Perception & Psychophysics
2003, 65 (5), 789-800

Audiovisual phenomenal causality

RAINER GUSKI and NIKOLAUS F. TROJE
Ruhr-Universität Bochum, Bochum, Germany

We report three experiments in which visual or audiovisual displays depicted a surface (target) set into motion shortly after one or more events occurred. A visual motion was used as an initial event, followed directly either by the target motion or by one of three marker events: a collision sound, a *blink* of the target stimulus, or the blink together with the sound. The delay between the initial event and the onset of the target motion was varied systematically. The subjects had to rate the degree of perceived causality between these events. The results of the first experiment showed a systematic decline of causality judgments with an increasing time delay. Causality judgments increased when additional auditory or visual information marked the onset of the target motion. Visual blinks of the target and auditory *clacks* produced similar causality judgments. The second experiment tested several models of audiovisual causal processing by varying the position of the sound within the visual delay period. No systematic effect of the sound position occurred. The third experiment showed a subjective shortening of delays *filled* by a clack sound, as compared with *unfilled* delays. However, this shortening cannot fully explain the increased tolerance for delays containing the clack sound. Taken together, the results are consistent with the interpretation that the main source of the causality judgments in our experiments is the impression of a plausible unitary event and that perfect synchrony is not necessary in this case.

Imagine an object moving horizontally toward a second object and then stopping abruptly right in front of it. At this moment, the target starts moving, while the first object remains motionless. When observing such a sequence on a computer screen, we tend to say that the launching of the second object was caused by the first one. This effect is called *phenomenal causality* (Michotte, 1946/1963) or *perceived causality* (Piaget, 1967) and has often been replicated, as has been reviewed by Scholl and Tremoulet (2000). Michotte stressed the systematic relationship between causality judgments and the spatial/temporal parameters of the displays. The main variable contributing to causality judgments in many fields is the temporal contiguity between the observed events (Oestermeier & Hesse, 2000; White, 1995). Temporal contiguity does not necessarily mean exact co-occurrence of the cause and its effect: For instance, Michotte observed "best" causality impressions in *launching* displays with delays of about 30–40 msec between the end of the first motion and the start of the target motion. He also claimed that "direct launching is only reported constantly for intervals up to 50 milliseconds, though it may still be reported occasionally up to 100 milliseconds" (Michotte & Thinès, 1963/1991, p. 70). Several experiments have shown that the delay may rise up to 200 msec and still produce reliable impressions of causality. Gruber, Fink, and Damm (1957) reported a "median temporal threshold of causality" of 95–98 msec in their pretests and of 150–200 msec in a series of delay trials (0–250 msec). Scholl and Nakayama (2000) reported that the "causal capture effect"—that is, facilitating the causal interpretation of an ambiguous streaming display by means of an additional launching display—is confined to a 0- to 100-msec overlap between the two displays. Larger delays result in significant decreases of judged causality.

In most experiments in this area, unimodal visual displays have been used. To our knowledge, no publication reporting the use of unimodal auditory displays has appeared in print until now, but Michotte (1946/1963, pp. 235–241) mentioned three audiovisual experiments—that is, experiments in which sounds were linked to visual motions. In all three experiments (Experiments 80–82), the sound was generated by an electrically controlled hammer enclosed in a wooden box and occurred 20 msec after one of the following visual events: For instance, in Experiment 80, a circle appeared, disappeared, changed its brightness, moved and then stopped abruptly, or moved and passed over another object, which was stationary. To Michotte's great surprise, "it was only in exceptional cases that observers used expressions implying that one event had an influence on the other" (p. 236). A few of the reported comments of the subjects point to the probability that some subjects judged the acoustic event to cause the visual events—a causal relation that Michotte seems to have "counted as negative" (p. 237). Experiment 81 used 6 subjects, with the classic launching display containing two circles and the sound occurring at the time of the contact between the circles. In some of the trials, rebounding of the first disk took place. "The

Correspondence concerning this article should be addressed to R. Guski, Department of Psychology, Ruhr-Universität Bochum, D-44780 Bochum, Germany (e-mail: rainer.guski@ruhr-uni-bochum.de).

