

Verbalizing events: Overshadowing or facilitation?

MARKUS HUFF AND STEPHAN SCHWAN
Knowledge Media Research Center, Tübingen, Germany

Verbal overshadowing refers to the surprising effect whereby additional verbal information about a visual stimulus hinders its subsequent recognition. In two experiments, we analyzed the validity of this effect for event recognition across various conditions of presentation and testing. Participants observed events that were either followed (Experiment 1) or preceded (Experiment 2) by a verbal description. Results showed that verbal overshadowing occurred when the verbal description was presented after the visual presentation, independent of the distractor type. However, when the verbal description preceded the event, recognition performance was seen to improve when distractor items incompatible with the verbal description were used. The findings were interpreted in terms of two interacting mental representations, which differ both in their level of abstraction and in their accessibility.

The question whether verbal descriptions of a visual stimulus foster or hinder its subsequent recognition has received much attention in cognitive psychology (e.g., Loftus, Miller, & Burns, 1978). During the last two decades, research has mainly focused on verbal overshadowing—that is, a decrease in recognition performance due to verbalization. Recently, the complementary possibility of an increase in recognition performance (verbal facilitation) by varying the experimental design came into the focus of research (e.g., Brown & Lloyd-Jones, 2005, 2006; Itoh, 2005). Hence, the present article aims to extend previous research on verbalization effects on recognition performance in two ways, namely by developing a unified account which includes both verbal overshadowing and its counterpart, verbal facilitation, and by investigating these effects for events instead of for static items (e.g., faces or objects) as constituting another important class of environmental stimuli.

Inspired by practical considerations of eyewitness testimony, Schooler and Engstler-Schooler (1990) introduced the verbal overshadowing paradigm, which demonstrates verbal influences on the recognition of visual stimuli. In this paradigm, participants are first presented with a visual stimulus (e.g., a face) which they are then required to verbally describe. Typically, in a subsequent recognition test, recognition accuracy of these participants is lower than that of viewers who were not required to offer verbal descriptions of the visual stimulus (Brown & Lloyd-Jones, 2002, 2003; Meissner & Brigham, 2001; Schooler, 2002). A similar verbal overshadowing effect has also been found when participants are presented with a verbal description (*passive verbalization*) rather than having to formulate a description for themselves (*active verbalization*; Dodson, Johnson, & Schooler, 1997) after observing a visual

stimulus. In their meta-analysis, Meissner and Brigham analyzed 29 studies of the verbal overshadowing phenomenon and found a small negative effect of verbalization processes on recognition performance (Fisher's $Z_e = -0.12$). However, several studies failed to replicate the verbal overshadowing effect (e.g., Meissner, Brigham, & Kelley, 2001), indicating the fragile nature of the verbal overshadowing effect.

Two different theoretical accounts of the verbal overshadowing phenomenon have been proposed. The first of these relates to a transfer-inappropriate processing shift, and the second to recoding interference or source confusion between two concurring mental representations.

According to the processing shift hypothesis, subsequent verbalization of visual stimuli forces participants to focus on visual details of the stimulus, resulting in a feature-based cognitive style. At least in the case of face recognition, this feature-based style proves suboptimal, compared with the holistic processing style spontaneously applied when observing faces. Verbalization therefore induces an inappropriate processing shift, which in turn leads to decreased recognition accuracy in a subsequent recognition test. In line with this argument, it was shown that other tasks inducing a feature-oriented processing strategy also result in diminished recognition accuracy (for a more detailed account, see Schooler, 2002). For example, Macrae and Lewis (2002) yielded findings, in the context of a global–local task, which are comparable to verbal overshadowing effects. Whereas focusing on local features led to an impairment of visual recognition performance, focusing on global features did not.

A recoding interference has also been postulated to be responsible for the impairment of visual recognition performance following active verbalization (Schooler &

M. Huff, m.huff@iwm-kmrc.de

Engstler-Schooler, 1990). According to this theory, in verbally describing and memorizing a previously observed stimulus, participants form an additional mental representation based on their verbal description. This representation differs from the original representation, which is based on the perceived visual stimulus, by interfering with the retrieval and use of the first representation. This in turn impairs recognition accuracy, since the quality of the information available in each mental representation differs.

Dodson et al. (1997) proposed a similar account in explaining the effects of passive verbalization. They assumed that source confusion concerning which mental representation (the visual or the verbal one) should form the basis of the recognition decision is responsible for this kind of verbal overshadowing effect. Again, recognition performance is impaired because the two mental representations do not contain the same amount of detail, the visual being substantially more detailed than the verbal representation. Accordingly, Dodson et al. were able to show that, in the case of passive verbalization, there was no verbal overshadowing effect for recognition performance when participants were asked to ignore their verbal representation.

Although the impairment of visual recognition performance following verbalization processes is a robust effect (Meissner & Brigham, 2001), its generalizability can be questioned on two accounts. First, with few exceptions (e.g., Adaval & Wyer, 2004; Fiore & Schooler, 2002), research on verbal overshadowing has mainly focused on single, permanent entities, such as individual faces or cars (Dodson et al., 1997; Schooler & Engstler-Schooler, 1990), as target stimuli. However, apart from static entities (objects or faces), a second equally important class of environmental stimuli should be considered: events. In everyday communication, it is not only static entities but also witnessed events that are often subject to verbalization, both in active forms (e.g., when a person describes an accident he has witnessed) and passive forms (e.g., when a person reads about a concert after attending it the evening before).

The main difference between static entities and events is in the dynamic nature of the latter, in the sense that events undergo characteristic changes over time (Johansson, von Hofsten, & Jansson, 1980). More specifically, events typically consist of several objects that continuously change their spatial relations, because distances and orientations among the objects, as well as between objects and the observer, are not constant over time. Hence, compared with static entities, events can be considered to constitute a substantially more complex and transient type of environmental stimuli. Accordingly, it has been argued that static objects and events represent the two most important yet cognitively distinct classes of environmental stimuli (Tversky, 2004). Whether or not the verbal overshadowing effect also holds for events is, therefore, a question that arises.

Research has shown that, in contrast to static entities (e.g., faces), events are often verbalized in a global manner, with both spatial and temporal details being abstracted (Zacks & Tversky, 2001). Additionally, similar processes of detail abstraction have been reported for the observa-

tion of visual events, even in the absence of verbalizations (Huff, Schwan, & Garsoffky, 2007; Jenkins, Wald, & Pittenger, 1978; Schwan & Garsoffky, 2008). In other words, in the case of events a certain degree of abstraction can be expected for both observation- and verbalization-based representations. Throughout the rest of the paper, we will refer to this type of global mental representation as an *event model*.

