Children’s Creative Imagination in Response to Radio and Television Stories

by Patti M. Valkenburg and Johannes W. J. Beentjes

The prevailing explanation for the experimental finding that radio stories elicit more novel responses than do television stories is that viewers have difficulty dissociating themselves from ready-made television images (visualization hypothesis). In this experiment, we investigated a rival hypothesis that radio stories elicit more novel responses than do television stories because they are less well remembered (faulty-memory hypothesis). We presented 64 children at two age levels (grades 1 to 2 and 3 to 4) with one radio story and one television story, and exposed half the children in both age groups to the radio story twice to stimulate their memory. Contrary to the faulty-memory hypothesis, double presentation of a radio story did not result in fewer novel ideas than did a single presentation. In the older age group, radio stories elicited more novel responses than did television stories. We found no medium difference in the younger age group.

Television viewing, it is often claimed, evokes less creative imagination than does radio listening (Greenfield & Beagles-Roos, 1988; Mander, 1978; Meline, 1976; Webb, 1980; Winn, 1985). According to the visualization hypothesis, television, unlike radio, presents viewers with ready-made visual images and leaves them little opportunity to form their own images. When engaged in

Patti M. Valkenburg (PhD, Leiden University, 1995) is an associate professor and Dutch Royal Academy fellow at the University of Amsterdam, the Netherlands. Her research interests include children and media, methodology, and political communication. Johannes W. J. Beentjes (PhD, Leiden University, 1990) is an associate professor of Communication at the University of Nijmegen, the Netherlands. His research interests include media and children’s culture. An earlier version of this article was presented at the 1996 annual meeting of the International Communication Association, Chicago, IL. The data were collected when the first author was working at Leiden University, the Netherlands. We thank Tom H. A. van der Voort and Leo J. Th, van der Kamp for their useful suggestions, Patricia Greenfield for the stimulus materials and scoring manuals used in the study by Greenfield, Farrar-Hartmann, and Beagles-Roos (1986), and three anonymous reviewers for their valuable comments.

Copyright © 1997 Journal of Communication 47(2), Spring. 0021-9916/97/$5.00
creative thinking, viewers find it hard to dissociate themselves from the images supplied by television, so that they have difficulty generating novel ideas. This hypothesis attributes the reductive effect of television on creative imagination to the medium's visual nature (Valkenburg & van der Voort, 1994).

Six media comparison experiments, conducted with children between the ages of 7 and 13, have investigated the effect of radio and television on creative imagination. In one study, researchers presented children with problems to be solved (Meline, 1976). In the remaining five studies they presented the children with stories (Greenfield & Beagles-Roos, 1988; Greenfield, Farrar, & Beagles-Roos, 1986; Kerns, 1981; Runco & Pezdek, 1984; Vibbert & Meringoff, 1981). The problems and stories were presented in either radio or television format. Because the sound tracks were the same for the radio and television versions, the television versions differed only in that they included images. In all studies, researchers defined *creative imagination* as the number of (novel) ideas produced in response to a creative assignment. Five different types of creative assignments were used:

1. Problem solutions (Meline, 1976). Researchers presented children, via radio or television, with four problems (e.g., getting people to stop smoking), then asked them to invent one creative solution for each problem. Children's problem solutions were scored on stimulus freedom, that is, the number of novel elements not present in the given stimulus information.

2. Drawings (Vibbert & Meringoff, 1981). After being exposed to a radio or television story, children were asked to create drawings depicting various selected story scenes. The researchers scored the drawings on stimulus freedom.

3. Questions about a story (Kerns, 1981). Following exposure to either a radio or a television story, children answered questions designed to elicit creative responses to the story. The responses were scored on stimulus freedom, fluency (the number of ideas), flexibility (the number of distinct and different conceptual categories), and originality (the number of ideas statistically infrequent relative to the ideas of the other children in the sample).

4. A just suppose test (Runco & Pezdek, 1984). Following a radio or television story, researchers gave children an adaptation of Torrance’s (1974) Just Suppose Test, which required them to generate ideas about what might have happened if the story had turned out differently. Responses were scored on fluency, flexibility, and originality.

5. Story completions (Greenfield & Beagles-Roos, 1988; Greenfield et al., 1986). Researchers presented, via radio or television, a story that was stopped just before the end. Then, the children were asked to think up an ending for the incomplete story. The story completions were scored on stimulus freedom.

With four of the five creative assignments, radio led to more novel responses than did television. Only the adapted Just Suppose Test used by Runco and Pezdek (1984) failed to yield a significant medium effect on creative imagination. However, this absence of an effect might be attributable to the low sensitivity of the test, which, compared to the other operationalizations of creative imagination, elicited remarkably few creative responses in both the radio and television conditions (Valkenburg & van der Voort, 1994). In the
above-mentioned experiments, the medium differences in favor of radio were explained with the visualization hypothesis. According to the authors, the television presentations led to fewer novel ideas than did the radio presentations because children in the television condition had difficulty dissociating themselves from television images during creative thinking.

