Memory & Cognition
2000, 28 (2), 214-223

Do film cuts facilitate the perceptual and cognitive organization of activity sequences?

STEPHAN SCHWAN, BÄRBEL GARSOFFKY, and FRIEDRICH W. HESSE
University of Tübingen, Tübingen, Germany

Film depictions of activities possess two kinds of structures—namely, the structural features of the depicted activities themselves and a formal structure defined by film cuts. The former structure is used by everyday observers for perceptually and cognitively unitizing the continuous flow of events into comprehensible entities. It seems conceivable that cuts can serve a similar unitizing purpose for film viewers. For each of two different activity sequences, two film versions were produced. Throughout each film version, cuts were placed either at breakpoints or at nonbreakpoints. In a 2 × 2 (activity sequence × film version) factorial design, 40 subjects segmented the film during viewing and recalled the film content after viewing in a detailed protocol. Segmentation behavior depended primarily on the occurrence of breakpoints and was largely unaffected by the occurrence of cuts. Cuts accompanying a breakpoint lead to more detailed recall protocols for these sections of the film.

At first sight, human behavior, as well as other observable events, consists of a continuous flow of dynamically changing states. These lack discernible structures that allow for the identification of separable units with sharply definable boundaries. Nevertheless, as the research of Newtson and his successors has shown, observers can readily segment these continuous *streams* of actions and events into disparate units, separated by boundaries, or *breakpoints*, with a high degree of intraindividual reliability and interindividual consistency (Massad, Hubbard, & Newtson, 1979; Newtson, 1973; Newtson, Engquist, & Bois, 1977). In addition, the resulting segments have been shown to have a close relationship to the structure of the subsequent cognitive representation of the observed event sequence in at least two ways (Newtson & Engquist, 1976). First, the resolution with which a given event sequence is segmented by an observer influences its memorability and retrievability. Compared with subjects who make a more coarse-grained segmentation, subjects who make more fine-grained segmentations show better recall of the observed event sequences (Hanson & Hirst, 1989; Lassiter, Stone, & Rogers, 1988). Second, those segment boundaries showing high interindividual agreement (breakpoints) typically possess perceptual characteristics that qualify them as adequate summaries of the preceding segments. This leads to higher recognition scores, as well as to an improved comprehensibility of breakpoints, as compared with points located within the segments (nonbreakpoints; Newtson & Engquist, 1976).

Newtson determined the breakpoints by having observers watch videotapes of familiar activity sequences and having them unitize the sequences by pressing a hand-held button. The observers preferably segmented the stream of actions at points at which salient changes of activity occurred—for instance, when an actor put the housing of a computer terminal to the side and took a screwdriver to remove a graphic card (an example of the activity sequence of upgrading a computer, used as experimental material in the following study). From similar findings, Newtson et al. (1977) concluded that observers base their segmentations on the monitoring of characteristics in the activity sequence, such as the type of activity, the involvement or orientation of parts of the body, or the manipulation of objects (Lasher, 1981; Wright, 1967). The occurrence of change in at least one of these characteristics then causes the observer to unitize—that is, to localize—a segment boundary within the activity sequence (Ginsburg & Smith, 1993; Reed, Montgomery, Schwartz, Palmer, & Pittenger, 1992).

Although Newtson relied on videotaped activity sequences for his studies, he presented them without any film-specific modifications, thus coming as close as possible to the conditions of natural, everyday observation. In contrast, ordinary films use specific techniques, or *formal features*, for the presentation of activity sequences—such as cuts, zooms, or pans (Huston & Wright, 1983). For example, during the computer upgrading sequence mentioned above, the viewer perspective could be changed repeatedly by abruptly cutting from one camera position to another. Similar to breakpoints, these cuts are more or less evenly dispersed across the film and are accompanied by abrupt visual changes. In addition, the film maker is largely free to choose the precise location of a cut within a given

This research was supported by the Deutsche Forschungsgemeinschaft, Bonn. Many thanks go to the Institut für den Wissenschaftlichen Film (IWF) for producing the experimental film materials and to Marcus Skupin-Alfa, Karen Oltersdorf, and Sandra Marjanovic for their help in carrying out the experiments and coding the protocols. Special thanks are due to Dan Levin and Robert Kraft for their critical suggestions in the first version of the article. Correspondence concerning this article should be addressed to S. Schwan, Department of Applied Cognitive Science, German Institute for Research on Distance Education (DIFF) at the University of Tübingen, Konrad-Adenauer-Straße 40, 72072 Tübingen, Germany (e-mail: stephan_schwan@diff.uni-tuebingen.de).

Copyright 2000 Psychonomic Society, Inc.

film sequence, so that cuts and other formal features can be conceived of as independent of the activity sequence the film depicts (Kraft, 1986). Thus, besides the structure of the depicted activity sequence, which is defined by breakpoints, an additional formal structure is imposed on the film content through the process of film editing; this structure is then defined by the cuts (Hochberg, 1986; Korac, 1988; Zettl, 1990). Therefore, from the perspective of Newtson's model, cuts may possibly qualify as features, the occurrence of which could be monitored by viewers for the purposes of segmenting the continuous flow of the events depicted in the film. In other words, we assume that, although they have no direct relation to the film content, cuts may serve a syntactical purpose similar to that of breakpoints—namely, guiding the viewer to separate the previously shown activity from the ongoing one while watching the film (Kraft, 1986).

This proposed view of cuts as structuring devices is at least partly held by a number of film scholars. The editing philosophy, especially, of classical Hollywood movies emphasizes that formal film devices, such as cuts, ought to be applied in such a manner that they pass mostly unnoticed, so as to avoid disruptions to the narrative flow. Nevertheless, a number of film scholars also assert that shots constitute the basic units of films and that the transitions between the shots serve as punctuation devices for the film content (Metz, 1974; Reisz & Millar, 1968). But by taking the syntactical purpose of cuts for granted, film scholars have primarily concentrated their discussion on their additional connotative purposes. More specifically, it is assumed that, through a careful placement of cuts, one is able to induce certain impressions in the spectator by imposing specific types of *rhythms* on the film content. For example, fast or accelerated cutting rates—such as those in the famous shower scene in Hitchcock's *Psycho*—are thought to render the impression of an event more interesting and varied than do films edited in a slow or decelerated pace (Bordwell & Thompson, 1993; Kraft, 1987; Reisz & Millar, 1968; Zettl, 1990).