The existence of event models as predominant representations of events carries some implications for the verbal overshadowing effect. Given the validity of the processing shift account, the global character of the verbalization should reduce the probability of an overshadowing effect for events, since such a global verbalization is not compatible with the feature-based processing orientation assumed to be responsible for the transfer-inappropriate processing shift. On the other hand, the recoding interference and source confusion accounts are based on the assumption of a difference which, in its traditional formulation, exists between the levels of abstraction of two competing mental representations, with the verbalization-based representation being less detailed than its observation-based counterpart (Itoh, 2005). In the case of events, both representations (event models) may well be at equal levels of abstraction, but may have different ways of abstraction, thereby placing emphasis on different features of the event. Given this difference, an overshadowing effect should also be observed for events; this can be attributed to the interference between competing event models derived from event observation and verbal event description, respectively.

The only empirical study to have dealt with this issue was carried out by Adaval and Wyer (2004). In accordance with the recoding interference assumption, they showed that when participants observed and subsequently described a social interaction, recognition performance in terms of both the statements of the target persons and their nonverbal behavior decreased. Although their experiment can be criticized on methodological grounds—they used only a single event, a 12-min excerpt from the Hollywood movie *Who's Afraid of Virginia Woolf?* (which, moreover, was not entirely visual, but also included a large number of verbal statements)—their results nevertheless provide a first indication that verbal overshadowing effects may also occur after events far more complex than static entities are described. A central goal of the present study was, therefore, to examine in a more systematic and stringent manner the existence of the verbal overshadowing effect for events.

As a further restriction, it has been shown that under certain circumstances the verbal overshadowing effect disappears, or even reverses. For example, Read (1979) showed that face recognition performance improved when verbal rehearsal processes were used. In the verbalization condition, Read observed both higher recognition accuracy and decreased response latency. Kitagami, Sato, and Yoshikawa (2002) found a verbal overshadowing effect when distractor items were designed to be highly similar to target items, whereas no verbal overshadowing effect appeared under low similarity. In an earlier study using rec-

alistic photos as stimulus material, Bartlett, Till, and Levy (1980) found that verbalization led to higher recognition performance when it enabled participants to distinguish between target and distractor items. Using a more ecological version of the verbal overshadowing paradigm, Itoh (2005) found a facilitating effect of verbalization processes on the visual recognition test; participants observed a stimulus without being instructed to memorize the target. Itoh therefore assumed visual memory of the target to be poor. Participants in the experimental condition were then asked to verbally describe the observed item. Recognition performance was significantly higher for participants in the verbalization condition than for those in the control condition.

Again, these findings are compatible with the assumption that poststimulus verbalization induces a second, less detailed representation, which is subsequently drawn upon in the recognition test. Normally, this second representation leads to recoding interference or source confusion, thereby decreasing recognition accuracy. If, however, the initial representation of the visual stimulus is itself coarse and lacking in detail (see Itoh, 2005), or if the verbal description is itself sufficient in discriminating between target and distractor items (see Bartlett et al., 1980, and Kitagami et al., 2002), the verbal overshadowing effect disappears or even reverses, and improved recognition is observed under conditions of verbalization.

A further, though empirically unaddressed, factor is the sequence in which the stimulus is observed and verbalized. Whereas active verbalization can occur only during (e.g., Brandimonte, Hitch, & Bishop, 1992) or after the visual stimulus, passive verbalization can also take place prior to its presentation. For example, in a multimedia learning environment for chemistry, a certain chemical reaction can be verbally described before or after a corresponding animation is presented to the learner. In this case, both processing-shift and source-confusion models predict an asymmetrical relationship of the two possible sequences. A transfer-inappropriate processing shift requires the verbalization to take place between observation of the event and execution of the recognition test. Verbal overshadowing is thus expected when the description follows the observation, but not when it precedes it. Similarly, in their version of the source confusion model, Adaval and Wyer (2004) describe event memory as a kind of stack in which different representations of a single referent are stored. The most recent representation is stored on top, and, consequently, is more accessible than are those representations constructed earlier (Adaval & Wyer, 2004; Wyer & Srull, 1989). Again, recognition failures are anticipated only when the less detailed representation is the last to have been formed. In sum, both theoretical accounts predict that verbal overshadowing arises only when verbalization follows visual stimulus observation.

In the case of passive verbalization of events, another potential type of ordering effect should be considered. Whereas the passive verbalization of an event already provides the listener with an abstract event model, the visual presentation of an event requires the observer to actively engage in an abstraction process with an event model as

a result. In other words, whereas the event model from the verbal description is externally given and fixed, the observation-based event model is self-generated, and can thus be adapted to the needs of the observer. This leads to an asymmetry in presentation sequence: If the visual presentation precedes the verbal description, the participant first develops an observation-based event model on his/her own, and then receives a second, externally given event model, which presumably diverges from the participant's own. Under these circumstances, verbal overshadowing due to interferences between both models can be expected. On the other hand, if the verbal description precedes the visual presentation, the participant first receives the externally given model, and then develops his own. Thus, he will be able to guide his own abstraction process in such a way that dissimilarities between the two event models are avoided.

Additionally, when presentation of the verbal event description precedes the observation of the event, it presumably influences the observation process itself, in particular by guiding attention to relevant aspects of the event (Boland, 2005; Grant & Spivey, 2003). Additionally, the literature on signaling has shown that guiding the attention of an observer by graphically indicating relevant aspects of a complex dynamic device can lead to an increase in recall and understanding (Hegarty, Kriz, & Cate, 2003). Hence, presenting verbal event description before event observation should have positive consequences for the formation of an observation-based event model. Accordingly, a verbal facilitation effect is predicted under verbalization-before-observation conditions.

Subsequently, these findings suggest that the very same verbal description of an observed event leads to either verbal overshadowing or verbal facilitation, depending on the particular circumstances—for example, the order of acquisition, the quality of the observation-based event model, and the type of difference between target and distractor items in the recognition test. Accordingly, a second main goal of the present study was to systematically investigate two of these factors—namely, sequential position and type of difference between target and distractor item—in their relationship to verbalization and recognition.

Overview of the Experiments

In sum, on the basis of current empirical findings, we assume that viewers of an event develop an abstracted event model that specifies its characteristic and relevant features. If the event is verbalized also, a second event model is formed. Furthermore, the order in which these models are acquired plays a significant role. If verbalization follows observation, the observation-based event model is left unchanged and the verbalization-based event model is prioritized for recognition. In contrast, if verbalization precedes observation, the observation-based event model is shaped according to the verbal description and is prioritized for recognition. In the following recognition test, event models are assessed according to their prioritization, in order to distinguish between target and distractor items. Distractor items that violate the prioritized event model are easily discarded as “not seen,” whereas distrac-

tor items compatible with the prioritized event model are likely to be confused with target items, leading to a decrease in recognition accuracy.