However, the results of the previous media comparison experiments can also be explained in a different way. According to a rival hypothesis, discussed by Greenfield and Beagles-Roos (1988), and Valkenburg and van der Voort (1994), radio presentations elicited more novel responses than television presentations, not because radio presentations were more stimulating for creative imagination, but because they were remembered less well. According to this faulty-memory hypothesis, radio listeners were less able to incorporate repetitive elements in their creative responses than were television viewers, because the radio presentation was less clear in their minds. The novel ideas produced by children in the radio condition might not have been creative responses but irrelevant fabrications invented to compensate for faulty memory.

The visualization and the faulty-memory hypotheses both provide an explanation for the finding in previous media comparison experiments that children produce more novel ideas following a radio presentation than following a television presentation. However, the two hypotheses fundamentally differ in the interpretation of the cognitive processes that underlie the production of novel ideas. The visualization hypothesis states that the novel ideas after a radio story are expressions of children's creativity. According to this hypothesis, radio leaves children more cognitive room to generate creative ideas than does television, because the production of novel ideas in the radio condition is not hindered by ready-made images. The faulty-memory hypothesis, however, disputing that the superior production of novel ideas after a radio presentation should be attributed to creativity. This hypothesis claims that the higher number of novel ideas after a radio presentation results from inferior story recall. The novel ideas produced after radio listening are not creative ideas, but merely inventions to fill in holes in a faulty memory.

A part of the faulty-memory hypothesis is that radio information is remembered less well than television information. This is supported by experimental evidence (Beagles-Roos & Gat, 1983; Gibbons, Anderson, Smith, Field, & Fisher, 1986; Greenfield & Beagles-Roos, 1988; Hayes, Kelly, & Mandel, 1986). However, to date, none of the media comparison experiments has investigated whether the relatively poor recall of radio information is responsible for the incorporation of fewer repetitive and more novel ideas in children's creative products. Therefore, the faulty-memory hypothesis cannot be excluded as a rival explanation for the differences in creative imagination elicited by radio and television found in earlier media experiments.

The purpose of the present experiment was to test whether the faulty-memory hypothesis should be maintained or excluded as a rival explanation for the relatively high number of novel ideas in creative products following a radio presentation. In addition, we investigated whether the differences found in creative imagination in favor of radio could be replicated. To enhance compara-
bility, we used the same stimulus materials employed in three of the six previous media comparison experiments. In these studies, children's creative imagination was operationalized by means of story completions (Greenfield & Beagles-Roos, 1988; Greenfield et al., 1986) or a just suppose test (Runco & Pezdek, 1984). We decided to use story completions, because Runco and Pezdek's (1984) test might have lacked sensitivity (Valkenburg & van der Voort, 1994). Following Greenfield and Beagles-Roos and Greenfield et al., we presented children from two different age groups (grades 1 to 2 and 3 to 4) with one radio story and one television story. Both stories were stopped just before the end. After each story, the children were asked to retell the story as well as they could. We assessed children's story recall on the basis of their retellings. Subsequently, we asked the children to think of an ending for the incomplete story. To determine children's creative imagination from their story endings, we followed the content analytic procedure developed by Greenfield and colleagues. We investigated how many ideas in the story completions were reproduced from the stimulus story (i.e., the number of repetitive ideas) and how many ideas in the story completions were not present in the given stimulus story (i.e., the number of novel ideas).

We investigated the faulty-memory hypothesis in two ways. First, we introduced an extra radio condition in addition to the usual radio and television conditions, in which children were exposed twice to the same radio story. Because there is ample evidence that repetitive stimulus presentation improves recall (e.g., Crowder, 1976; Richardson-Klavehn & Bjork, 1988), we expected that double presentation of a radio story would stimulate children's recall. If this expectation is verified, the validity of the faulty-memory hypothesis can be investigated by comparing children's creative story completions following a double radio presentation with those following a single radio presentation. After all, if the faulty-memory hypothesis is valid, improved recall of a radio story should result in fewer novel and more repetitive responses in the story completions.

A second way in which we investigated the faulty-memory hypothesis was by using an alternative scoring method for assessing children's creative imagination that takes the quality of the novel ideas into account. Earlier media comparison experiments focused only on the number of novel ideas produced and failed to pay attention to the quality of these novel ideas. Therefore, it was possible that the novel ideas produced by children who listened to the radio story were irrelevant fabrications invented to compensate for faulty memory.