The few studies that have addressed the question of a syntactical function of cuts experimentally provide mixed empirical support. In studies carried out by Carroll and Bever (1976) and Kraft (1986), viewers watched short film versions of simple activity sequences, in which the position of both cuts and activity changes was systematically varied. In both studies, the subsequent memorability of the activity sequence was largely unaffected by the occurrence of cuts but, rather, depended primarily on the number and position of activity changes shown in the films. From these findings, Kraft (1986) concluded that cuts possess no syntactical function for unitizing the flow of events depicted in the film. In a previous study by Schwan, Hesse, and Garsoffky (1998), commercially available instructional videos of considerable length and complexity were presented, and the viewers were instructed to segment the film content into meaningful units by pressing a hand-held button while watching the film. Contrary to the studies by Carroll and Bever (1976) and Kraft (1986), it could be shown that the frequency of segmentations shortly after cuts was significantly greater, in contrast to that at points in the films at which no formal film features occurred. This indicated a close relationship between cuts and the unitization processes of the viewers.

To resolve these differences, the present study was conducted in an attempt to integrate the relevant aspects of the previous studies into a single experimental design. Similar to the study directed by Schwan et al. (1998), complex and enduring activity sequences were selected with which the viewers were unfamiliar, in order to maximize the chance that cuts, as formal features, would show relevance for the segmentation behavior. The cuts used in the editing of the film material were restricted to changes of camera distance and viewing angle, both because they were used in all three studies mentioned above and because they have been shown to be by far the most frequent types of cut (Messaris, 1994). Furthermore, analogous to the study by Carroll and Bever (1976), two film versions of the same activity sequence were constructed, in which the position of cuts systematically varied in relation to the occurrence of breakpoints—that is, the salient changes in the activity sequence itself. In the film version with *cuts at breakpoints*, cuts were inserted at points at which a breakpoint occurred in the depicted activity. In the film version with *cuts at nonbreakpoints*, cuts were inserted at points at which a nonbreakpoint occurred in the depicted activity. The cuts-at-breakpoints version thus contained both breakpoints accompanied by a cut and nonbreakpoints without a cut, whereas the cuts-at-nonbreakpoints version contained the same breakpoints without a cut and the same nonbreakpoints accompanied by a cut. Finally, and most important, to allow for a direct comparison between segmentation behavior during viewing and during subsequent recall, both measures were used in the present study.

On the assumption that cuts constitute a second, additional set of cues besides the breakpoints inherent in the depicted activity sequence, in the context of the present study, the syntactical purpose of cuts could be one of four possibilities. First, in accordance with Kraft's (1986) findings, viewers could ignore cuts as syntactical devices and rely solely on breakpoints (*irrelevance hypothesis*). Conversely, and as a second possibility, viewers could consider cuts rather than breakpoints as the primary syntactical cues (*substitution hypothesis*). Third, viewers could consider both cuts and breakpoints as syntactical cues in parallel (*supplementation hypothesis*). Fourth, viewers could attend primarily to breakpoints, but in a more accentuated manner when accompanied by cuts, whereas cuts alone would have no segmentational effect (*accentuation hypothesis*; Carroll & Bever, 1976).

Furthermore, the syntactical function of cuts not only should manifest itself in the segmentations during viewing but should also influence the subsequent cognitive representation of film content. Newtson and his successors were able to establish a strong empirical correspondence between segmentation behavior and the characteristics of subsequent recall, in terms of both its detail and its completeness (Hanson & Hirst, 1989; Lassiter et al., 1988; Newtson & Engquist, 1976). Thus, in the context of the present study, the detail and completeness of recall can be expected to vary according to the frequency and position of segmentations during viewing, which, in turn, should bear a relationship to the occurrence of cuts, depending on their syntactical purpose, as specified by one of the four hypotheses above.

More specifically, the film version that elicits greater fine-grained segmentation behavior should also lead to a more detailed overall recall. A similar relationship should also be demonstrable at a local level. Ideally, a point in the film leading to a segmentation during viewing should also lead to a unitization at the level of cognitive representation, which would be indicated by a new entry in the recall protocol. In other words, local sections of a film containing such segmentation-eliciting points should lead to a more detailed recall—that is, to a greater number of different entries within a given recall protocol, when compared with local sections containing no such points. Finally, Newtson and Engquist (1976) found that breakpoints—that is, event-inherent structural boundaries—summarize the preceding event segment in a particularly informative way, which therefore leads to better recognition scores, in contrast to nonbreakpoints. In the light of these findings, one might conjecture that segmentations departing from these highly informative breakpoints could possibly distract the viewers and lead to a reduced completeness of the subsequent recall.

With regard to the syntactical function of cuts, for each of the four different hypotheses mentioned above, a different cluster of results can be predicted.

In the case of the irrelevance hypothesis, viewers ignore cuts and base their segmentation solely on breakpoints. Therefore, because the two film versions are similar with regard to their breakpoints, they ought to lead to a roughly similar number of segmentations, as well as to similar detail and completeness of recall. Conversely, in either version, local sections containing a breakpoint (i.e., an event-inherent boundary) should elicit both more segmentations and more entries in the recall protocols than do local sections containing no breakpoint, irrespective of the occurrence of a cut.

In the case of the substitution hypothesis, viewers base their segmentation on cuts and ignore breakpoints. Since both films contain an equal number of cuts, again, no differences in the overall amount of segmentation and in the overall detail of recall would be expected. But within the films, local sections containing a cut should elicit both more segmentation decisions and more entries in the recall protocols than do local sections containing no cut, irrespective of the occurrence of a breakpoint. In addition, the version in which the cuts are placed at less informative nonbreakpoints should lead to a less complete recall than does the version in which the cuts are placed at breakpoints.