In order to test these assumptions, two experiments were conducted. In Experiment 1, viewers either saw a video clip of a short event (video-only), or they both saw a video clip *and* read a verbal description of the depicted event (video-before-verbal), or they only read the verbal description (verbal-only). All participants took part in a recognition test immediately following the presentation. The test consisted of target items (video stills from the original event) and two types of distractor item. The first type did not depict the original event, but was nevertheless compatible with the verbal description, and therefore was, presumably, in accordance with the verbalization-based event model (compatible distractors). The second type of distractor item again did not depict the original event, but at the same time violated the verbal description, and therefore was, presumably, incompatible with the verbalization-based event model (incompatible distractors). Experiment 2 was similar to Experiment 1, though the order of video clip presentation and verbal description was reversed; that is, viewers read the verbal description of the event prior to watching the video clip (verbal-before-video).

In Experiment 1, we expected that under video-before-verbal conditions, viewers would develop first an observation-based event model, then, from the verbal description, adopt a second, differing, event model, whereupon the latter would be prioritized in the recognition test by order of acquisition. In contrast to the “video-only” condition, where only an observation-based event model is formed, this condition should result in model interference, which should in turn lead to verbal overshadowing; that is, a decrease in recognition accuracy, in particular for distractor items compatible with the verbalization-based event model.

In Experiment 2, under verbal-before-video conditions, we expected that viewers would first adopt an event model from the verbal description, which would guide their attention during the subsequent observation of the event, and which would also shape the development of the observation-based event model, with dissimilarities

between both models being avoided. We also assumed that the latter would be prioritized in the recognition test, due to the order of acquisition. In this case, differences in recognition performance between video-only and verbal-before-video conditions thus have to be attributed to differences in the observation-based event models. Accordingly, assuming that the prior verbal description increases the appropriateness of the observation-based event model by means of attentional guidance, we expected a verbal facilitation—that is, an increase in recognition accuracy, especially with respect to distractor items incompatible with the observation-based event model.

EXPERIMENT 1

Method

Participants. Eighteen students (11 female, 7 male) at the University of Tübingen in Germany participated in return for payment. Their average age was 23 years.

Apparatus. Experimental procedures were controlled using a Microsoft computer and programmed using MediaLab and directRT (www.empirisoft.com). Video clips and video stills were presented on a black background in the middle of a 17-in. CRT monitor.

Stimulus Material and Design. As stimulus material, short events¹ of four balls moving on parallel courses toward a white line were created, using 3ds Max (www.autodesk.com/3dsmax/). Each ball had a different starting position and a different movement characteristic (constant velocity, accelerating, or decelerating). Each race lasted 8 sec and was rendered in such a way that the balls moved either away from or toward the observer (see Figure 1). In total, 12 video clips were designed for the main experiment. From each video clip, 3 versions were designed, 1 for every verbalization condition. No participant in this experiment viewed a video clip more than once.

The first independent variable was the verbalization, realized as a blocked within-subjects design. In the video-only condition, participants received no additional verbal information. In the video-before-verbal condition, participants received a short verbal description after watching the film. In the verbal-only condition, they received only the verbal description, which was included as a control condition, to show the efficacy of the event-model compatibility of the distractor items. The verbal description summarized the behavior of a single ball, including its starting position, its passing the other balls, and its finishing position; for example, “The red ball starts behind all other balls, overtakes every other ball and finally wins.” The verbal description was designed to establish a verbalization-based event model containing generalized information about the event.

The second independent variable was compatibility with the verbalization-based event model. It was realized by use of the dis-

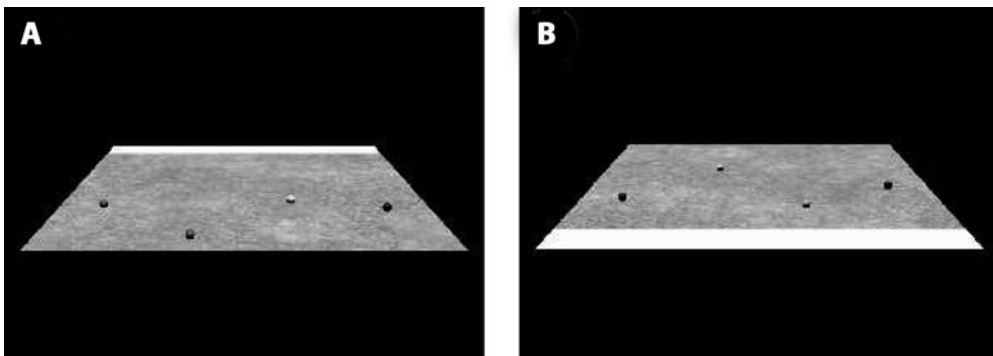


Figure 1. Stimulus material. (A) Balls moving away from the observer. (B) Balls moving toward the observer.

tractor items in the recognition test. Two types of distractor item were created: event-model compatible items and event-model incompatible items.² The event-model incompatible items violated the verbalization-based event model. In the example above, a possible event-model incompatible distractor item could consist of moving the blue ball toward the finishing line, making it impossible for the red ball to win the race. In contrast, event-model compatible items did not violate the verbalization-based event model. In the example above, this involved the yellow ball being moved. This modification is in line with the verbal description and thus with the verbalization-based event model—that is, that the red ball starts in the rearmost position, passes all the other balls, and finally wins the race. The ball that was moved in order to get the distractors was shifted 20%–25% of the length of the plane either toward the finishing line or backward from it. This manipulation was the same for both compatible and incompatible distractor items. As a third independent variable, the deviation between the viewpoint of the dynamic scene in the learning phase and that of the video stills in the testing phase was varied. Possible deviations were 0°, 45°, 90°, and 135° (see Figure 2). Finally, as a fourth independent variable, recognition items were recorded at 4 different points of time during the scene (3.0, 4.5, 6.0, and 7.5 sec after the onset of the scene). The latter two variables were introduced to measure several points of view at a number of points over the duration of the dynamic scene, on the one hand, and to collect enough data for analysis, on the other.³

Procedure. All participants were tested individually and received instructions via computer monitor. Descriptions of the kind of events used, the verbalization conditions, and the subsequent recognition test were provided. Following the instruction phase, all participants completed a training phase, in which each and every condition was presented. These data were excluded from analysis. The following experimental phase consisted of three blocks, one block for each verbalization condition (video-only vs. verbal-only vs. video-before-verbal). These blocks were presented in a balanced manner. A rest period of 5 min was established between the blocks in order to avoid exhaustion and possible carryover effects.

In the video-only condition, participants were shown an event twice. After viewing the video, participants were then shown a progress bar for 10 sec. In the video-before-verbal condition, participants watched a video twice before reading a verbal description of the event, which appeared on the monitor for 10 sec. In the verbal-only condition, participants viewed a progress bar instead of the video for 17 sec. A verbal description of the event then appeared for 10 sec.

