is inherently a more effective transmitter of news information than television. Despite the fact that we ensured that the readers in our experiment had ample opportunities to control their processing of information (they were all proficient readers, they all expected to be tested, and were given as much time as they needed to read the stories), information presented in print induced considerably lower recall than the same information presented in a “good” television news format. Furthermore, although some readers took more time processing the news stories and recalled somewhat more information than other readers, the recall advantage of the “good” television presentations proved to be about the same for slower and faster readers.

The news stories that were used in the present study not only differed in their use of semantically redundant versus nonredundant pictures, but also in features such as topic category, story length, placement in the presentation sequences, and difficulty level. Although several of these features were varied deliberately and balanced across the “good” and “poor” conditions, these features could have affected our results. As presented in Table 3, two news items were recalled significantly less well than the other items. Overall lower recall of these items could be due to features such as those described above. Although we think that these explanations are unlikely because the two stories that were less well remembered differed in length, in their placement in the presentation sequences, and did not seem to be particularly more difficult than the other items, we cannot rule out the possibility that a combination of factors accounted for their lower recall. Nevertheless, the most important observation is that for all news items that were used in our experiment, the results of the
television-print comparisons were in the predicted direction. Thus, despite possible
(and in part consciously varied) differences between our news stories, “good”
television items led to greater recall whereas “poor” television items led to poorer
recall compared to their printed versions.

The present study not only provided a test of the semantic overlap hypothesis, but
also a comparison of memory for three different text-picture relations (direct,
indirect, talking head) to memory for the same information presented in print.
Although intra-media comparison studies have generally shown that directly related
audiovisual information may enhance recall, while indirect text-picture correspon-
dences may divert viewers’ attention away from the verbal message (e.g., Brosius et
al., 1996), such studies have not painted a clear picture of the position of talking
heads. From a limited capacity perspective, it should be expected that memory for
talking heads and directly related audiovisual information is the same because the
talking head picture does relate to the verbal message (you see and hear a person
talking) and would therefore require little capacity to process (Lang, 1995). On the
other hand, dual-processing theories predict that the talking head picture does not
contribute to learning because it is not semantically related to the verbal message
and therefore fails to provide the viewer with an additional memory code. Based on
the latter view, one should expect semantically related audiovisual information to be
recalled better than talking head messages.

The results of the present study showed that, in line with dual-processing theories,
semantically redundant information induced greater memory gains than talking head
information and that talking head information did not contribute to the superior
recall of the “good” television items compared with the printed versions. In addition,
it was found that talking head information and indirectly related audiovisual infor-
mation equally accounted for the inferior recall of the “poor” television news items.
The latter result suggests that both talking head information and indirectly related
audiovisual information may have equally diverted viewers’ attention away from the
verbal message.

Limitations of the Study

Four limitations that may affect the ecological validity of our study should be noted
here. A first limitation is that we used literal transcripts of the television narratives as
print materials. As was the case in other television-print news comparison experi-
ments, for reasons of internal validity, the printed texts in our study were identical to
the television texts. Although the texts were perfectly comprehensible without
pictures, we may have underutilized the potential advantages of printed texts
because the stories were not originally written as newspaper stories. In previous
television-print news comparisons, the potential underutilization of the print condi-
tions did not challenge the obtained results because the printed texts were recalled
better than the television versions. In our case, however, further research should
establish whether the relative recall of television versus print news is still best
described by the semantic overlap hypothesis if memory for different television news stories is compared not only with memory for literal transcripts but also with memory for stories originally written in print, that is, stories that include the journalistic strengths that are typical of "real" newspaper stories.

Second, our sample consisted of a very homogeneous group of university students. Although this choice was again motivated by our intention to replicate previous television-print news comparisons as much as possible, we are aware of the fact that our sample choice does not allow us to simply generalize our findings to all groups of news consumers. On the one hand, compared to older adults, university students may be less interested in the news. Although for the present study we did not measure students' news consumption behavior, our students often reveal that they do not follow the news on a daily basis and that they are not regular newspaper readers. It could be that such factors influence students' news reading proficiency and cause them to profit more from television news that is supplemented with semantically related pictures than people with a greater interest in news who are more used to news information presented in print. On the other hand, however, it is reasonable to assume that well-educated university students are more technically skilled readers than adults with lower levels of education. If less skilled readers would profit more from news presented via television than skilled readers, inclusion of a more heterogeneous sample of readers in the experiment could lead to an even stronger recall advantage of the "good" television news stories. To resolve this issue, further studies of the semantic overlap hypothesis should include participants of different age groups, from different educational and socio-economic backgrounds, with different daily news consumption behaviors.