To assess both the novelty and the quality of children's story completions, we used Amabile's (1982) consensus method for assessing creative products. This method assumes that the creative value of a product (i.e., story completions) cannot be assessed through objective analysis. Instead, the most suitable method is to rely on subjective judgments of creative imagination made by observers familiar with the domain in question. If these observers agree on the creative value of the individual products, then their judgments can be accepted as a theoretically valid measure of creative imagination. Similarly, in everyday life, a product is labeled as creative because appropriate judges collectively
agree on this evaluation (Amabile, 1982; Sternberg, 1988; Sternberg & Lubart, 1991). The consensus method has been applied to creative products of both adults (e.g., poems) and children (e.g., collages, stories made up to accompany a picture book). In all these studies, there was consensus on the creative value of the individual creative products, evidenced by high interjudge reliabilities (Amabile, 1982; Amabile & Gittomer, 1984; Amabile, Hennessy, & Grossman, 1986; Hennessy & Amabile, 1988; Lubart & Sternberg, 1991)

Hypotheses
The faulty-memory hypothesis postulates that inferior story recall results in story completions with fewer repetitive and more novel ideas. To stimulate recall, we presented children with a radio story twice instead of once. If these children remember the radio story better than do children presented with the radio story once, the faulty-memory hypothesis predicts that:

H1a: The story completions of children presented with the radio story twice will contain fewer novel ideas than those of the children who listen to the radio story once.

H1b: The story completions of children presented with the radio story twice will contain more repetitive ideas than those of the children who listen to the radio story once.

The faulty-memory hypothesis can be investigated experimentally, and correlationally. The hypothesis predicts that the level of story recall is related negatively to the number of novel ideas and positively to the number of repetitive ideas in the story completions. The correlational prediction of the faulty-memory hypothesis results in the following hypothesis.

H2: As children remember the stimulus story less well, they produce more novel and fewer repetitive ideas in their story completions.

The faulty-memory hypothesis also predicts that faulty memory results in a higher number of irrelevant fabrications in the story completions. We investigated this prediction by comparing the quality of the story completions of children who heard the radio story twice with that of the children who heard the story once. We assessed the quality of the story completions through Amabile’s (1982) consensus method. After having checked whether the double radio presentation caused superior story recall, we investigated the following hypothesis.

H3: The story completions of children presented with the radio story twice are of a better quality than those of the children who listen to the radio story once.

Because various media comparison experiments found that radio stories elicited more novel (Greenfield & Beagles-Roos, 1988; Greenfield et al., 1986;
Kerns, 1981; Meline, 1976; Vibbert & Meringoff, 1981) and fewer repetitive (Greenfield & Beagles-Roos; Greenfield et al.) responses than did television stories, we investigated the following hypothesis.

**H4a:** In response to a radio story, children generate more novel elements in their story completions than in response to a television story.

**H4b:** In response to a radio story, children generate fewer repetitive elements in their story completions than in response to a television story.

In the present study, the number of novel ideas in children's story completions was scored in two different ways: by means of the content analytic method of Greenfield et al. (1986) and by means of the consensus method of Amabile (1982). The use of two scoring methods offered the opportunity to investigate the extent to which the content analytic and the consensus methods produced the same results. The content analytic method might yield biased results in two ways. First, there is a risk that a story completion that contains a high number of irrelevant fabrications receives an undeservedly high creative-imagination score. Second, through the content analytic method, a short story completion gets a low creative-imagination score by definition, although the story completion might actually be very creative. Addition of the consensus method offers the opportunity to investigate the possible scoring artifacts of the content analytic method. Therefore, in this study we investigated the following research question.

**RQ:** To what extent do the results of the content analytic scoring method agree with those of the consensus method?

**Method**

**Sample**

We recruited the subjects, 32 first and second graders (6.5 to 8 years old) and 32 third and fourth graders (8.5 to 10 years old), from eight primary schools in an urban district in the Netherlands. The sampling method used insured that the sample of children was representative of Dutch society in terms of socioeconomic status. The children were all native speakers of Dutch without learning, visual, speech, or hearing disabilities. Boys and girls were equally represented in the two age groups.

**Story Materials**

Three of the earlier media comparison experiments (Greenfield & Beagles-Roos, 1988; Greenfield et al., 1986; Runco & Pezdek, 1984) used two animated stories produced by Weston Woods Studios. They were *Strega Nonna* (De Paola, 1975), a story about a kind-hearted old witch with a magic pasta pot, and *A Story, A Story*, an African folktale. To enhance comparability with earlier media comparison experiments, we asked Patricia Greenfield to send us these stories. How-
ever, after inspection of the two stories, we agreed with Greenfield et al.’s (1986) conclusion that A Story, A Story is too difficult for younger children to comprehend. Accordingly, we used only Strega Nonna in the present study. Instead of A Story, A Story, we selected Doctor De Soto (Steig, 1983), another animated story produced by Weston Woods Studios. Doctor De Soto has an easier story line, about a mouse dentist and his wife who get a hungry fox as a patient.

A female Dutch narrator translated and recorded the text of the U.S. sound track of Strega Nonna. To keep the voice and style of narration similar in the two stories, the same narrator also recorded Doctor De Soto. The sound tracks were identical for the radio and television versions of each story. As in the studies by Greenfield and colleagues, each story was stopped approximately 1 minute before the end. The incomplete versions of both Strega Nonna and Doctor De Soto each had running times of approximately 7 minutes.