In all conditions, a video still depicting one time point of the previously seen event was shown to indicate the commencement of the recognition test, which included 48 video stills: 16 target items showing the original event from 4 different viewpoints at 4 different points in time, 16 event-model incompatible distractor items showing a distractor event (also 4 different viewpoints and 4 different points in time), and 16 event-model compatible distractor items

(4 different viewpoint deviations and 4 different points in time). Here participants were instructed to answer “yes” by keypress if the video still depicted a constellation of the original video clip, and “no” if the video still depicted a new constellation.

Altogether a 3 (verbalization) \times 2 (event-model compatibility) \times 4 (viewpoint deviation) \times 4 (point in time) design was realized.

Results

In order to compute the signal detection parameter sensitivity A' (Pollack & Norman, 1964) and B'' (Grier, 1971), the mean hit rate (“yes” answers to target items) and the mean false alarm rate (“yes” answers to distractor items) for event-model incompatible and event-model compatible distractor items were calculated for every condition.

Sensitivity. Across all conditions and participants, a mean of $A' = .697$ was calculated. An ANOVA with repeated measurement was calculated, including the independent variables verbalization (video vs. verbal vs. video-before-verbal), compatibility with the verbalization-based event model (incompatible vs. compatible), viewpoint deviation (0° vs. 45° vs. 90° vs. 135°), and point in time (3.0 vs. 4.5 vs. 6.0 vs. 7.5 sec after onset of the scene).⁴ For an overview of the results, see Table 1.

There was a significant main effect for verbalization condition. In the video-only condition, an A' value of .782 was observed. In the video-before-verbal condition, A' amounted to .700, and to .601 in the verbal-only condition. Single Scheffé comparisons revealed significant differences between all three conditions ($p < .01$). The significant main effect of event-model compatibility revealed that incompatibility led to higher recognition performance ($A' = .779$) than compatibility ($A' = .615$). The main effect of point in time was also significant (A' for 1, .636; 2, .678; 3, .733; and 4, .743). For this effect, there was a significant linear trend indicating a recency effect [$F(1,17) = 27.975, p < .001, \eta_p^2 = .622$].⁵

The interaction between verbalization and event model compatibility was significant. Single Scheffé comparisons revealed the following significant differences ($p < .01$): Under event-model incompatibility, there was a significant difference between the video-only condition (.816) and the video-before-verbal condition (.750). For event-model compatible distractor items, significant differences

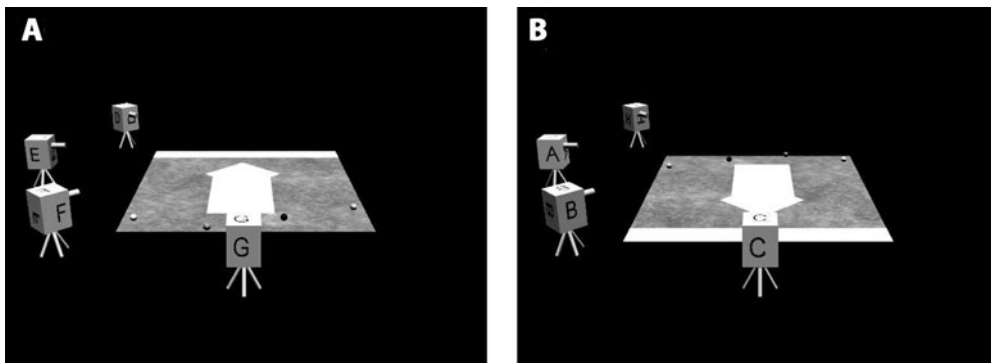


Figure 2. Stimulus material with cameras indicating the viewpoints from which the recognition items were recorded.

Table 1
Results of the ANOVA in Experiment 1, Including the Within-Subjects Factors

Main Effects or Interactions	<i>A'</i>	<i>B''</i>
Verbalization (V)	$F(2,34) = 13.59, p < .001, \eta_p^2 = .44$	$F(2,34) = 7.06, p < .01, \eta_p^2 = .29$
Event-model compatibility (EM)	$F(1,17) = 28.97, p < .001, \eta_p^2 = .63$	$F(1,17) = 23.55, p < .001, \eta_p^2 = .58$
Viewpoint deviation (VD)	$F < 1$	$F(3,51) = 3.45, p < .05, \eta_p^2 = .17$
Point in time (T)	$F(3,51) = 14.52, p < .001, \eta_p^2 = .46$	$F(3,51) = 5.11, p < .001, \eta_p^2 = .39$
V × EM	$F(2,34) = 10.37, p < .001, \eta_p^2 = .379$	$F(2,34) = 4.33, p < .05, \eta_p^2 = .203$
V × T	$F(6,102) = 4.68, p < .001, \eta_p^2 = .22$	$F(6,102) = 3.58, p < .01, \eta_p^2 = .174$
EM × VD	$F(3,51) = 3.09, p < .05, \eta_p^2 = .15$	$F(3,51) = 1.21, \eta_p^2 = .07$
EM × T	$F(3,51) = 4.71, p < .01, \eta_p^2 = .22$	$F(3,51) = 3.17, p < .05, \eta_p^2 = .16$
V × EM × T	$F(6,102) = 11.10, p < .001, \eta_p^2 = .40$	$F(6,102) = 6.07, p < .001, \eta_p^2 = .26$

were found among all conditions. Single comparisons regarding event-model compatibility across verbalization conditions revealed that event-model incompatibility led to higher recognition rates in every verbalization condition ($p < .01$; see Figure 3). A planned comparison of the predicted difference between the video-only and the video-before-verbal conditions, with respect to event-model compatibility, revealed no differences; there was no significant interaction between verbalization (video-only vs. video-before-verbal) and event-model compatibility ($F < 1$). Recognition performance in the verbal-only condition under event-model compatibility was at chance level ($A' = .5$) [$t(17) = -1.111, p = .282$]. The result pattern concerning the verbal-only condition demonstrates the efficacy of the visual distractor items regarding the concept of the compatibility with the verbal description.

The interaction between event-model compatibility and viewpoint deviation was significant. Post hoc analysis (Scheffé) revealed no differences between different view-

point deviations under event-model incompatibility or event-model compatibility. Event-model incompatibility led to superior recognition performance at 0°, 45°, and 135°, but not at 90° viewpoint deviation ($p < .01$; see Table 2).

The interaction between event-model compatibility and point in time at which the video still was recorded (3.0, 4.5, 6.0, and 7.5 sec after onset of the scene) was significant. Post hoc analysis showed the following pattern of results (see Table 3): Event-model incompatibility led to lower recognition performance at the first than at the other points in time. There were no differences between time points 2, 3, and 4. Event-model compatibility led to a significant difference between the first and the last two points in time ($p < .01$).

There was also a significant interaction between verbalization condition and point in time (see Table 3). Whereas there were no differences between points in time in the video-before-verbal condition, several differences were revealed in the verbal-only and the video-only conditions.