 Copyright 2003 Psychonomic Society, Inc.

results showed considerable variation. In some cases the events seemed independent, while in others the impact seemed to 'produce' the noise. For present purposes, however, the most important point is that for more than one observer the 'production' of the noise was much more clearly marked *when rebounding took place*" (p. 237f.). Experiment 82 was identical to Experiment 81, except that the second disk moved at a considerably slower speed after the collision. "Twenty-one people . . . took part in this experiment; twelve of them, or 57%, stated, when we asked them specifically, that the impact produced the noise" (p. 238). Michotte seemed to be troubled by the fact that (1) virtually no subject spontaneously mentioned the sound and (2) the sound was not synchronous with the contact between the experimental objects. From both aspects, he concluded that the "impression was not really a causal one" (p. 240), and he stopped experimenting on audiovisual displays, mainly because of technical problems.

In the light of more recent data, Michotte (1946/1963) probably drew a false conclusion, because audiovisual events are judged to be most synchronous when the sound occurs 50 msec *after* the visual event connected with the sound (Kohlrausch & van de Par, 2000; Rudloff, 1997). In addition, he seemed to believe that the audiovisual system is less tolerant with small delays between the end of the initial movement and the start of the auditory event than the visual system is with small delays between the stop of the initial movement and the start of the effect movement. This is strange, because it is generally known that sound waves travel at rather low velocities and that echoes are considerably behind their cause (e.g., hand claps or shouts).

There are some historic papers that have been concerned with auditory influences on motion perception. We will mention them here briefly, because their line of reasoning is different from that of Michotte (1946/1963). For instance, Zietz and Werner (1928) used pairs of drawings (e.g., a drawing of a bent arrow and a drawing of a small circle) and presented them repeatedly one after the other, with a delay of several hundred milliseconds. Without sound, the majority of subjects seemed to perceive the two drawings in isolation, but when the presentation of each drawing was accompanied by a collision sound, the subjects tended to see the visual objects in motion—jumping, expanding, or rotating in space. Maass (1938) extended the experimental approach of Zietz and Werner to a visual display consisting of 12 tiny electric bulbs, which could be programmed in order to represent rapidly changing geometric forms, and three electric hammers, which also could be programmed and produced 64 different rhythms. Without sound, the optic displays produced a huge variety of movement descriptions, and the author did not see a specific relation between the visual stimuli and the movement descriptions of the subjects. When the optic displays were connected with sounds, the movement descriptions the subjects produced were less variable, and the light points were experienced as belonging more closely together.

There are some recent papers that did not describe audiovisual causality experiments but were essentially related to them.

1. Sekuler, Sekuler, and Lau (1997) reported that an ambiguous visual display is perceived as a collision if a collision sound is added. When two surfaces move toward each other in a straight line, coincide, and then move apart, this display is consistent with two different interpretations: After coincidence, either the two objects could have continued to move in their original directions, or they could have collided and then bounced, reversing directions. Adding (acoustic) clicks at the time of the visual coincidence increased the perception of collision significantly. Watanabe (2000) reported that the bounce-inducing effect of 3-msec low-intensity clicks worked from 250 msec before to 150 msec after the visual coincidence and within a separation of up to 30° of visual angle between the sound and the visual display. Watanabe and Shimojo (2001b) found that the bounce-inducing effect of very brief clicks was attenuated when other identical sounds (auditory flankers) were presented before and after the simultaneous sound. The attenuation occurred only when the simultaneous sound and auditory flankers had similar acoustic characteristics and the simultaneous sound was not salient. The authors conclude that (1) the bounce-inducing effect of the sound is a genuine audiovisual effect and (2) there exists a saliency-assigning process that is context sensitive and can be utilized by the visual system for resolving ambiguity. In a second series of experiments on the streaming/bouncing display using a sound, a flash, or a pause after the coincidence of the two surfaces, Watanabe and Shimojo (2001a) found that each of the additional events increased the *bouncing* judgment but that "the percentage of the bouncing percept did not reach its maximal level immediately after the coincidence but increased as a function of postcoincidence trajectory duration up to 150–200 msec" (p. 16). The authors hypothesized that the perceptual systems need a certain amount of time to integrate the (more or less conflicting) information into a unitary concept.