**Design**
The experiment used a 2 (age group: grades 1 to 2 and 3 to 4) x 2 (radio presentation: single vs. double presentation of radio story) x 2 (medium: radio vs. television story) mixed design. Medium was varied within subjects; age group and radio presentation were between-subjects variables. We presented each child in each age group with two stories, one in radio format and another in television format. Half the subjects in each age group listened to the radio story just once; the other half heard it twice. Within each age group and radio condition, half the subjects were exposed to the television version of Strega Nonna and the radio version of Doctor De Soto. The other half listened to the radio version of Strega Nonna and watched the television version of Doctor De Soto. To control for possible order effects, half the children in each age group and radio condition received the television story first, and half received the radio story first.

**Procedure**
A female experimenter in an empty room in the child’s primary school individually tested each child. To prevent the children from discussing the stories with classmates, we selected only two children from each classroom. We chose the first child randomly, then asked the teacher to select a second child of the opposite sex with the same level of linguistic competence as the first. Finally, we assigned the two children randomly, one to the double and one to the single radio story. This matching method was applied to reduce the error variance of the between-subjects variable radio condition (single vs. double radio story), and hereby to improve the power.

During the experimental session, the child sat to the right of the experimenter at a table. A television monitor and a radio cassette player sat on the table in front of the subject. Each session began with an informal conversation of approximately 5 minutes to help the child feel at ease, and to make certain that he or she understood that the experiment had nothing to do with formal grades or testing. The first story was introduced as follows: “We are going to
listen to/watch a story. The story is called Strega Nonna /Doctor De Soto. After we have heard/seen this story, we are going to talk about it.” Then the incomplete story was presented.

At the conclusion of the story presentation, the experimenter asked if the child had ever heard or seen the story before. Children who were familiar with either story were replaced by a classmate with the same level of linguistic competence. Subsequently, we gave the following recall instruction: “Now I would like you to tell me what the story was about. It was a long story and nobody can remember everything about a story, but I would like you to tell me everything that happened in the story from the beginning to the end.”

After the child finished his or her retelling of the story, we gave the creative imagination instruction. With respect to Strega Nonna, the instruction was as follows: “At the end of the story Strega Nonna blew the three kisses, and with a sputter the pasta pot stopped boiling. ‘Oh grazie, thank you,’ the people cried. And then the story ended. But in reality the story still continues. And now I would like you to tell me a story about what you think is going to happen next.” With respect to Doctor De Soto, we gave the same instruction, but with this introduction: ‘At the end of the story the dentist said to the fox: ‘You will never have to visit us again.’ And the fox thought: ‘he he, no one will ever visit you again.’ And then the story ended. . . .”

When the child had finished the first creative imagination assignment, the experimenter introduced the second story. The procedure for the second story was identical to that for the first story. However, a child who had listened to a radio story first received the second story in television format. A child who first had seen a television story received the second story in radio format. Children’s retellings and story completions were audiotaped for later transcription.

**Measures**

*Free recall.* To score the children’s retellings, we conducted a propositional analysis as developed by Kintsch (1974). In a first step, we converted the transcribed sound track of each stimulus story into an ordered list of verbal propositions. A verbal proposition consists of a predicate (e.g., verb, adjectives, adverbs) and one or more arguments (e.g., subject, object, or another proposition).¹ The verbal propositional analysis yielded 284 verbal propositions for Strega Nonna and 312 for Doctor De Soto.

To measure the memory of visual story information, in a second step, the visual part of each stimulus story was converted into an ordered list of visual propositions (i.e., stimulus propositions conveyed visually). However, because Kintsch (1974) focused only on verbal texts, we had to develop an additional procedure. First, we developed a visual observation instruction in which we described clear criteria concerning how to score visual propositions about characters, objects, time and location settings, and actions. Subsequently, the first author and another rater analyzed the visual parts of each stimulus story in

¹ For a detailed description of the conventions used in converting verbal texts into propositions, see Kintsch (1974) and Kintsch and Van Dijk (1978).
accordance with the criteria established in the visual observation instruction. The interrater reliability for this procedure, based on the ratio of the number of commonly scored propositions to the total number of scored propositions, was .87. The visual content analysis yielded 542 visual propositions for *Strega Nonna* (57 shots) and 694 for *Doctor De Soto* (67 shots).

In a final step, we scored children’s transcribed retellings. Each retelling was converted into a list of propositions in accordance with the conventions of Kintsch (1974). The interrater reliability for the number of propositions counted in each free-recall protocol, conducted on a sample of 16 retellings, was .95. We scored each recall proposition in a child’s retelling on whether it was semantically equivalent to a verbal or visual stimulus proposition. The interrater reliability of this scoring procedure, conducted on a sample of 40 free-recall protocols, was .94. A child’s total recall score was the sum of the verbal and visual stimulus propositions that were recalled. This scoring procedure implied that children’s recall after the television presentation was made up of verbal as well as visual propositions, whereas children’s recall after the radio presentation was based on only verbal propositions. One could argue that this scoring procedure unjustly privileged children in the television condition and that this bias should be controlled. However, the point of media comparison experiments is to investigate differences between media. Therefore, one should not control for differences between the two media. After all, dual information provision is inherent to television.