In the case of the supplementation hypothesis, viewers attend both to breakpoints and to cuts. In the cuts-at-breakpoints version, cuts and breakpoints are placed together, whereas in the cuts-at-nonbreakpoints version, their position differs. Therefore, for the latter version, a greater overall amount of segmentation, as well as greater overall detail of recall, would be expected. In addition, the completeness of recall should be lesser in the cuts-at-nonbreakpoints version, owing to the distracting effect of cut-induced boundaries at less informative nonbreakpoints. Differences between the two versions can also be expected on a local level. In the cuts-at-nonbreakpoints version, the local sections containing either a breakpoint without a cut or a cut without a breakpoint should lead to roughly equal numbers of entries in the recall protocols. On the other hand, the corresponding sections of the cuts-at-breakpoints version contain a combination either of a breakpoint with a cut or of a nonbreakpoint without a cut and, in the latter case, should therefore lead to a less detailed recall than in the former case.

Finally, according to the accentuation hypothesis, viewers attend primarily to breakpoints, but in a more distinct manner when accompanied by cuts, whereas cuts alone have no segmentational effect. Thus, in contrast to the cuts-at-nonbreakpoints version, for the cuts-at-breakpoints version, a greater overall amount of segmentation, as well as greater overall detail of recall, would be expected. The completeness of recall should, however, be comparable in both versions. On a local level, in both versions, sections containing a breakpoint (i.e., an event-inherent boundary) should elicit both more segmentation decisions and more recall entries than do local sections containing no breakpoint. But in the cuts-at-breakpoints version, the number of recall entries at sections containing a breakpoint should be higher than in the cuts-at-nonbreakpoints version.

METHOD

Subjects

A total of 40 subjects unfamiliar with the activities depicted in the films (33 female and 7 male students) participated in the study. Their average age was 25. They were paid for their participation.

Stimulus Materials

To determine the syntactical relevance of both event-inherent structural boundaries (breakpoints) and formal filmic means, a systematic variation of their relative positions was required. Accordingly, first, two different event sequences were videotaped; their breakpoints were then determined in conformity with the procedure proposed by Newtson (1973), after which two film variants were edited from each event sequence, systematically differing in the position of cuts relative to the breakpoints.

In order to control for possible effects of the depicted event sequence, two different activities were selected for video recording. These activities consisted of the upgrading of a computer and the cleaning of a police pistol. Each activity lasted approximately 15.5 min, the action of which included a number of different action steps. Each activity was performed by an actor routinely familiar with the procedures involved. He was seated at a table, with the computer or the pistol placed in front of him. In the case of the computer, he dismounted the case, removed a number of inserted cards, exchanged the mainboard, and, finally, reassembled it. In the case of the pistol, the actor disassembled it, carefully cleaned the parts with brushes and cloths, and finally reassembled it. Each activity was simultaneously recorded from three camera positions: a medium shot, taken from a position in front of the actor; a medium close-up, taken from a high angle behind the actor over his right shoulder; and a medium close-up, taken from about 45° to one side of the front position.

The video record of the upgrading of the computer, as seen from the frontal perspective, was then shown to a group of 17 computer experts familiar with the depicted activity. Likewise, the video record of the pistol cleaning, as seen from the frontal perspective, was then shown to a group of 17 policemen. The sessions were run individually. Each viewer was instructed to watch the film attentively and to press a button whenever, in his or her opinion, a meaningful part of the activity was concluded and a new part began. The occurrences of buttonpresses during the viewing of the film were automatically recorded by a computer program.

From this data, breakpoints were determined. Breakpoints represent locations at which observers perceive a boundary of a segment in the depicted activity with high interindividual agreement. To compute the breakpoints, the video record of the activity was divided into short 2-sec intervals, and for each interval, the observed frequency of buttonpresses was determined. In accordance with the computing procedure of Newtson and his successors (Hanson & Hirst, 1989; Newtson, 1973), those intervals showing frequencies of buttonpresses that lay at 1.65 standard deviations above the mean were considered to be breakpoints. Intervals containing at least four buttonpresses for the activity of upgrading a computer or at least five buttonpresses for the activity of cleaning a pistol counted as breakpoints. By contrast, intervals with frequencies of buttonpresses at 1.65 standard deviations below the mean were considered nonbreakpoints. For both activities, intervals containing no buttonpresses counted as nonbreakpoints. According to these definitions, the activity of upgrading a computer contained 60 breakpoints and 163 nonbreakpoints, whereas the activity of cleaning a pistol contained 58 breakpoints and 184 nonbreakpoints.

As a next step, two different versions of each activity (upgrading a computer and cleaning a pistol) were edited. Both versions of each activity depicted the complete activity at full length, but with abruptly changing perspectives by cutting back and fourth between the frontal perspective and one of the other camera positions, which corresponded to the most frequent type of cut (Messaris, 1994). The versions differed with respect to the points of time at which these cuts occurred. In the cuts-at-breakpoints versions, cuts were inserted at points at which a breakpoint occurred in the activity. In addition, care was taken to ensure that the period between adjacent cuts was at least 4 sec, that the periods between adjacent cuts varied considerably in length, and that the occurrence of cuts did not form a rhythmic pattern. Taking these principles into account, the cuts-at-breakpoints versions of upgrading a computer and cleaning a pistol contained 34 and 32 cuts; these corresponded to 2.2 and 2.1 cuts per minute, respectively.

Accordingly, in the cuts-at-nonbreakpoints versions, cuts were inserted at points at which a nonbreakpoint occurred in the activity. Again, care was taken to ensure that the period between adjacent cuts was at least 4 sec, that the cuts did not form a rhythmic pattern, and that the mean period between cuts, as well as the proportions of the different camera positions in the films, were matched with the cuts-at-breakpoints versions. In addition, the cuts were placed so that all the breakpoints were seen from the frontal perspective. Taking these principles into account, the cuts-at-nonbreakpoints versions of upgrading a computer and cleaning a pistol both contained 32 cuts, which corresponded to 2.2 cuts per minute.