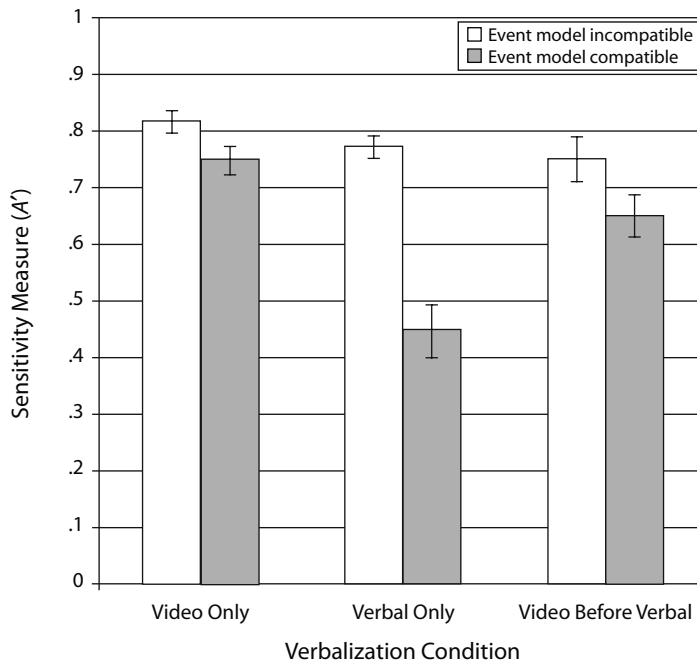


Figure 3. Interaction of verbalization and event model. Error bars indicate the standard error of the mean.

Table 2
Interaction Between Viewpoint Deviation and Event-Model Compatibility (Sensitivity Measure A')

Viewpoint Deviation	Event Model	
	Incompatible	Compatible
0°	.795	.611
45°	.772	.601
90°	.770	.636
135°	.781	.613

In the video-only condition, there were significant differences between time points 1, 3, and 4, as well as between 2 and 3. In the verbal-only condition, significant differences were found between time points 1, 3, and 4, and between 2 and 4 ($p < .01$).

There was a significant interaction between verbalization condition, event-model compatibility, and point in time. Single comparisons, according to Scheffé, revealed the following pattern of results: Event-model incompatibility did not lead to significant differences between time points in the video-before-verbal or the video-only condition. In the verbal-only condition, lower recognition performance was observed at time point 1, than at all other time points. Event-model compatibility once again did not lead to differences among time points in the video-before-verbal condition. In the video-only condition, lower recognition performance was observed at time point 1 than at time point 3. In the verbal-only condition, recognition performance at time point 2 also proved lower than at time point 4 ($p < .01$, see Table 3).

Response bias. The arithmetic mean for the response bias measure B'' was $-.280$. For statistical details, see Table 1. The significant main effect for verbalization showed that the verbal-only condition led to the most liberal reactions of the participants, whereas the video-only condition produced the most conservative reactions; the video-before-verbal condition was in between. Event-model compatibility led to more conservative reactions than did incompatibility, as the main effect for event-model compatibility showed. In line with the significant main effect for viewpoint deviation, there was a significant linear trend indicating that the greater the deviation between learning and test viewpoint the more conservative the participants. A similar effect was observed with

Table 3
Interaction Between Verbalization Condition, Event-Model Compatibility, and Point in Time (Sensitivity Measure A')

Event-Model Compatibility	Point in Time			
	1	2	3	4
Video-Only Condition				
Yes	.760	.792	.872	.840
No	.614	.748	.869	.758
Verbal-Only Condition				
Yes	.607	.781	.837	.863
No	.474	.347	.416	.553
Video-Before-Verbal Condition				
Yes	.760	.784	.739	.717
No	.599	.615	.663	.725

the factor point in time. The earlier a video still was recorded, the more liberal were the reactions of the participants. The significant interaction between verbalization and event model showed the following result pattern (see Table 4). Pairwise comparisons with the Scheffé procedure ($p < .01$) showed no differences regarding event-model incompatibility between the video-only and the video-before-verbal conditions. However, at event-model compatibility, the participants' reactions were more liberal in the video-before-verbal condition. The results in the verbal-only condition were more liberal for both event-model compatibility and incompatibility than they were in the video-only and the video-before-verbal conditions.

Discussion

The main focus of this experiment was to examine whether the verbal overshadowing phenomenon also holds for events. Recognition performance was lower in the video-before-verbal than in the video-only condition. In the control condition, in which participants only read a verbal description of the event, recognition performance was lowest. This finding suggests that reading a verbal description after viewing an event decreases recognition in a way similar to the verbal overshadowing effect found for static entities, such as a face or a car. Recognition sensitivity in the video-before-verbal condition was significantly higher for distractor items incompatible with the verbalization-based event model than it was for those that were event-model compatible. A comparable difference was also found in the video-only condition.

EXPERIMENT 2

In contrast to the first experiment, Experiment 2 reversed the order of the visual presentation and the verbal description. We expected that under the verbal-before-video condition, viewers would acquire similar verbalization-based and observation-based event models whereby the latter would be prioritized in the recognition test, due to order of acquisition. It was further assumed that the prior verbal description would direct the attention of the viewers to relevant parts of the event, thereby enhancing the appropriateness of the observation-based event model. Accordingly, a verbal facilitation (i.e., an increase in recognition accuracy) was hypothesized under the verbal-before-video condition, in particular for distractor items incompatible with the observation-based event model.

Method

Participants. Eighteen students (15 female, 3 male) of the University of Tübingen in Germany participated in return for payment. Their average age was 23 years.

Apparatus, Stimulus Material, and Design. Setting, stimuli, and design were identical to those used in Experiment 1.

Procedure. The procedure resembled that employed in Experiment 1, with the exception of the presentation order of the video clip and verbal description. The video-only condition was adapted such that, after watching the event twice (for 17 sec each time), participants were asked to begin the recognition test by pressing the space bar. In the verbal-only condition, the verbal description was presented, after which participants were requested to press the space bar. A progress

Table 4
Interaction Between Verbalization and Event-Model Compatibility (Response Bias Measure B'')

Event Model	Verbalization Condition		
	Video Only	Verbal Only	Video Before Verbal
Compatible	-.220	-.660	-.325
Incompatible	-.114	-.246	-.114

bar then appeared on screen for 17 sec. The recognition test started again when the participant pressed the space bar. In the verbal-before-video condition, participants first read a verbal description of the event, which they were subsequently shown after pressing the space bar. Immediately after the 17-sec video clip, participants were asked to begin the recognition test by pressing the space bar.

Results

As in Experiment 1, the dependent variables A' and B'' were calculated.

Sensitivity. Across all participants and conditions a mean sensitivity rate of $A' = .730$ was obtained. An ANOVA with repeated measurement was conducted using the independent variables verbalization (video-only vs. verbal-only vs. verbal-before-video), event-model compatibility (incompatible vs. compatible), viewpoint deviation (0° vs. 45° vs. 90° vs. 135°), and point in time (3.0 vs. 4.5 vs. 6.0 vs. 7.5 sec after onset of the event). For an overview of the results, see Table 5.