During data screening, we located one outlier with an extreme value on the recall score. Because outliers may exert a negative influence on statistical power (Stevens, 1992), we assigned the outlier a score that was one unit higher than the next highest score in the cell in which it was detected (Tabachnick & Fidell, 1989, p. 70).

**Creativity.** We coded the children’s story completions according to the protocol in the scoring manual of Farrar-Hartman (1984), which was sent to us by Patricia Greenfield. Greenfield et al.’s (1986) basic coding unit was the proposition. A proposition in the story completions was considered novel if it was not found in the stimulus story, and repetitive if it was semantically equivalent to a verbal or visual stimulus proposition. In our study, we replicated the propositional analysis of Greenfield and colleagues. The story completions were converted into a list of propositions by two independent raters. The interrater reliability for this procedure was .96. Then, each proposition was coded as either novel or repetitive. The interrater reliability for this procedure was .90. Data screening revealed four outliers, cases with an extremely high number of novel propositions in response to both the television and the radio story. We assigned each outlier a score that was one unit higher than the next highest score in the cell in which it was detected (Tabachnick & Fidell, 1989).

In addition to using the content analytic method of Greenfield and colleagues, we evaluated the story completions using Amabile’s (1982) consensus method. The story completions were rated on five dimensions, two concerning the novelty of the story completions and three concerning the quality of the story completions. The two novelty dimensions (creativity and novelty) were
adopted from Amabile (1982) and Hennessey and Amabile (1988). We had to develop the three quality dimensions ourselves, because story completions require different quality criteria than do the creative products that were used in the Amabile studies (e.g., poems, collages). We asked 16 developmental psychologists to read three story completions and to indicate which quality criteria a creative story completion has to meet to be deemed high in quality. According to these psychologists, a story completion was high in quality if it satisfied the following three criteria: (a) It maintained continuity with the preceding stimulus story, (b) the events within the story completion were coherent, and (c) the story completion provided an ending to the previous, incomplete stimulus story.

We recruited five developmental psychologists to judge the story completions. The judges were unaware of the research hypotheses and made their assessments independently. Before judging the story completions, we familiarized the judges with the stimulus stories and the dimensions of judgment. Then, we asked them to read the story completions once. They were then instructed to go through the story completions a second time and to judge each story completion on a discrete 7-point rating scale on the following dimensions:

• Creativity: Using your own, subjective definition of creativity, assess how creative the story completion is.
• Novelty: How novel is the story completion? Does the story completion have an original twist, or is it only a variation on the preceding story theme?
• Continuity: Does the story completion maintain continuity with the preceding stimulus story or is it irrelevant?
• Coherence: To what extent are the events within the story completion coherent? Do the events within the story completion fit together well, or is there no clear connection between them?
• Completion: Does the story completion round off well or does it stop abruptly?

The judgment procedure took 2 days. On the first day, judges evaluated the 64 completions of *Strega Nonna*, and on the second day, the 64 completions of *Doctor De Soto*, with story completions and judgment dimensions presented to the judges in random order. The judgments were made per dimension. Therefore, a judge evaluated all story completions on each dimension before going to the next.

Following Amabile (1982), we calculated a Spearman-Brown interjudge reliability for each judgment dimension (Nunnally, 1967). For *Strega Nonna*, the reliability coefficients of the dimensions were as follows: creativity .87, novelty .90, coherence .80, continuity .79, and completion .81. For *Doctor De Soto*, they were as follows: creativity .92, novelty .86, coherence .80, continuity .79, and completion .81.

To investigate the extent to which the scores on the five judgment dimensions reflected underlying factors, a principal component analysis with varimax rotation was carried out. A clear two-factor pattern emerged. The creativity and novelty dimensions loaded high on the first factor, which could by labeled novelty. The continuity, coherence, and completion dimensions loaded high on
the second factor, which could be labeled quality. Together, the two factors accounted for 86.5% of the variance in the five judgment dimensions (see Table 1).

Because of the high amount of variance accounted for by the factor solution, we created a novelty scale and a quality scale on the basis of the five judgment dimensions. We formed the novelty scale by totaling the child’s unweighted scores on the creativity and novelty dimensions. We constructed the quality scale by summing the child’s unweighted scores on the continuity, coherence, and completion dimensions. The reliabilities of the novelty and quality scales were satisfactory. Cronbach’s alpha values were .86 for the novelty scale and .81 for the quality scale. The correlation between the novelty and quality scale was, \( r = .19, p > .05 \).

### Results

**Correlation Between the Two Creative-Imagination Measures**

Our research question asked to what extent the results of the two scoring methods for creative imagination, the content analytic method (i.e., the number of novel propositions) and the consensus method (i.e., novelty judgments), would agree. To investigate this research question, we calculated a Pearson correlation between the number of novel propositions and the scores on the novelty scale found through principal component analysis. The number of novel propositions correlated strongly with the scores on the novelty scale (\( r = .76, p < .01 \)), showing that the results of the two scoring methods agreed to a large extent. Because of the high correlation between the content analytic and the consensus method, the two scoring methods cannot be considered as independent from each other. Therefore, in the analyses below, the results from the two scoring methods were simultaneously included in a multivariate analysis of variance (Tabachnick & Fidell, 1989).