2. Experiments on the perception of audiovisual synchrony (Dixon & Spitz, 1980; Kohlrausch & van de Par, 2000; Lewkowicz, 1996; Rudloff, 1997; van de Par & Kohlrausch, 2000) have shown that a sound may be judged to be synchronous with a visual event even when the sound occurs after a short delay. In contrast, a preceding sound is less likely to be judged as synchronous. The tolerance for synchronism depends on the content of the audiovisual events (e.g., speech or collisions between objects) and lies in the area between ca. −130 msec (sound preceding) and +250 msec (sound following). Optimal synchronism between sound and image is perceived if the sound follows 80–100 msec after the visual event (cf. Rudloff, 1997; van de Par & Kohlrausch, 2000).

3. Lewald, Ehrenstein, and Guski (2001) varied the degree of synchronism between moving visual and stationary auditory pulse stimuli in an auditory localization task. They reported that the auditory localization of pulse stimuli with a 2-Hz repetition rate is most strongly influenced by a visual stimulus that is ±50 msec synchronous with an auditory stimulus.

Taken together, the results that have been reported show that auditory and visual events seem to be linked if they occur within a time frame of less than 200 msec. White (1995, 1999) discussed a similar time frame under the heading of iconic processing—that is, he proposed that most visual information has been lost after 250–300 msec and that iconic processing operates "on material integrated over a span of not more than 250 msec" (White, 1995, p. 40). Following this line of reasoning, visual events will not be causally linked together when they are more than 250 msec apart. Crowder (1993) also reported 250 msec for echoic processing—without discussing phenomenal causality. When audiovisual events are considered, it is an open question whether the causality time span is (1) shorter than the shorter of both the echoic and the iconic memory stores, (2) the shorter of the two, or (3) longer than the longer of the two. The empirical evidence presented so far suggests that visual and auditory information interact at least during the time frame of about 100–200 msec in a way that would be advantageous for audiovisual phenomenal causality, and it seems to be even more favorable when the visual event precedes the auditory event.

Several questions have been open until now, especially with respect to the role of auditory information in visual events associated with phenomenal causality. For instance, we do not know whether sudden auditory events, timed in close contiguity with the "collision" of the two visual objects, will enhance the causal impression or not. Furthermore, if they enhance the causal impression, will they simply increase the degree of judged causality, or will they expand the tolerated delay between the two visual motions? If auditory events enhance the impression of causality in displays containing a delay between the two motions, what is the optimal point in time for doing so? Would this enhancement be due to the distraction of attention from the visual display or to some other mechanism? Earlier experiments have shown that delays *filled* with events often are judged to be shorter than *unfilled* delays of the same duration (for a review, see Ihle & Wilsoncroft, 1983). This may also happen with delays in launching displays.

The experiments to be described here relate mainly to the question of whether the phenomenal causality of visual *launching* displays can be altered by means of additional impulsive sounds or additional visual impulsive events. The expected increase in judged causality could be due simply to the temporal contiguity of visual and auditory events (a temporal cross-modal information link). Alternatively, it could be due to associations with real-life collisions, because real collisions between two objects usually cause impulsive sounds (a "reality" link). A third hypothesis relates to the *assumption of unity* in the sense proposed by Welch and Warren (1980): The discrepancy of cross-modal information may be partially ignored when structural factors (e.g., amount of stimulus discrepancy) and cognitive factors (e.g., instructions) support the impression of a common cause of the cross-modal information. There is a fourth hypothesis: The enhancement of causality impressions is due neither to cross-modal nor to "realistic" information but is caused simply by an increase of the total available information about sudden changes in a stream of events. Therefore, the effect of a sudden color change—which is not a typical effect of real collisions—instead of the sound was tested in Experiment 1.

In Experiment 1, we also varied the time delay between the end of the first movement and the start of the second one (delay). In Experiment 2, we varied the time delay between the end of the first visual movement and the start of the auditory event. In Experiment 3, we used a two-alternative forced choice (2AFC) time comparison task and tested the hypothesis that delays *filled* by a clack sound are subjectively shorter, as compared with *unfilled* delays. (See Figure 1 for a schematic display.)