In short, the cuts-at-breakpoints versions contained both breakpoints accompanied by a cut and nonbreakpoints without a cut, whereas the cuts-at-nonbreakpoints versions contained the same breakpoints without a cut and the same nonbreakpoints accompanied by a cut. During the upgrading of a computer sequence, for example, a breakpoint occurred when the actor placed the encasement of the computer terminal to the side and proceeded to remove the graphic card. In the cuts-at-breakpoints version, this was accompanied by a cut that switched the camera position from the frontal perspective to a position approximately 45° to one side of the frontal perspective. By contrast, in the cuts-at-nonbreakpoints version, the whole activity was consistently shown without any cut from the frontal perspective. Similarly, a nonbreakpoint occurred while the actor connected several wires to the mainboard. In the case of the cuts-at-nonbreakpoints version, this was accompanied by a cut that shifted the angle of the camera from 45° to one side to the frontal perspective, whereas in the cuts-at-breakpoints version, the whole activity was continuously shown from a 45° angle.

Design and Procedure

The experiment was run as a complete two-factorial design, including film version (cuts-at-breakpoints vs. cuts-at-nonbreakpoints) and depicted activity (upgrading a computer vs. cleaning a pistol) as factors, with each cell containing 10 subjects.

The subjects were tested individually in experimental sessions, each of which lasted about 1 h. They were told that they were participating in an experiment about the perception of activities. In accordance with the procedure of Newtson and his successors, they received the following instructions: "In this project, we are interested in how observers perceive work procedures. You are going to watch a video film that lasts about fifteen minutes. A work procedure can be divided into a number of individual steps. Such a division can be done in a more or less detailed manner, i.e. you can divide the whole procedure into many small steps or into a number of coarser steps. Let us take the preparation of a mixed salad as an example. One can distinguish such steps as 'taking the bottle with salad oil off the shelf,' 'opening the bottle,' 'taking a spoonful of oil' . . . 'washing the paprika,' 'slicing the paprika,' and so on. One might also summarize the small steps into 'pouring salad oil into a bowl' and 'preparing the paprika.' Of course, smaller or larger steps can also be defined. We are interested in which divisions make the most sense to you. There are no right or wrong divisions. We are interested in your personal opinion regarding the division of work steps. Please watch the video carefully. For this task, you have a hand-held button, which you can press during the video presentation. Always press the button only when, in your opinion, one work step ends and a new one begins. Please watch the video attentively because afterwards your memory of its content will be tested."

Special care was taken in the instructions to avoid mentioning formal aspects of the film, so as not to direct the attention of the viewers to cuts, but, rather, to emphasize the content of the film.

The subjects were seated in front of a television monitor and were given a hand-held button. After a short practice video, the experimental film was shown. The time-stamps of the buttonpresses, as indicators of segmentation decisions during the film presentations, were recorded automatically on an Apple Macintosh synchronized with the videotape by using the software package PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993).

After a short break following the viewing of the film, the subjects were instructed as follows: "Please try to recall the content of the video you have just watched as accurately as possible. Write down all activities carried out by the person in the film you can remember. Write down the activities in their order of appearance. Use the first line of the sheet of paper for the first activity shown in the film, use the second line for the second activity, and so on. Do not describe the film by a few general activities but instead try to divide the film content into as many steps as possible. Specify every step with a short sentence. If the person used an object unfamiliar to you or whose name you do not know, describe it in your own words."

After finishing their account, the subjects were required to guess the number of cuts that they had seen in the film. Finally, at the end of the session, the subjects were debriefed and received payment for their participation.

Analysis of the Recall Protocols

To provide a basis for relating the recall protocols to the filmic content, each film was transcribed into a chronological and detailed list of the actions it depicted. Actions were considered to be specific manipulations of objects and were defined both by the objects that were involved in the action and by the accompanying movement patterns. Thus, a change of object or a substantial change in the movement pattern counted as the beginning of a separate action—for example, *opening the wrapping*, *taking the mainboard*, *inserting the mainboard into a slot*, *wiring the mainboard*, and so on. As a result, for each film, a list of actions was drawn up, which consisted of a short description of the action, the time at which it began, its length, and the time it ended. For the upgrading a computer film, the list consisted of 255 actions, and for the cleaning a pistol film, the list consisted of 272 actions.

Two coders then mutually assigned each entry in the recall protocol recorded by each of the subjects to either a single corresponding action or a sequence of successive actions in the film transcription. From a total of 1,651 recall protocol entries, 1,418, or 86%, could then be successfully assigned to actions or action sequences listed in the film transcription. Thus, each recall protocol was mapped onto the film transcription. This specified which of the individual actions or action sequences in the film were recalled by each of the subjects. To determine intercoder reliability, a third coder analyzed 30% of the recall protocols independently. With regard to the number of actions of the film transcription recounted in the recall protocols, the interrater correlation was $r = .89$. With regard to the mean length of sequences of actions in the film transcription addressed by a single entry in the recall protocol, the interrater correlation was $r = .95$.

In principle, the number of recall protocol entries can be conceived of as a function of two distinct features of the recall—namely, its completeness and its detail. For instance, a subject may possess an almost complete but molar cognitive representation of the film content, thus describing the entire film by means of a few entries, each encompassing large sections of the film. Conversely, a similar number of recall entries may stem from a subject who possesses a more fine-grained but incomplete cognitive representation of the film content. The film would then be described by means of a few detailed entries encompassing small sections of the film, leaving much of the film content unmentioned.

By mapping the recall protocols onto the film transcription, it was possible to calculate independent indices of recall completeness and recall detail. In order to measure recall completeness, for each subject, the overall number of actions that his or her recall protocol contained was determined. In order to measure recall detail, for each subject, the mean length of his or her recall protocol entries, in terms of the number of successive actions they encompassed, was determined. For both the cuts-at-breakpoints and the cuts-at-nonbreakpoints versions, recall completeness and recall detail were largely independent ($r = .07, p > .10$, and $r = -.33, p > .10$, respectively).

The segmentational function of breakpoints or cuts should also be demonstrable on a local level. As was indicated above, local sections of a film containing a segmentation-eliciting point (i.e., a breakpoint or a cut) should lead to a more detailed recall—namely, to an increased number of different entries within a given recall protocol—unlike local sections that contain no such points. To address this issue, two sets of sections with a similar length were defined in the film transcriptions—namely, sections around breakpoints containing a cut in the cuts-at-breakpoints version, and sections around nonbreakpoints containing a cut in the cuts-at-nonbreakpoints version. Since the number of entries in the recall protocols corresponding to these sections should depend not only on the occurrence of a breakpoint or a cut, but also on the number of actions contained in a section, the sections were selected on the basis of the following two criteria. First, to allow for differences in the number of recall protocol entries, each section should contain at least two distinct actions. Thus, the length of the sections was set at twice the mean length of the actions depicted (7 sec). Second, both sets of sections were matched for the number of actions they contained according to the film transcriptions, resulting in two sets of 13 intervals in the case of upgrading a computer and two sets of 6 intervals in the case of cleaning a pistol.