<table>
<thead>
<tr>
<th>Judgment Dimension</th>
<th>Factor 1 Novelty</th>
<th>Factor 2 Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1: Novelty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creativity</td>
<td>.94</td>
<td>.19</td>
</tr>
<tr>
<td>Novelty</td>
<td>.97</td>
<td>-.01</td>
</tr>
<tr>
<td>Factor 2: Quality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuity</td>
<td>-.31</td>
<td>.87</td>
</tr>
<tr>
<td>Coherence</td>
<td>.23</td>
<td>.89</td>
</tr>
<tr>
<td>Completion</td>
<td>.28</td>
<td>.80</td>
</tr>
<tr>
<td>Eigenvalues</td>
<td>2.22</td>
<td>2.08</td>
</tr>
</tbody>
</table>
Manipulation Check

To investigate the faulty-memory hypothesis, we introduced an extra radio condition in which the radio story was presented twice to stimulate the child’s story recall. A major premise of this study was that a double radio presentation would be remembered better than a single one. To check if this occurred, we performed an analysis of variance on the number of correctly recalled propositions in children’s retellings following a radio story, with age group (grades 1 to 2 and 3 to 4) and radio condition (single presentation and double presentation) as between-subjects variables. Table 2 presents the mean number of propositions recalled from a single or double presentation of the radio story in the two age groups.

The ANOVA revealed a significant main effect for radio story, $F(1,60) = 10.43$, $p < .01$, $\eta^2 = .15$, and age group, $F(1,60) = 4.65$, $p < .05$, $\eta^2 = .07$. The interaction between radio story and age group was not significant, $F(1,60) = 3.59$, $p > .05$. The main effect of radio story implied that a double presentation of a radio story caused better story recall than a single presentation. The main effect of age group meant that the older children recalled more from both the single and double presentations than did the younger children.

To compare children’s story recall following a single or double presentation of the radio story with story recall in response to a television story, two separate ANOVAs were conducted. Both analyses included age group as a between-subjects variable and medium (radio vs. television) as a within-subjects variable. In the first ANOVA we investigated the differences in story recall following a radio story presented once and a television story. The difference in recall between a television story ($M = 43.47$, $SD = 22.10$) and a single radio story ($M = 40.53$, $SD = 20.69$) was not statistically significant, $F(1,30) = .83$, $p = .37$. In the second ANOVA we investigated the differences in story recall in response to a radio story presented twice and a television story. The analysis revealed that, in both age groups, children remembered a radio story presented twice ($M = 55.94$, $SD = 19.31$) better than a television story presented once ($M = 40.59$, $SD = 18.41$), $F(1,30) = 11.54$, $p < .01$, $\eta^2 = .28$.

Table 2. Mean Number of Recalled Stimulus Propositions Following a Single or Double Radio Story by Age Group

<table>
<thead>
<tr>
<th></th>
<th>Grades 1–2</th>
<th>Grades 3–4</th>
<th>Grades 1–2</th>
<th>Grades 3–4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Single Radio Story</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>30.88</td>
<td>50.19</td>
<td>55.31</td>
<td>56.56</td>
</tr>
<tr>
<td>$SD$</td>
<td>18.98</td>
<td>18.03</td>
<td>21.74</td>
<td>17.22</td>
</tr>
</tbody>
</table>

Number of Novel and Repetitive Ideas in the Story Completions

Two research hypotheses made predictions about the number of novel ideas in the story completions. H1a predicted that the story completions of children presented with a radio story twice would contain fewer novel ideas than those
Children’s Creative Imagination

of children who listened to a radio story once. H4a predicted that the story completions in response to a radio presentation would contain more novel ideas than those in response to a television presentation. To test H1a and H4a, we conducted a MANOVA on the number of novel propositions and the scores on the novelty scale with age group (grades 1 to 2 and 3 to 4) and radio condition (single and double radio story) as between-subjects variables and medium (radio vs. television) as a within-subjects variable. The MANOVA did not reveal a significant interaction between medium and radio condition, $F(1,60) = .03, p = .86$. This implied that, contrary to H1a, a radio story presented twice elicited as many novel ideas as a radio story presented once. The MANOVA did reveal a significant main effect of age group, $F(1,60) = 19.21, p < .001, \eta^2 = .24$, and a significant interaction between medium and age group, $F(1,60) = 4.23, p < .05, \eta^2 = .07$. Table 3 presents the mean numbers of novel propositions and the mean scores on the novelty scale for the two age groups and the two media.