Another limitation of our study is that our participants were given the opportunity to watch and read the news under relatively quiet experimental conditions. In the home situation, both viewers and readers usually consume the news under less optimal conditions. However, whereas readers are able to shut out the outside world by withdrawing themselves behind the newspaper, viewers of television newscasts are more likely to be distracted by other ongoing activities. An observational study of British families watching television in their homes (Gunter, Furnham, & Lineton, 1995) has shown that viewers' attention might well be divided when watching television newscasts. It is possible, therefore, that even if news items would contain more redundant audiovisual information, the recall advantage of television news would be less favorable under normal home viewing conditions.

Fourth, in the present experiment, we refrained from using news items that contained violent or otherwise emotionally arousing pictures. Yet, many news stories deal with crime and violence, and television news items about violence have become increasingly graphic in the past decades (Johnson, 1996; Slattery & Hakanen, 1994). It is generally accepted that the presence of violent images increases arousal in viewers, which in turn may affect viewers' memory for news stories (e.g., Furnham & Gunter, 1985; Gunter, Furnham, & Gietsoson, 1984; Lang, Newhagen, & Reeves, 1996). Nonetheless, in the present study, our main goal was to test the
semantic overlap hypothesis. Because thus far it is not quite clear how memory for different text-picture combinations is affected by the presence of violent or otherwise arousing images, the decision was made to use low-arousing stimuli. Possible interactions between violent television pictures and different text-picture combinations are being investigated in an experiment in progress.

Finally, an additional limitation of our study concerns the differences in length of exposure time between the television and print conditions. In the television condition, length of exposure was determined by the duration of the television broadcast, whereas participants in the print condition were allowed to read the news at their own pace, which resulted in a considerable variation in reading time. To overcome this limitation, future television-print news comparisons should also include an audio-only version of the television narratives (see Walma van der Molen & van der Voort, 2000b). That way, exposure times may be held constant at least for the television and audio-only conditions, while the advantage of redundant audiovisuals may be further compared to different verbal-only presentations. In addition, future research could consider using other measures of mental effort. We are aware of the fact that taking reading time as an indicator of invested cognitive effort is not an optimal measure to establish participants’ information processing. Thus far, however, within memory research, an objective index of depth of processing is not yet readily available. Vincent, Craik, and Furedy (1996) suggest that psychophysiological measures, such as cardiovascular responses, together with self-reports of invested mental effort could be used to explore the relation between levels of processing and memory performance, but these measures were beyond the scope of the present study.

Conclusion

Despite the above-mentioned limitations, we believe that the findings from the present study argue against the common presumption drawn from previous television-print comparisons that television is necessarily a less effective information medium than print. Our study suggests that supplementing about 40-50% of the verbal information with semantically related visuals may be enough to bring about an overall recall advantage of televised information compared with the same information presented in a print modality. We would like to stipulate, however, that in our view these results do not imply that television news could replace newspapers simply by showing more semantically related pictures. Newspapers may present more background information than regular television newscasts and they may confront consumers with more abstract issues and ideas, which are important tools for deeper communication of news information. Nevertheless, given the fact that many people rely on television for their daily news consumption, the results of the present study do suggest that for those people that predominantly use television, news learning could be improved. The ability to supplement verbal information with dynamic pictures is a unique attribute of television that other available news media
(newspapers and radio) do not possess. If used intelligently, this attribute could make the transfer of television news information as effective as or even more effective than the same information presented in print.

Note

Examples of questions about directly related audiovisual information are, “One of the news stories was about marines smuggling drugs. Where did the drugs come from? And to which country were the drugs transported?” The answers, “Aruba” and “Great Britain,” had been presented on television via a voiceover accompanied by pictures of a map that showed both parts of the world and that highlighted both locations with an arrow pointing from Aruba to Great Britain. An example of a question about the same news story that pertained to indirectly related audiovisual information is, “How long has this clandestine operation been going on?” The answer, “For three years,” had been presented via a voiceover accompanied by standard pictures of the island of Aruba.

References