There was a significant main effect of verbalization. Single comparisons with the Scheffé procedure revealed significant differences among conditions. In the verbal-only condition, the sensitivity measure was .671; in the video-only condition, .744; and in the verbal-before-video condition, .775. As in Experiment 1, there was a significant main effect of event-model compatibility. Incompatibility with the verbalization-based event model led to higher sensitivity measures (.792) than event-model compatibility (.668). The main effect for point in time was significant (A' for 1, .683; 2, .710; 3, .739; 4, .789). Here, a significant linear trend indicated a recency effect [$F(1,17) = 35.172, p < .001, \eta_p^2 = .665$].

The interaction between verbalization and event-model compatibility was significant. Single Scheffé comparisons showed the following significant differences ($p < .01$): Event-model incompatibility led to higher sensitivity measures in the verbal-before-video condition than in both the video-only and verbal-only conditions. However, no difference between the video-only and verbal-before-video

conditions was observed for event-model compatibility. The verbal-only condition showed lower performances than both of the other conditions. There are no significant differences with regard to event-model compatibility in the video-only condition. However, in the verbal-only and the verbal-before-video conditions, event-model incompatibility led to higher recognition performance (see Figure 4). Recognition performance in the verbal-only condition for event-model compatibility was above chance level ($A' = .5$) [$t(17) = 3.921, p < .01$].

Response bias. The arithmetic mean across all participants and conditions was $B'' = -.285$. In sum, the results of Experiment 2 regarding the response bias measure were quite similar to those of Experiment 1. The significant main effect for verbalization indicated that there were no differences between the video-only and the video-before-verbal conditions. In contrast, in the verbal-only condition, more liberal reactions were observed (Scheffé, $p < .01$). As in Experiment 1, participants were more liberal at event-model compatibility, as indicated by the latter's significant main effect. The significant main effect for viewpoint deviation in combination with the significant linear trend for this variable [$F(1,17) = 6.201, p < .05, \eta_p^2 = .267$] indicates that the greater the viewpoint deviation, the more conservative participants' answers tend to be. The significant main effect for point in time in interplay with the significant linear contrast [$F(1,17) = 4.557, p < .05, \eta_p^2 = .211$] indicated that the earlier a test item was recorded, the more liberal were the reactions. There was a significant verbalization \times event-model compatibility interaction (see Table 6). Pairwise comparisons with the Scheffé procedure ($p < .01$) evinced the same result pattern as Experiment 1. There were no differences regarding event-model incompatibility between the video-only and the verbal-before-video conditions; however, at event-model compatibility, the participants' reactions were more liberal in the verbal-before-video condition. The results in the verbal-only condition were more liberal for both event-model compatibility and incompatibility than in the video-only and the verbal-before-video conditions.

Discussion

The main goal of this experiment was to show that the same verbal description that led to a verbal overshadowing effect in Experiment 1 would lead to a verbal facilitation effect when the order of presentation was reversed and the verbal description was presented before the video clip

Table 5
Results of the ANOVA in Experiment 2, Including the Within-Subjects Factors

Main Effects or Interactions	A'	B''
Verbalization (V)	$F(2,34) = 7.61, p < .05, \eta_p^2 = .31$	$F(2,34) = 30.41, p < .001, \eta_p^2 = .64$
Event-model compatibility (EM)	$F(1,17) = 75.26, p < .001, \eta_p^2 = .82$	$F(1,17) = 51.80, p < .001, \eta_p^2 = .75$
Viewpoint deviation (VD)	$F < 1$	$F(3,51) = 4.03, p < .05, \eta_p^2 = .19$
Point in time (T)	$F(3,51) = 19.97, p < .001, \eta_p^2 = .50$	$F(3,51) = 4.17, p < .05, \eta_p^2 = .20$
V \times EM	$F(2,34) = 5.67, p < .01, \eta_p^2 = .25$	$F(2,34) = 10.09, p < .001, \eta_p^2 = .37$
V \times T	$F(6,102) = 1.09, \eta_p^2 = .06$	$F(6,102) = 2.28, p < .05, \eta_p^2 = .12$
V \times VD	$F(6,102) = 2.66, p < .05, \eta_p^2 = .14$	$F < 1$
EM \times T	$F(3,51) = 2.84, p < .05, \eta_p^2 = .14$	$F(3,51) = 1.57, \eta_p^2 = .08$
V \times EM \times T	$F(6,102) = 6.84, p < .001, \eta_p^2 = .29$	$F(6,102) = 2.98, p < .05, \eta_p^2 = .15$

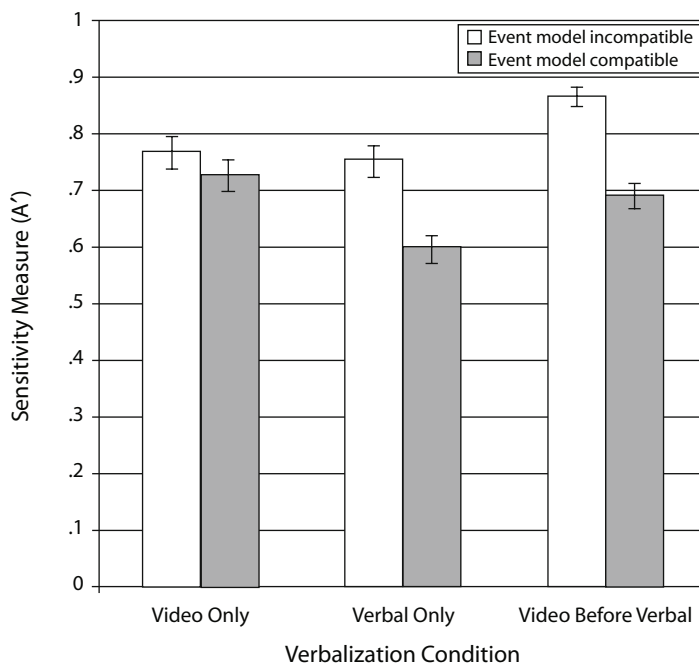


Figure 4. Interaction of verbalization and event model. Error bars indicate the standard error of the mean.

of the event was shown. The findings corroborate this hypothesis. According to the significant interaction between verbalization and event-model compatibility, viewers in the verbal-before-video conditions benefited from the verbal description and were better able to identify distractor items incompatible with the verbal event description. More specifically, in the verbal-before-video condition, these distractor items were easier to identify than were distractor items that were in accordance with the verbal summary, whereas in the video-only condition, no corresponding difference between distractor items was found. Additionally, the sensitivity rate for incompatible distractor items was significantly higher for the verbal-before-video than for the video-only condition, thus showing a verbal facilitation effect. Finally, the verbal-only condition, serving as a control condition, indicated that the improvement of recognition performance in the verbal-before-video condition could not simply be attributed to the assessment of the verbalization-based event model, since recognition accuracy was significantly lower in the verbal-only than in the verbal-before-video condition. However, in contrast to Experiment 1, recognition performance in the verbal-only condition at event-model compatibility was above chance level. A possible explanation of this finding may involve the 17-sec time delay in

Experiment 2 between encoding and test phase. Although there was no filler task, we cannot preclude visualization or recoding processes taking place, finally leading to recognition performances above chance level.