Causality judgments were recorded by means of continuous probability ratings—that is, subjects rated the probability that the target motion was caused by a preceding perceptual event. We did not try to distinguish between *judged* and *perceived* causality, as proposed by Schlottmann and Shanks (1992), because we were convinced that the second and each of the following reactions of subjects in a repeated measures design would not reflect immediate impressions any more. Rather, they must reflect judgmental processes—for example, comparisons between experimental stimuli (cf. Parducci, 1974; Sarris, 2000) and, in our case, additional comparisons between experimental stimuli and memorized collision events.

Generally, we expected causality judgments of visual collision events to increase if synchronous collision sounds or visual events were present. This increase could be either

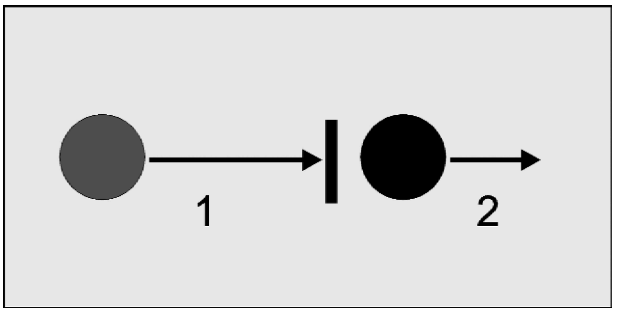

Figure 1. Schematic display of the motions in the animations of the *launching* type (prestudy and Experiments 1–3). The background color was green, the left disk was yellow (prestudy) or blue (Experiments 1–3) and the right disk was red. Arrows, the vertical bar, and numerals were not displayed in the original.

additive or multiplicative. In the first case, the increase in judged causality would be constant and, therefore, independent of other experimental variables. In the second case, the increase induced by the additional marker event would depend on the judged causality without the marker event. If judged causality was already high (e.g., because the delay was short), adding a marker event would have less effect than it would in a case in which the initial causality judgment was low. We further expected that causality judgments would be greater when the collision sound occurred in the middle of the visual delay, as compared with the margins of the delay, and that the increase in causality judgments with additional sound would be due partly to the subjective shortening of time delays *filled* with sound, as compared with *unfilled* delays.

Prestudy

In a classroom experiment, a series of animations was presented by means of an LCD projector, most of them depicting the classic Michotte (1946/1963) launching display. (Several other displays were used in the same study, but they will not be reported here.) The two disks moved with a constant speed of 16 cm/sec, both before and after the "collision." The delay between the stop of the first movement and the start of the target movement varied systematically between 0 and 400 msec, in steps of 80 msec. A small gap separated the disks even at the closest position. (It had been shown earlier that a small gap does not decrease causality judgments; see, e.g., Michotte, 1946/1963, p. 99; Yela, 1952.) Half of the trials contained collision sounds: 10-msec *clack* sounds in the middle of the visual delay, recorded while a wood block was hit with a drumstick (peak frequency = 880 Hz) and emitted by means of active loudspeakers in front of the projection screen.

A total number of 30 subjects (students) were naive with respect to the experimental questions. They were shown several examples of the movies and received separate answer sheets for each video clip, each containing the same question: "How probable is it that the movement of the red object (disk or ball) is caused by a perceivable event immediately before?" This question leaves the kind of causal event (e.g., sound or motion) open to the subjects. The numeric answer scale comprised 11 steps from 0 to 10, and both of the extremes were marked: *not at all* (0) and *very probable* (10).

Results. As was expected, the median scores of causality judgments decreased with an increasing time delay between the two movements in both conditions (Figure 2). The median scores in the sound condition (aggregated over all delays) were significantly above those without sound (Wilcoxon $Z = 4.192$, $p < .001$). The medians at all delays above 0 msec in the sound condition lay within the 95% confidence interval of the no-sound condition, if the sound effect was subtracted. Considering half of the full range of the response scale (0 . . . 10 = 5) to be the critical value for causality judgments, we computed the tolerated visual delay as the point at which a quadratic regression of the medians intersected with the critical causality value. As a result, 180 msec were tolerated in the condition without sound, whereas 340 msec could pass in the condition with sound, without a decrease below the critical value.