RESULTS

Segmentation of the Film Content

For each subject, the number of segmentations (i.e., buttonpresses) during viewing was determined and compared across film version (cuts-at-breakpoints vs. cuts-at-nonbreakpoints) and the activity depicted (upgrading a computer vs. cleaning a pistol) in a 2 $\times$ 2 between-subjects analysis of variance (ANOVA). The mean square error (MS_e) was 495.6. Cleaning a pistol led to significantly more segmentations than did upgrading a computer [49.5 vs. 33.3 segmentations; $F(1,36) = 5.26, p < .05$]. Neither version ($F < 1$) nor interaction of film version with depicted activity ($F < 1$) had a significant effect on the overall number of segmentations during the viewing of the films.

To assess the respective influence of breakpoints and cuts on the segmentation decisions, an index of the proportion of segmentation decisions (i.e., the probability of pressing a button) taken by each subject was computed for two different sets of intervals 2 sec in length. The first set consisted of all those intervals at breakpoints containing a cut in the cuts-at-breakpoints version. Thus, in the cuts-at-nonbreakpoints condition, these intervals only contained a breakpoint, whereas, in the cuts-at-breakpoints condition, they contained both a breakpoint and a cut. Conversely, the second set consisted of all those intervals at nonbreakpoints containing a cut in the cuts-at-nonbreakpoints version. Thus, in the cuts-at-breakpoints condition, these intervals contained neither a breakpoint nor a cut, whereas, in the cuts-at-nonbreakpoints condition, they contained only a cut.

For each subject, the index computed the proportion of the percentage of his or her segmentations in the specific set of intervals (i.e., intervals at breakpoints and intervals at nonbreakpoints, respectively) to the percentage of his or her segmentations in all the intervals of the film. It was determined according to the formula $(seg_{int}/n_{int})/(seg_{all}/n_{all})$, where seg_{int} is the number of the subject's segmentations in the set of intervals, n_{int} is the number of intervals of the set, seg_{all} is the overall number of the subject's segmentations, and n_{all} is the overall number of

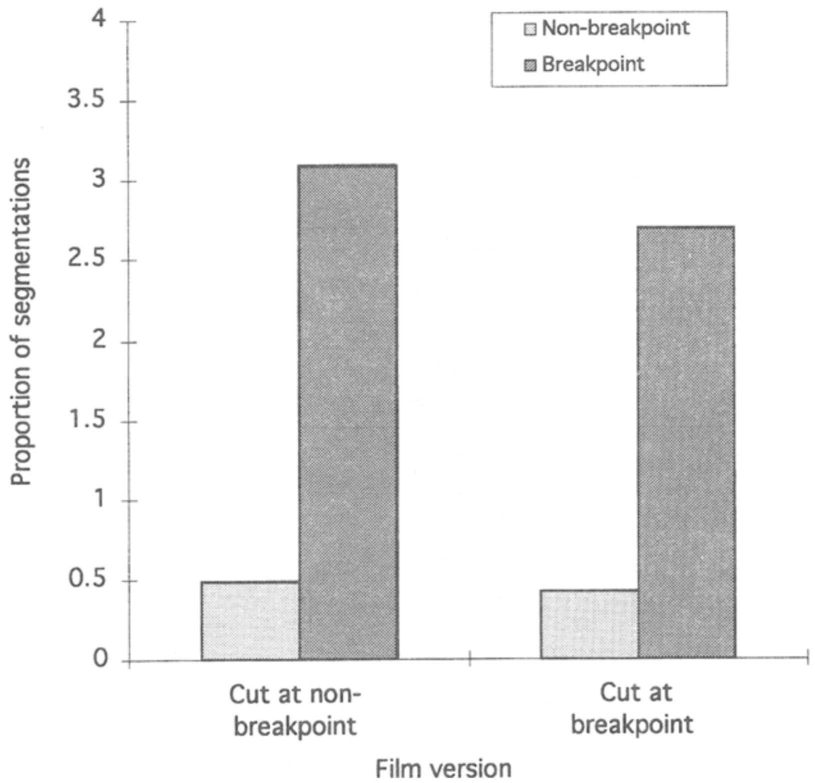

Figure 1. Number of segmentations as a function of nonbreakpoint and breakpoint intervals for the two film versions.

intervals in the film. Thus, the formula accounts for differences in the overall number of segmentations between subjects. It can also be interpreted as the proportion of observed segmentations in relation to the number of segmentations expected on the basis of the overall number of the subject's segmentations and the number of intervals contained in the set; the indices below 1 represent a relative decrease in frequency, and indices above 1 signal a relative increase in frequency, respectively.

The resulting indices were analyzed according to a 2 (type of interval: breakpoint vs. nonbreakpoint) × 2 (film version: cuts-at-breakpoints vs. cuts-at-nonbreakpoints) × 2 (depicted activity: upgrading a computer vs. cleaning a pistol) ANOVA, with type of interval as a within-subjects factor and film version and depicted activity as between-subjects factors. For intervals at nonbreakpoints, the mean index was markedly below 1 (0.45; $SD = 0.57$), whereas for intervals at breakpoints, the mean index was markedly above 1 (2.89; $SD = 1.23$). This difference was highly significant [$F(1,36) = 100.05$, $MS_e = 1.19$, $p < .001$]. Whereas, on average, about 25% of the intervals at breakpoints led the viewer to segment, segmentation was carried out for only 4% of the intervals at nonbreakpoints (see Figure 1). No other factor, especially film version [$F(1,36) = 1.44$] and the interaction of interval type with film version ($F < 1$), was significant. Thus, whereas the presence or absence of an event-inherent structural feature led to substantial differences in the segmentation behavior of the viewers, the occurrence of a cut had no influence on their segmentation behavior.