The response bias measures showed the same result pattern as did those in Experiment 1.

GENERAL DISCUSSION

In the present article, the results of two experiments that contribute to the understanding of the consequences of verbalization processes for visual stimulus recognition were presented. Previous research on this topic has focused on (1) verbalizing single entities (Brown & Lloyd-Jones, 2002, 2003; Schooler & Engstler-Schooler, 1990), and (2) rather than also considering possible facilitation effects, primarily describing those cognitive processes that were made accountable for the impairment of recognition performance (Dodson et al., 1997; Schooler, 2002). Furthermore, it becomes apparent from the literature that the type of distractor items has often been disregarded (Brown & Lloyd-Jones, 2002; Schooler & Engstler-Schooler, 1990) and that—in terms of presentation sequence—verbalization processes always occurred either during or after the visual presentation (Brandimonte et al., 1992; Dodson et al., 1997).

Hence, the present experiments extended previous research by focusing on events instead of on static entities, systematically considering the type of distractor item and order of presentation, and including an additional control condition in which the verbal event description was presented without visual event observation.

To summarize the findings: It was first shown that the verbal overshadowing effect is not restricted to single, static

Table 6
Interaction Between Verbalization and Event-Model Compatibility (Response Bias Measure B'')

Event Model	Verbalization Condition		
	Video Only	Verbal Only	Verbal Before Video
Compatible	-.107	-.672	-.348
Incompatible	-.026	-.548	-.007

entities (e.g., faces or cars), but also holds for more complex and transient events. In particular, observing an event and then reading a corresponding event description led to effects similar to those reported by Dodson et al. (1997), with static entities as stimulus material. Second, it was possible to demonstrate that the effect of an event verbalization is dependent on the order of presentation and type of distractor item. Unlike simply viewing an event, receiving a verbal description *after* observation decreased recognition accuracy regardless of the compatibility of distractor and verbal description, thus indicating a general verbal overshadowing effect. In contrast, receiving a verbal description *before* viewing the event increased recognition accuracy for distractors that were incompatible with the verbal description, thus indicating verbal facilitation; recognition accuracy for compatible distractors remained unaffected.

Third, the experiments also provided evidence that participants engage in abstraction processes even in the absence of a verbal event description, in particular under conditions of delayed recognition testing. This abstraction process is manifest in the observed difference in sensitivity between event-compatible and event-incompatible distractors in Experiment 1.

But what is the underlying cognitive account? One possible explanation of these findings is that changing the serial order of visual and verbal presentation affects participants' expectations; but we can exclude motivational influences, such as expectations, on the visual recognition performance. Apart from the fact that sensitivity measure corrects for response bias influences, the response bias as indicator for motivational effects shows the same result pattern in the two reported experiments.

In the verbal overshadowing domain, the most prominent approach in explaining negative influences on visual recognition performance is the processing shift account, which assumes that verbalizing visual memory leads to a general shift from global to local processing (Schooler, 2002). At first glance, results from Experiment 1 appear to be compatible with this account; reading a verbal summary shifted the processing strategy from a global to a local one. The resulting processing style, in turn, hinders subsequent visual recognition. Consequently, this would be initial evidence of passive verbalization (i.e., reading a verbal description) leading to a processing shift.

However, the processing shift account fails to predict results from Experiment 2. Reading the verbal summary induces a local strategy; watching the video afterward shifts the processing strategy to a more global one. The processing strategy in the video-only condition is global, too. Therefore, according to the processing shift account, no differences between those two conditions should be expected. However, at event-model incompatibility, recognition performance in the verbal-before-video condition was substantially higher than in the video-only condition.

In general, the present findings are in accordance with an account that can be deduced from the existing literature on verbal overshadowing and event cognition (Zacks, Tversky, & Iyer, 2001). The central point of this account is the assumption that, as well as encoding a concrete mental representation containing surface information, viewers of

an event develop a second mental representation, in terms of an abstracted event model that specifies its characteristic and relevant features. If the event is further verbalized, a second abstracted event model is formed. Interference processes between the two abstract mental representations (as described by the recoding interference account of verbal overshadowing) take place because there is no way of integrating them (Adaval & Wyer, 2004; Schooler & Engstler-Schooler, 1990). It is conceivable to locate source confusion, as described by Dodson et al. (1997), at this specific point. Reversing the order of visual and verbal presentation, as in the verbal before-video condition in Experiment 2, should prevent interference effects, because the abstract mental representation induced by reading the verbal summary already exists when participants are watching the video. Attentional processes are likely to guide attention during subsequent viewing of the event, leading to an observation-based mental representation shaped according to the verbal description.

One important assumption of the described account is that, even under video-only conditions, observers generate an abstract event model. Although the events in both experiments were chosen in order to allow the abstraction of many different event models, there was a chance that, in some circumstances, observers would develop an observation-based event model more or less similar to the verbal event description provided in the other experimental conditions. In this case, distractors incompatible with the verbal event description should be more easily identifiable than compatible ones, a pattern found for the video-only condition in Experiment 1.

A caveat has to be made here: Although the results seem to fit the proposed model, the latter only makes prognoses for passive verbalizations. The influences of active verbalization processes were not addressed by this account. More research is needed to investigate the influences of active verbalization on visual recognition performance of events.

From the proposed account, other consequences, which should be subject to future experimental scrutiny, can be deduced. In particular, similar overshadowing effects can be expected for pairs of purely visual event presentations, provided that each presentation accentuates different characteristics of the event (e.g., by showing the event from different viewpoints), thus inducing two differing—and, therefore, interfering—event models. In contrast, suppressing the building of an observation-based event model—for example, by concurrent articulatory suppression, as introduced by Brandimonte et al. (1992)—should reduce verbal overshadowing, because under such circumstances, the verbalization-based event model will not be accompanied by an interfering observation-based counterpart.

In sum, the present findings indicate that the recognition of complex and transient events is affected by verbalization, as is the case with stable entities. However, verbalization should not simply be equated with a performance-decreasing verbal overshadowing effect. Instead, depending on the specific conditions, a complex interplay of observation-based and verbalization-based event models, which can lead to verbal facilitation as well as to verbal overshadowing, must be assumed.