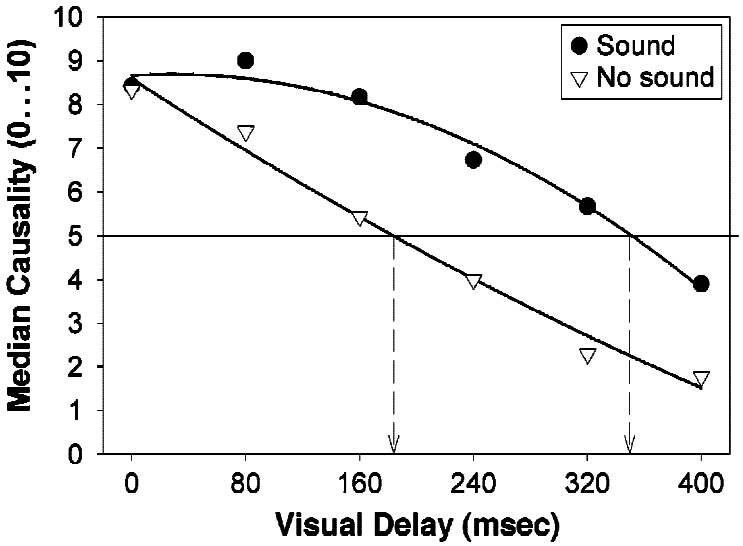

Figure 2. Median causality judgments with Michotte launching displays (prestudy). The term *visual delay* denotes the interval between the stop of the first disk and the start of the target motion. The solid lines show the quadratic regressions of the medians.

EXPERIMENT 1

Parts of Experiment 1 served as a reliability test of the prestudy, with a refined technique. In addition, we asked whether the observed increase of causality judgments in the sound conditions of the prestudy was due to a specific acoustic event that might easily be attributed to a collision or whether it was due more generally to an unspecific sudden event. Therefore, we used a *blink* (i.e., a short color change of the target from red to white and back to red) in one of the experimental conditions.

There were three reasons for using *blinks* as an alternative to *clacks*. First, it was nonacoustic information with a similar impulse character, like clacks, and second, it would probably be less related to preexperimental knowledge of "real" collisions than clack sounds would be. Whereas clack sounds, especially hitting wood, might easily be associated with collisions of billiard balls, a sudden flash of one of the balls would not be expected. Third, Watanabe and Shimojo (1998, Experiment 1) used *blinks* in the middle of the streaming/bouncing bistable display (similar to that of Sekuler et al., 1997) and found enhanced *bouncing* impressions, as compared with control situations (no *blinks*) and nonsynchronous *blinks*. In contrast, Schlottmann and Shanks (1992, p. 321) reported that "subjects' ratings of the degree of perceived causality were not affected by the color change."

Method

Subjects. Ten undergraduate students of psychology (6 women and 4 men; mean age = 22.8 years) served as subjects for part of their course credit. All were naive as to the purposes and hypotheses motivating the study. They reported normal or corrected-to-normal vision and hearing.

Stimuli. The classic Michotte (1946/1963) *launching* display was used (as in the prestudy). The visual stimuli were 26 mm in diameter (1.89º of visual angle at viewing distance), their velocity was 180 mm/sec (13.2 deg/sec, both before and after the "collision"), and the gap size was 2.5 mm (0.176º). The visual stimuli (blue and red disks) were displayed on a 21-in. Silicon Graphics screen. The delay between the stop of the first disk's movement and the start of the second disk's movement was varied between 0 and 400 msec in nine steps. Right in the middle of this delay, an additional marker event could occur: the same 10-msec clack sound as in the prestudy (sound condition); a blink event, during which the (red) target disk changed its color to white for 50 msec before returning to the red color (blink condition); or a combined sound plus blink event with both the visual and the acoustic markers occurring at the same time (sound + blink condition). The acoustic background level in the laboratory was 45 dB(A). The auditory stimuli were presented monophonically via headphones (Sennheiser HD 450 II). Their peak level was 75 dB(A), measured by means of an artificial ear (Brüel & Kjaer 4152; position: "fast") and the measuring amplifier (Brüel & Kjaer 2609).