Recall of the Actions Depicted in the Film

The percentage of recalled actions was determined as an indicator of recall completeness by mapping the subject's recall protocol entries onto the film transcript. The data were compared across film version (cuts-at-breakpoints vs. cuts-at-nonbreakpoints) and the depicted activity (upgrading a computer vs. cleaning a pistol) in a 2 × 2 ANOVA, with film version and depicted activity as between-subjects factors. On average, the subjects recalled about 57% of the actions depicted in the films. The difference between the depicted activities was highly significant [$F(1,36) = 68.1$, $MS_e = 669.4$, $p < .001$], with upgrading a computer leading to a higher percentage of recalled actions than cleaning a pistol (71.9% vs. 42.6%). Neither film version ($F < 1$) nor the interaction of film version with depicted activity ($F < 1$) had a significant effect on the completeness of recall.

Whereas, for cuts-at-breakpoints, no relationship between the completeness of the recall protocols and the frequency of the segmentation could be confirmed ($r = -.14$, $p > .10$), for the cuts-at-nonbreakpoints version, a substantial negative correlation was found ($r = -.66$, $p < .01$).

The mean length of the subject's recall protocol entries, in terms of the number of actions they encompass, was determined as an indicator of detail of recall. Again, the

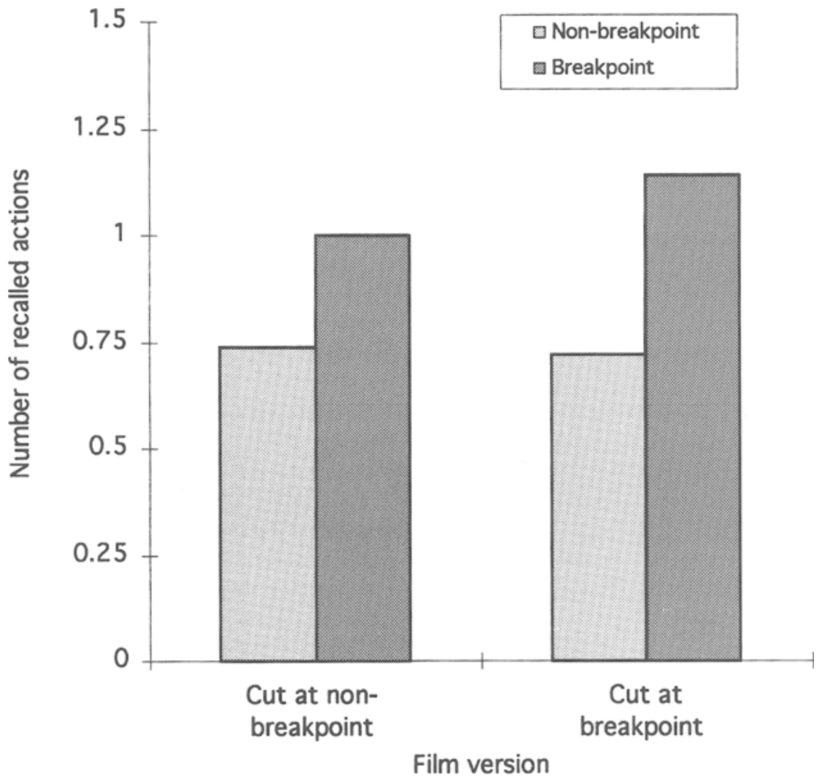

Figure 2. Number of recalled actions as a function of nonbreakpoint and breakpoint intervals for the two film versions.

detail scores were compared across film version (cuts-at-breakpoints vs. cuts-at-nonbreakpoints) and depicted activity (upgrading a computer vs. cleaning a pistol) in a 2 × 2 ANOVA, with film version as a within-subjects factor and depicted activity as a between-subjects factor. Overall, the mean length of the action sequences recalled was 7.8 actions. The difference between the depicted activities approached significance [$F(1,36) = 3.26$, $MS_e = 14.3$, $p < .10$], with upgrading a computer leading to slightly longer recalled action sequences than cleaning a pistol (8.86 vs. 6.70). Neither film version ($F < 1$) nor the interaction of film version with depicted activity ($F < 1$) had a significant effect on the detail of recall. In summary, the different placement of cuts in the two versions affected neither the overall completeness nor the detail of the subsequent recall protocols.

The mean length of the action sequences recalled indicated a substantial negative relationship to the frequency of segmentation, both in the cuts-at-breakpoints version ($r = -.60$, $p < .01$) and in the cuts-at-nonbreakpoints ($r = -.50$; $p < .05$) versions. Thus, a finer unitization, indicated by more segmentations, also leads to a more detailed recall, indicated by shorter action sequences in the recall protocols.

To assess the respective influence of breakpoints and cuts on a local level, the number of entries in the recall protocols was determined for sections around breakpoints containing a cut in the cuts-at-breakpoints version and for sections around nonbreakpoints containing a cut in the cuts-at-nonbreakpoints version, as was described in the Method section. For each subject, the mean number of recall protocol entries at each section type was computed and analyzed. This was conducted according to a 2 (type of interval: breakpoint vs. nonbreakpoint) × 2 (film version: cuts-at-breakpoints vs. cuts-at-nonbreakpoints) × 2 (depicted activity: upgrading a computer vs. cleaning a pistol) ANOVA, with type of interval as a within-subjects factor and film version and depicted activity as between-subjects factors. As Figure 2 shows, the subjects recalled significantly more actions at intervals with breakpoints than at intervals with nonbreakpoints [$F(1,36) = 79.70$, $MS_e = 0.03$, $p < .001$]. In addition, a significant interaction between type of interval and film version [$F(1,36) = 4.79$, $MS_e = 0.03$, $p < .05$] revealed that, at intervals with breakpoints, the subjects recalled more actions if the breakpoint was accompanied by a cut. Again, the depicted activity revealed a significant difference [$F(1,36) = 6.0$, $MS_e = 0.09$, $p < .05$], with upgrading a computer leading to more actions recalled than cleaning a pistol (0.98 vs. 0.82).