The main effect of age group meant that the story completions of the older children contained more novel ideas than those of the younger children. To explain the interaction between medium and age group, we conducted a simple effects analysis (Tabachnick & Fidell, 1989). This analysis showed that for older children, in agreement with H4a, the story completions following a radio story contained more novel ideas than those following a television story, $F(1,60) = 4.67, p < .05, \eta^2 = .07$. For the younger children, however, there was no significant difference in the number of novel ideas in response to a radio or television story, $F(1,60) = .63, p = .43$.

Two hypotheses pertained to the number of repetitive ideas in the story completions. H1b predicted that the story completions of children who heard a radio story twice would contain more repetitive ideas than those of children who listened to a radio story once. H4b predicted that the story completions in response to a radio presentation would contain fewer repetitive ideas than those in response to a television presentation. H1b and H4b were tested by means of an ANOVA on the number of repetitive propositions in the story.

### Table 3. Mean Number of Novel Propositions and Novelty Judgments by Age Group and Medium

<table>
<thead>
<tr>
<th>Creative-Imagination Measures</th>
<th>Grades 1–2 (n = 32)</th>
<th>Grades 3–4 (n = 32)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Radio</td>
<td>TV</td>
</tr>
<tr>
<td>Number of Novel Propositions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>12.56</td>
<td>13.22</td>
</tr>
<tr>
<td>$SD$</td>
<td>9.26</td>
<td>7.56</td>
</tr>
<tr>
<td>Novelty Judgments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>34.59</td>
<td>36.72</td>
</tr>
<tr>
<td>$SD$</td>
<td>13.01</td>
<td>13.34</td>
</tr>
</tbody>
</table>
completions, with age group and radio condition as between-subjects variables and medium as a within-subjects variable. The ANOVA did not yield a significant interaction between medium and radio condition, $F(1,60) = .44, p = .51$. This result implied that, contrary to H1b, the story completions after a double radio story ($M = 3.45, SD = 3.10$) did not contain more repetitive ideas than those after a single radio story ($M = 5.39, SD = 4.77$). The ANOVA showed only a significant interaction between medium and age group, $F(1,60) = 5.89, p < .05, \eta^2 = .09$. A simple effects analysis indicated that H4b held only for the older age group. The story completions of the older children contained more repetitive propositions following a television story ($M = 5.46, SD = 4.31$) than following a radio story ($M = 3.81, SD = 3.25$), $F(1,60) = 4.73, p < .05, \eta^2 = .07$, regardless of whether the children heard the radio story once or twice. For the younger children, we found no significant difference in the number of repetitive propositions between the radio and television presentations, $F(1,60) = 1.72, p = .19$.

Correlations of Story Recall with the Number of Novel and Repetitive Ideas in the Story Completions

The faulty-memory hypothesis contained in H2 predicted that level of story recall would be related negatively to the number of novel ideas and positively to the number of repetitive ideas in the story completions. To investigate this prediction, we correlated story recall with the number of novel propositions, the scores on the novelty scale, and the number of repetitive propositions. Contrary to H2, level of story recall was related positively to the number of novel propositions in the story completions ($r = .47, p < .01$) as well as to the scores on the novelty scale ($r = .52, p < .01$). Also contrary to H2, level of story recall was not significantly related to the number of repetitive propositions ($r = .17, p > .05$).

Quality Differences in Story Completions Following a Single or Double Radio Presentation

The faulty-memory hypothesis contained in H3 predicted that the quality of the story completions following a double presentation of a radio story would be superior to that of story completions following a single radio presentation. To test H3, an ANOVA was conducted on the quality judgments of the story completions, again with age group and radio condition as between-subject variables and medium as a within-subject variable. Contrary to the faulty-memory hypothesis, the quality of the story completions following a double radio story ($M = 64.44, SD = 16.96$) was not superior to that of the story completions following a single radio story ($M = 61.56, SD = 15.86$), $F(1,60) = .84, p = .36$. The quality of a radio story completion after a single presentation also was not inferior to that of a television story completion ($M = 64.20, SD = 14.23$). The ANOVA revealed only a main effect of age group, $F(1,60) = 4.72, p < .05, \eta^2 = .08$. Older children generated story completions of higher quality than did the younger children, regardless of whether they were presented with a television story or a (single or double) radio story.
Discussion

The basic purpose of this study was to test the faulty-memory hypothesis, which postulates that radio presentations elicit more novel ideas than do television presentations, because radio information is less well remembered. In this hypothesis, novel ideas are seen as fabrications to compensate for a faulty memory. The faulty-memory hypothesis challenges the prevailing visualization hypothesis, which states that children generate relatively few novel ideas in response to a television presentation because they have difficulty dissociating themselves from television’s ready-made images. In the visualization hypothesis, novel ideas are considered expressions of children’s creativity.