Procedure and Design. In contrast to the prestudy, Experiment 1 was run with individual subjects, using a Silicon Graphics workstation. The stimuli were computed and displayed in real time with a precision of about 1 msec. The distance from the subjects' eyes to the screen was 76 cm. The head was supported by a chinrest. The instructions were the same as those in the prestudy, but a 9-point response scale (1 to 9) was used, and the subjects reported the causality ratings verbally. In four experimental conditions, the kind of additional events was varied. In the control condition, no additional event was used. In the sound condition, a 10-msec clack sound was presented. In the blink condition, the target disk changed its color for 50 msec. In the sound + blink condition, the clack sound was presented together with the blink of the target disk. The additional sound/blink occurred exactly in the middle of the delay between the stop of the first disk's motion and the beginning of the target disk's motion. The delay varied in nine steps of 50 msec between 0 and 400 msec.

At the beginning of the experiment, the subjects were presented with 12 demonstration trials, 3 from each condition (containing minimum, maximum, and intermediate delays, respectively). After the verbal instructions, the order of the trials was randomized individually for each subject. Each condition and each delay was measured three times, resulting in a total of 108 trials (4 conditions × 9 delays × 3 repetitions). The repetitions were evenly distributed within the four quarters of the total individual session. At the end of the session, each subject answered several questions about his/her experiences during the experiment. The questions we consider relevant to be discussed here were (1) whether the subjects had the impression that the disks were "real" objects and (2) whether they thought that the motions of the two disks belonged to the same event.

Results

As in the prestudy, the median causality ratings declined with an increasing delay between the motions of the disks (Figure 3), and the lowest causality ratings were obtained in the control condition—that is, without any additional (sound or blink) information. Because of many skewed frequency distributions of the causality judgments, we decided to perform nonparametric statistical analyses. The Friedman test with causality ratings, aggregated for all the delays tested, showed a significant effect of the experimental conditions ($\chi^2 = 23.4$; $N = 10$; $df = 3$; $p < .001$), with that of the sound + blink condition being highest and that of the control condition lowest. Separate post hoc analyses showed the sound and the blink conditions not to be significantly different from each other with respect to causality ratings (Wilcoxon $Z = 0.65$, $p = .5$), whereas their respective differences from the control condition were significant (Wilcoxon Z of about 2.75, $p < .01$). The differences between aggregated causality ratings in the sound + blink condition and the respective sound and blink conditions were statistically significant (Wilcoxon Z for sound vs. sound + blink = 2.67, $p < .01$; Z for blink vs. sound + blink = 1.99, $p < .05$).

The regression lines of the median causality ratings over the 9 delays showed significant linear trends mainly ($r^2 = .95$–$.97$), except for the control condition, which showed a marked quadratic trend ($r^2 = .99$). Unlike the prestudy, all the noncontrol conditions showed somewhat shallower slopes—that is, the majority of the subjects kept their causality judgments above the midpoint of the scale (1 . . . 9 = 5) for all conditions with additional events up to delays of 280 msec. In the sound + blink condition, they even tolerated 390 msec! The general effect was independent of the particular choice of the criterion: A more strict criterion for causality would use only median ratings exceeding two thirds of the response scale (i.e., 7). With this criterion, our subjects did not tol-