DISCUSSION

Although films are sometimes described as simple reproductions of reality, they differ from reality in a number of fundamental aspects through which the specificity of film, as a medium, is defined (Arnheim, 1957). One important difference between reality and its filmic representation lies in their respective structures. In reality, the observable flow of activities or events is structured by event-inherent features—namely, by breakpoints marking boundaries between separate events. Narrative films also possess this content-related structure. However, this structure is accompanied by the formal structure of the film material, primarily defined by the occurrence of cuts that mark the boundaries between separate shots. Thus, in the case of films, cuts constitute a second, media-specific set of features, to which viewers could possibly attend so as to segment the continuous stream of film content into separate units, thereby performing a syntactical function similar to breakpoints.

The results of the present study only partly confirm this assumption. The two film variants, which differed in the position of cuts in relation to breakpoints, did not lead to any substantial differences, either in the overall number of segmentations during the film presentation or in the overall completeness and detail of subsequent recall. Nevertheless, in the case of cuts-at-nonbreakpoints, a significant negative correlation between the number of segmentations and the completeness of the recall was found. In addition, on a local level, the occurrence of a breakpoint evoked both more segmentations and a higher number of distinctly recalled activities. The presence or absence of a cut, however, led neither to more segmentations nor to a higher number of recalled activities in general. Yet, compared with the occurrence of a breakpoint alone, the accompaniment of a breakpoint by a cut led to a higher number of recalled activities.

These findings may be interpreted in terms of the four hypotheses specifying the possible relationships between the structure of the film content and the structure of the film material. Viewers could either ignore cuts and only attend to breakpoints (irrelevance hypothesis) or could rely on cuts instead of breakpoints (substitution hypothesis); they could consider both types of cues simultaneously (supplementation hypothesis), or they could ignore isolated cuts but pay more attention to breakpoints when accompanied by a cut (accentuation hypothesis).

With regard to the segmentation behavior of the spectators during viewing, the results were mostly consistent with the irrelevance hypothesis. Overall, the cuts-at-breakpoints and cuts-at-nonbreakpoints versions elicited a roughly equal number of segmentations. Independent of version, the frequency of segmentations was about seven times higher shortly after the occurrence of a breakpoint, when compared with film intervals containing no such breakpoint. By contrast, the frequency of segmentations remained largely unaffected by the occurrence of a cut. This result stands in contrast to the notions of a substitutional, a supplementary, or an accentuating function of cuts. Taken together, these results indicate that, in order to perceptually segment the film content during viewing, spectators attended predominantly to event-inherent breakpoints and mostly ignored the formal structure of the film material, as represented by cuts.

This absence of a structuring effect of cuts during viewing stands in contrast to the previous study of Schwan et al. (1998). Here, for commercially available instructional films, higher segmentation rates were obtained shortly after the occurrence of different types of formal filmic means. The latter included cuts that implied a change of perspective and distance (similar to those used in the present study). But the two studies also differ with respect to the relationship between the filmic means and the event-inherent structural features: In the present study, the breakpoints were determined and their position relative to the cuts was controlled, whereas this was not the case in the previous study. Thus, one possible explanation for the discrepancy in the findings is that the observed increase of segmentations shortly after a cut in the Schwan et al. study was based on a coincidence of cuts with event-inherent structural boundaries and that not the cuts themselves, but the accompanying breakpoints, caused the spectators to segment the film content. In other words, it seems plausible that professional film makers intuitively place cuts-at-breakpoints in order to adapt the formal film structure to the already existing structure of the depicted flow of events.

Turning now to the structure of the subsequent recall, most of the findings in the present study are compatible with the accentuation hypothesis. Again, both versions—cuts-at-breakpoints and cuts-at-nonbreakpoints—elicited a roughly equal overall detail and completeness of recall. Thus, they were not influenced by the position of cuts in relation to the breakpoints. Also, viewers recalled a greater number of distinct action steps in intervals containing a breakpoint than in intervals containing no breakpoint. But this segmentational effect of breakpoints was heightened where the breakpoint was accompanied by a cut, indicating that the cut accentuated the existence of the event-inherent structural boundary.

The present findings differ from Kraft's (1986) study, where the number of actions recalled strongly depended on the number of depicted activities but was not influenced by the presence or absence of cuts. But because Kraft's study did not control the position of the cuts in relation to the event-inherent structures, its findings are difficult to interpret with regard to the assumption of an accentuating function of cuts. In line with this argumentation, the study of Carroll and Bever (1976), in which the position of cuts and event-inherent boundaries was controlled, reported an accentuating effect of cuts, similar to the present study. This was indicated by more pronounced differences in the reaction times between film probes that were separated by a change of activity and an accompanying cut, in contrast to an activity change alone.

Taken together, these findings suggest that both the segmentation processes during viewing and the structural features of the subsequent cognitive representations were, to a large extent, dependent on the event-inherent features of the depicted activity sequences. Although the viewers were not accustomed to the activities depicted, they were obviously able to identify those salient structural aspects in the flow of activities that experts consider unit boundaries. In contrast, cuts alone—without being accompanied by a corresponding event-inherent structural feature—had no substantial effects on the syntactical processing of the film content. Thus, no empirical evidence was found by which cuts could either substitute or supplement breakpoints with regard to their syntactical purpose.

With regard to the third type of syntactical purpose of cuts outlined above—namely, cuts that accompany breakpoints, thereby facilitating their structural processing—the results are not so clear-cut. Whereas an accentuating effect of cuts could not be confirmed for the segmentation processes during viewing, for the subsequent cognitive representation of the content, it was found that the latter processing stage seemed to be more sensitive to the occurrence of cuts than was the former. This divergence indicates that the relationship between both levels of syntactical processing cannot simply be conceived of as a one-to-one mapping, as the Newtson model suggests. Nevertheless, both process stages show strong dependencies, as the substantial correlation between the frequency of segmentations and the recall detail makes clear.

Depending on the position of the cuts in relation to breakpoints, this dependency was further demonstrated by the differing influence of segmentation frequency on the completeness of the subsequent recall. Whereas in the cuts-at-breakpoints version, recall completeness was largely unaffected by the frequency of segmentations, in the cuts-at-nonbreakpoints version the recall completeness showed an inverse relationship to segmentation frequency.