In this study, we derived three predictions from the faulty-memory hypothesis. The first prediction, that double presentation of a radio story leads to fewer novel (H1a) and more repetitive ideas (H1b) in the story completions than does a single radio presentation, was not confirmed. Two radio presentations elicited as many novel and repetitive elements in the story completions as a single radio presentation. The second prediction, that superior story recall is associated with fewer novel and more repetitive elements in the story completions (H2), also was not supported. Superior story recall was related to significantly more novel ideas in the story completions and was not significantly related to the number of repetitive ideas in the story completions. The third prediction, that the story completions in response to a double presentation of a radio story are of a higher quality than those in response to a single radio presentation (H3), was not confirmed either. Our results showed that the quality of story completions in response to a single presentation of a radio story was not inferior to the quality of story endings after two presentations. Because none of the predictions derived from the faulty-memory hypothesis was confirmed, there is no reason to assume that the hypothesis provides a plausible rival explanation for the finding in previous experiments that radio presentations elicit more novel ideas than do television presentations.

Because the faulty-memory hypothesis proved to be implausible, the visualization hypothesis has lost its rival. The visualization hypothesis, which states that children have difficulty generating novel ideas because they cannot easily dissociate themselves from television images, has never been directly investigated. The visualization hypothesis did, however, receive indirect support. In studies by Greenfield and colleagues, and in our study, children incorporated more repetitive ideas in their television story completions than in their radio story completions. The relatively high number of repetitive ideas in television story completions cannot be attributed to a better story recall, because the number of repetitive elements following a television story was also higher than that following a double presentation of a radio story. Because the television story differed from the radio story only in that it included images, it is not illogical to ascribe the relatively high number of repetitive ideas following a television story to an inability to dissociate oneself from ready-made television images.
A second purpose of our study was to investigate the RQ about the extent to which the content analytic method (i.e., the number of novel propositions) agreed with the novelty judgments obtained through the consensus method. As discussed in the introduction, the content analytic method seems susceptible to scoring artifacts because it does not evaluate the quality of story endings. Despite this shortcoming, the content analytic method proved to correlate highly with the consensus method. In addition, both scoring methods yielded a similar pattern of results for (a) the older and younger children and (b) radio and television. The agreement between the results of the content analytic and consensus methods is an indication of the convergent validity of both methods and increases confidence in the results obtained by the content analytic method.

A final purpose of this study was to investigate whether the differences in creative imagination in favor of the radio that were found in earlier experiments could be replicated. On the basis of previous research, we predicted that a radio story would elicit more novel (H4a) and fewer repetitive ideas (H4b) in children’s story completions than a television story would. This hypothesis held only for the older age group. In the younger age group, we found no significant differences between the numbers of novel and repetitive ideas following a radio or television story.

The different findings for the younger and the older children might be explained in terms of children’s cognitive development. According to Piaget and Inhelder (1969), two kinds of mental images develop during childhood: reproductory images and anticipatory images. Reproductive images are images in which scenes that have been observed earlier are brought into mind again. Anticipatory images are images of movements and transformations, as well as their possible outcomes, that were not previously observed. According to Piaget and Inhelder, it is not until the level of concrete operations (7 to 8 years old) that children are capable of creating anticipatory images. The ability to form anticipatory images could be important for the creative-imagination assignment in the present study. After all, we asked children to complete an unfinished story, a task that calls for the ability to imagine scenes that have not been observed before. Because children in grades 1 and 2 may have difficulty forming anticipatory images, completing an unfinished story is a difficult assignment, regardless of whether the story is conveyed via radio or television. This may explain why the medium used did not affect the number of novel elements produced by the younger children. Older children, who are already capable of forming anticipatory images, apparently experience greater difficulty generating a novel story completion after a television story than after a radio story.

Our finding about the difference between younger and older children can be compared only with the media comparison experiments by Greenfield and colleagues (Greenfield & Beagles-Roos, 1988; Greenfield et al., 1986), because the children in the other studies (Kerns, 1981; Meline, 1976; Runco & Pezdek, 1984; Vibbert & Meringoff, 1981) were older than those in our youngest age group. The interaction between medium and age group found in the present
study does not agree with the results of the studies by Greenfield and colleagues. They found that radio stories elicited more novel ideas than did television stories in both age groups. This discrepancy between the findings of Greenfield and colleagues and the present results might also be explained in terms of the Piagetian theory. The younger children in both our study and the two studies by Greenfield and colleagues were between 6.5 and 8 years old and, therefore, on the threshold of the level of concrete operations (7 to 8 years). This stage is, as mentioned before, important for the creative-imagination assignment employed in the studies involved. Maybe the younger children in the studies by Greenfield and colleagues were just a little further in their cognitive development than the younger children in our study, so that the younger children in our study might have experienced greater difficulty completing the creative-imagination assignment.

As in earlier media comparison experiments, the results of the present study are consistent with the visualization hypothesis. However, because the question of whether children have difficulty dissociating themselves from television images has never been investigated, the evidence for the visualization hypothesis is not conclusive. In future media comparison studies into the effect of radio and television on children's creative imagination, the validity of the visualization hypothesis can be investigated by distinguishing between elements of the stimulus story that have been visualized (audiovisual elements) and elements that have been conveyed only via audio track (audio-only elements). The visualization hypothesis predicts that children's creative imagination is hindered only if creative assignments concern audiovisual elements of the stimulus story, but not if assignments pertain to audio-only elements.

References