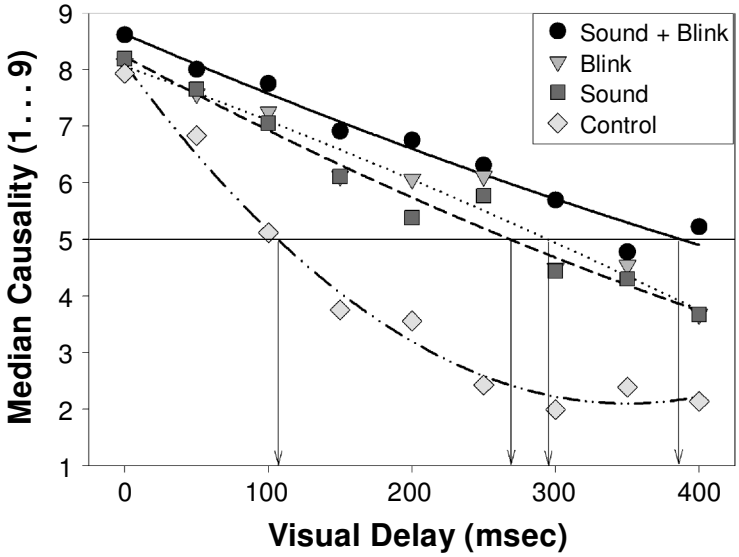

Figure 3. Median causality judgments in Experiment 1. The *x*-axis shows the delay between the stop of the first disk and the onset of the target motion.

erate a 50-msec delay in the control condition, but they tolerated up to a 100-msec delay in the sound/blink conditions and 150 msec in the sound + blink condition.

We also tested for learning effects during this experiment. For analytic purposes, we divided the whole session into four quarters and compared the median causality ratings between them. It turned out that neither the overall test nor specific tests between each of the possible pairs produced any significant result (Wilcoxon Z between 0.1 and 0.77; p between .9 and .4, respectively). In addition, we hypothesized that causality judgments would be greater for trials that occurred after control condition trials, as compared with trials occurring before. We tested this by using pairs of causality judgments from noncontrol-type trials before control-type trials against judgments from the same type of trials after control-type trials. Contrary to the hypothesis, the median causality judgments seemed to be somewhat lower after control-type trials than before, but statistical tests did not show conclusive results (Friedman $\chi^2 = 3.6$, $p = .06$; Wilcoxon $Z = 2.2$, $p < .05$).

The postexperimental interview produced two useful results. Eight of the 10 subjects said, in response to Question 1 (real objects or not), that they perceived "(billiard) balls, coins, or bowls." Only 1 person spontaneously mentioned the blink: She was reminded of an explosion. Answering Question 2 (inquiring whether one or two events were perceived), 7 of the subjects explained that they sometimes perceived two events. Elaborating on this topic, 4 of these 7 people mentioned either the delay between the two motions or "two independent motions," and the other 3 mentioned the asynchrony between sound and motion. We will come back to this issue.

Discussion

The results of Experiment 1 and the prestudy both showed a clear continuous decrease of causality judgments with an increasing delay between the two visual movements involved. Furthermore, they both showed that an additional perceivable event in the middle of the visual delay enhanced the impression of causality with Michotte-type launching displays. Considering the continuous monotonic decrease of the causality ratings with increasing delays in each of the experimental conditions of Experiment 1 and the prestudy, we tend to believe that with respect to experimental launching events, there is no clear-cut temporal *causality time span*. Instead, it seems that different aspects of the stimulus information (e.g., the delay, the additional impulsive sound event, the additional impulsive visual event, the size of the disks, the relative speed of the disks, and directions) all contribute more or less to a continuous causality judgment about artificial stimuli that remind one of collisions (cf. Anderson, 1990, pp. 161f.).

Experiment 1 also tested the hypothesis that the increased tolerance for delays in the case of audiovisual displays is a genuine audiovisual phenomenon. We found that a visual event has the same effect as an auditory event. Sekuler et al. (1997) suggested that the increase of *bouncing* interpretations in an ambiguous visual display—which is seen mostly as two motions crossing each other—is due to "an acoustic event that signals a collision between moving objects" (p. 308). In the same paper, the authors reported that a visual change—a pause in the visual motion at the moment of crossing—enhances the bouncing impression. In the case of launching events, the majority of our subjects agreed with the causal interpretation,

as long as the delay between the two motions was short. The addition of a sudden sound or a sudden blink during the delay helped to increase the causal interpretation, and it turned out that there was virtually no difference in causality judgments between displays using a brief clack sound and displays using a brief blink of the target. This demonstrates that the increase in the duration of delays tolerated is no genuine cross-modal effect; it may be due merely to attention distraction or to the amount of information connected with launching or collision situations.