One possible explanation could be that this difference between film versions is related to the respective cognitive capacity required to identify and process a breakpoint as a structural boundary. This argument rests on a number of assumptions. First, it is assumed that the identification of a breakpoint is a capacity-consuming process, especially for observers unfamiliar with the activity sequence, as was the case in the present experiment. Second, the identification of a breakpoint marks a point at which the actions of the preceding activity segment are recoded and integrated into an enduring cognitive representation (Carroll & Bever, 1976), a process also requiring some amount of cognitive capacity. As both processes compete for cognitive capacity, the capacity required for constructing the cognitive representation may exceed the capacity limit of a viewer, owing to the cognitive effort of identifying a breakpoint. Under these circumstances, the number of breakpoints a viewer identifies in a given activity sequence should be inversely related to the completeness of his or her subsequent recall, as was the case in the cuts-at-nonbreakpoints version. If, on the other hand, a breakpoint is accompanied by a cut, the breakpoint should be more salient for the viewer, and, therefore, the cognitive effort required to identify it would be substantially reduced, leaving enough cognitive capacity for building up the cognitive representation of the activity sequence. In addition, the occurrence of a cut leads to an orienting reaction that enhances the capacity available for cognitive processing, as was shown by the studies of Lang, Geiger, Strickwerda, and Sumner (1993) and Garsoffky, Schwan, and Hesse (1998). Under these circumstances, the memorability of an event sequence should remain, to a large extent, unaffected by the number of identified breakpoints, as was the case in the cuts-at-breakpoints version.

To sum up, these tentative considerations indicate that cuts can facilitate the processing of a breakpoint both by making it more salient and by activating additional cognitive capacity. Although this view is compatible with the accentuation hypothesis, it should be kept in mind that issues concerning the cognitive capacity requirements of breakpoints, cuts, or their combinations were beyond the scope of this study and should be separately addressed in future investigations.

On the whole, the pattern of results in the present study suggests that formal film structure, as constituted by cuts, should not be conceived of as a content-independent device for shaping the segmentational processes of the viewers. Rather, it would seem to serve its syntactical purpose primarily by being coupled with the structure of the film content, thereby highlighting the structural boundaries of the depicted events and activities.

REFERENCES

Arnheim, R. (1957). *Film as art.* Berkeley: University of California Press.

Bordwell, D., & Thompson, K. (1993). *Film art: An introduction* (4th ed.). New York: McGraw-Hill.

Carroll, J. M., & Bever, T. G. (1976). Segmentation in cinema perception. *Science*, **191**, 1053-1055.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavioral Research Methods, Instruments, & Computers*, **25**, 257-271.

Garsoffky, B., Schwan, S., & Hesse, F. W. (1998). Zum Einfluß von Filmschnitt und Bildausschnittsgröße auf die Aktivation sowie auf die Gedächtnisrepräsentation formaler Bildcharakteristika [The influence of film cuts and camera distance on viewer activation and on the cognitive representation of formal film characteristics]. *Medienpsychologie*, **10**, 110-130.

Ginsburg, G. P., & Smith, D. L. (1993). Exploration of the detectable structure of social episodes: The parsing of interaction specimens. *Ecological Psychology*, **5**, 195-233.

Hanson, C., & Hirst, W. (1989). On the representation of events: A study of orientation, recall, and recognition. *Journal of Experimental Psychology: General*, **118**, 136-147.

Hochberg, J. (1986). Representation of motion and space in video and cinematic displays. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 22-1–22-64). New York: Wiley.

Huston, A. C., & Wright, J. C. (1983). Children's processing of television: The informative functions of formal features. In J. Bryant & D. R. Anderson (Eds.), *Children's understanding of television* (pp. 35-68). New York: Academic Press.

Korac, N. (1988). Functional, cognitive and semiotic factors in the development of audiovisual comprehension. *Educational Communication & Technology Journal*, **36**, 67-91.

Kraft, R. N. (1986). The role of cutting in the evaluation and retention of film. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 155-163.

KRAFT, R. N. (1987). Rules and strategies of visual narratives. *Perceptual & Motor Skills*, **64**, 3-14.

LANG, A., GEIGER, S., STRICKWERDA, M., & SUMNER, J. (1993). The effects of related and unrelated cuts on television viewers' attention, processing capacity, and memory. *Communication Research*, **20**, 4-29.

LASHER, M. D. (1981). The cognitive representation of an event involving human motion. *Cognitive Psychology*, **13**, 391-406.

LASSITER, G. D., STONE, J. L., & ROGERS, S. L. (1988). Memorial consequences of variation in behavior perception. *Journal of Experimental Social Psychology*, **24**, 222-239.

MASSAD, C. M., HUBBARD, M., & NEWTSON, D. (1979). Selective perception of events. *Journal of Experimental Social Psychology*, **15**, 513-532.

MESSARIS, P. (1994). *Visual literacy: Image, mind, and reality*. Boulder, CO: Westview.

METZ, C. (1974). *Film language: A semiotics of the cinema*. New York: Oxford University Press.

NEWTSON, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality & Social Psychology*, **28**, 28-38.

NEWTSON, D., & ENGQUIST, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology*, **12**, 436-450.

NEWTSON, D., ENGQUIST, G., & BOIS, J. (1977). The objective basis of behavior units. *Journal of Personality & Social Psychology*, **35**, 847-862.

REED, E. S., MONTGOMERY, M., SCHWARTZ, M., PALMER, C., & PITTENGER, J. B. (1992). Visually based descriptions of an everyday action. *Ecological Psychology*, **4**, 129-152.

REISZ, K., & MILLAR, G. (1968). *The technique of film editing* (2nd. ed.). London: Focal Press.

SCHWAN, S., HESSE, F. W., & GARSOFFKY, B. (1998). The relationship between formal filmic means and the segmentation behavior of film viewers. *Journal of Broadcasting & Electronic Media*, **42**, 85-97.

WRIGHT, H. F. (1967). *Recording and analyzing child behavior*. New York: Harper.

ZETTL, H. (1990). *Sight sound motion: Applied media aesthetics* (2nd ed.). Belmont, MA: Wadsworth.

(Manuscript received May 12, 1998;
revision accepted for publication March 25, 1999.)