AUTHOR NOTE

M.H. and S.S. thank Maria Brandimonte, Jonathan Schooler, and Jeffrey Zacks for their helpful comments on an earlier version of the manuscript. We also thank Bärbel Garsoffky for her helpful comments and suggestions. Parts of this article were presented at the 28th Annual Conference of the Cognitive Science Society in Vancouver. It was supported by a grant from the DFG (Deutsche Forschungsgemeinschaft). Correspondence concerning this article should be addressed to M. Huff, Knowledge Media Research Center, Konrad-Adenauer-Str. 40, D-72072 Tübingen, Germany (e-mail: m.huff@iwm-kmrc.de).

REFERENCES

- ADAVAL, R., & WYER, R. S., JR. (2004). Communicating about a social interaction: Effects on memory for protagonists' statements and nonverbal behaviors. *Journal of Experimental Social Psychology*, **40**, 450-465.
- BARTLETT, J. C., TILL, R. E., & LEVY, J. C. (1980). Retrieval characteristics of complex pictures: Effects of verbal encoding. *Journal of Verbal Learning & Verbal Behavior*, **19**, 430-449.
- BOLAND, J. E. (2005). Visual arguments. *Cognition*, **95**, 237-274.
- BRANDIMONTE, M. A., HITCH, G. J., & BISHOP, D. V. (1992). Verbal recoding of visual stimuli impairs mental image transformations. *Memory & Cognition*, **20**, 449-455.
- BROWN, C., & LLOYD-JONES, T. J. (2002). Verbal overshadowing in a multiple face presentation paradigm: Effects of description instruction. *Applied Cognitive Psychology*, **16**, 873-885.
- BROWN, C., & LLOYD-JONES, T. J. (2003). Verbal overshadowing of multiple face and car recognition: Effects of within- versus across-category verbal descriptions. *Applied Cognitive Psychology*, **17**, 183-201.
- BROWN, C., & LLOYD-JONES, T. J. (2005). Verbal facilitation of face recognition. *Memory & Cognition*, **33**, 1442-1456.
- BROWN, C., & LLOYD-JONES, T. J. (2006). Beneficial effects of verbalization and visual distinctiveness on remembering and knowing faces. *Memory & Cognition*, **34**, 277-286.
- DODSON, C. S., JOHNSON, M. K., & SCHOOLER, J. W. (1997). The verbal overshadowing effect: Why descriptions impair face recognition. *Memory & Cognition*, **25**, 129-139.
- FIORE, S., & SCHOOLER, J. (2002). How did you get here from there? Verbal overshadowing of spatial mental models. *Applied Cognitive Psychology*, **16**, 897-909.
- GARSOFFKY, B., SCHWAN, S., & HESSE, F. W. (2002). Viewpoint dependency in the recognition of dynamic scenes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 1035-1050.
- GRANT, E. R., & SPIVEY, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science*, **14**, 462-466.
- GRIER, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, **75**, 424-429.
- HEGARTY, M., KRIZ, S., & CATE, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition & Instruction*, **21**, 325-360.
- HUFF, M., SCHWAN, S., & GARSOFFKY, B. (2007). The spatial representation of dynamic scenes—An integrative approach. In T. Barkowsky, M. Knauff, G. Ligozat, & D. R. Montello (Eds.), *Spatial cognition V: Reasoning, action, interaction* (Lecture Notes in Artificial Intelligence, No. 4387, pp. 140-155). Berlin: Springer.
- ITOH, Y. (2005). The facilitating effect of verbalization on the recognition memory of incidentally learned faces. *Applied Cognitive Psychology*, **19**, 421-433.
- JENKINS, J. J., WALD, J., & PITTINGER, J. B. (1978). Apprehending pictorial events: An instance of psychological cohesion. In C. W. Savage (Ed.), *Perception and cognition: Issues in the foundations of psychology* (pp. 129-163). Minneapolis: University of Minnesota Press.
- JOHANSSON, G., VON HOFSTEN, C., & JANSSON, G. (1980). Event perception. *Annual Review of Psychology*, **31**, 27-63.
- KITAGAMI, S., SATO, W., & YOSHIKAWA, S. (2002). The influence of test-set similarity in verbal overshadowing. *Applied Cognitive Psychology*, **16**, 963-972.
- LOFTUS, E. F., MILLER, D. G., & BURNS, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 19-31.
- MACRAE, C. N., & LEWIS, H. L. (2002). Do I know you? Processing orientation and face recognition. *Psychological Science*, **13**, 194-196.
- MEISSNER, C. A., & BRIGHAM, J. C. (2001). A meta-analysis of the verbal overshadowing effect in face identification. *Applied Cognitive Psychology*, **15**, 603-616.
- MEISSNER, C. A., BRIGHAM, J. C., & KELLEY, C. M. (2001). The influence of retrieval processes in verbal overshadowing. *Memory & Cognition*, **29**, 176-186.
- POLLACK, L., & NORMAN, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science*, **1**, 125-126.
- READ, J. D. (1979). Rehearsal and recognition of human faces. *American Journal of Psychology*, **92**, 71-85.
- SCHOOLER, J. W. (2002). Verbalization produces a transfer inappropriate processing shift. *Applied Cognitive Psychology*, **16**, 989-997.
- SCHOOLER, J. W., & ENGSTLER-SCHOOLER, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, **22**, 36-71.
- SCHWAN, S., & GARSOFFKY, B. (2008). Event segmentation and memory. In T. F. Shipley & J. M. Zacks (Eds.), *Understanding events: How humans see, represent, and act on events*. New York: Oxford University Press.
- TABACHNICK, B. G., & FIDELL, L. S. (1989). *Using multivariate statistics*. New York: Harper Collins.
- TVERSKY, B. (2004). Narratives of space, time, and life. *Mind & Language*, **19**, 380-392.
- WYER, R. S., JR., & SRULL, T. K. (1989). *Memory and cognition in its social context*. Hillsdale, NJ: Erlbaum.
- ZACKS, J. M., & TVERSKY, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, **127**, 3-21.
- ZACKS, J. M., TVERSKY, B., & IYER, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, **130**, 29-58.

NOTES

1. Examples of the stimulus materials are available at www.iwm-kmrc.de/cybermedia/voe/.

2. Although the concept of "compatibility" introduced in this paper looks quite similar to the notion of "consistency" used by the literature on misleading information, it refers to another mechanism. According to Loftus et al. (1978), inconsistent items were consistent with the postvisual verbal information but inconsistent with the visual information. Our concept of "compatibility" refers to the similarity of verbal information and type distractor items.

3. Since the results regarding the viewpoint deviation did not show any significant influences on the visual recognition performance, we did not discuss these influences in this article. The failure of replicating the viewpoint deviation effect as reported by Garsoffky, Schwan, and Hesse (2002) can be due to design issues (e.g., carryover effects).

4. We calculated additional analyses with the serial order of the blocks as between-subjects factor. Since this factor did not show an influence on the results of the reported within-subjects ANOVA, we report this calculation.

5. In the reported experiments the partial η^2 is reported as effect size measure, since it is more appropriate for designs with more than one independent variable (Tabachnick & Fidell, 1989).