In summary, the tolerated delays can be stretched, if a sudden impulsive event is inserted in the middle of the delay between the two visual events: At a 340-msec visual delay, the subjects still consider the events to be causally related. Although this seems to be a clear effect, we do not know the reasons for its occurrence. Three different lines of thought may apply.

1. Watanabe and Shimojo (1998) suggested that the increase of *bouncing* judgments in their streaming/bouncing display with the addition of a blink at the time and location of the overlap was due to attention distraction. This interpretation may apply to our results, too. In Experiment 2, we tested some of their assumptions.

2. It is uncertain which events are causally related in the audiovisual condition. Most of the stimuli used in the previous experiment consisted of three parts: the movement of the first disk, the marker event, and the movement of the second disk. It is uncertain which of these events were causally related. It is conceivable that the subjects did not perceive a unitary collision event but, rather, one of the three possible dyadic cause–effect relations: (1) the launching of the target was caused by movement (and stopping) of the first disk; (2) the launching of the target was caused by the marker event; and (3) the marker event was caused by the movement of the first disk (cf. Figure 4). The reported distribution of answers to Question 2 reflects two different aspects: (1) the range of stimuli used during the whole session contained several scenes that did not produce a clear impression of causality, and (2) in some cases, the subjects may not have experienced a direct causal link between the first motion, the marker, and the second motion. Experiment 2 explicitly accounted for this question.

3. It is well known that information presented during an interval may decrease the judged duration of the interval (e.g., Ihle & Wilsoncroft, 1983). Therefore, a clack or a blink presented during the delay between the two motions in the launching display may decrease the functional delay and, therefore, result in increased causality judgments. Experiment 3 explored this interpretation.

Taken together, the results of Experiment 1 showed a continuous decrease of causality ratings with increasing delays between cause and effect. The majority of the subjects rejected the notion of causality with visual launching displays using delays of more than 100–160 msec in the control condition. The decrease of causality ratings with an increasing delay could be partially compensated for by additional (auditory or visual) information: Even delays of more than 300 msec were tolerated if sounds or blinks were added to the classic visual launching information. Experiment 1 did not analyze the specific contributions of specific components of the total information to the global causality judgments. This was the aim of the following experiments.

EXPERIMENT 2

In the prestudy and Experiment 1, the marker event (clack sound and/or color change) always occurred in the middle of the visual delay. We assumed that this condition provided a maximum temporal link between the stop of the first movement and the start of the second one, at least in the range of delays used (up to 400 msec). In Experiment 2, we asked whether the position of the marker event within the delay period has a significant influence on the causality ratings. Since Experiment 1 showed that a visual event has the same effect as an acoustic event, we used only the latter in Experiment 2.

In the case of a unitary collision event, we would expect no effect of the exact timing of the marker event (see Figure 4). The three possible dyadic cause–effect relations, however, should all be influenced by changing the exact time of the marker event within the delay period. If the predominant cause–effect relation is the one between the stop of the first movement and the onset of the second movement (*a* in Figure 4), we hypothesize that a maximal temporal link between the two events is achieved if the marker event occurs right in the middle of the delay period. This hypothesis is based on the assumption that the causal impression decreases with an increasing delay, the decrease being almost complete at about 200 msec. The causality impression can be partially *refreshed* by means of additional information connected to collisions. After refreshing the impression, the decrease starts again.

If the important cause–effect relation is the one between the marker event and the launching of the second disk (*b* in Figure 4), causality judgments should increase the closer these two events are connected in time. We should expect low causality ratings if the marker event occurs at the beginning of the delay period and high ratings if it occurs at the end of the delay period. In contrast, if the main relation is the one between the movement of the first disk and the marker event (*c* in Figure 4), a high causality rating would be expected if the marker event occurs early and, therefore, temporally close to the stopping of the first disk, and a low rating would be expected if it occurs late.

There is another line of reasoning, quite different from the one just discussed: Extrapolating the results of Watanabe and Shimojo (1998), it can be hypothesized that sound attracts attention and that the momentary distraction hinders the complete pickup of visual information, thus leading to a greater contribution of inference processes at the cost of accurate perception. The effect

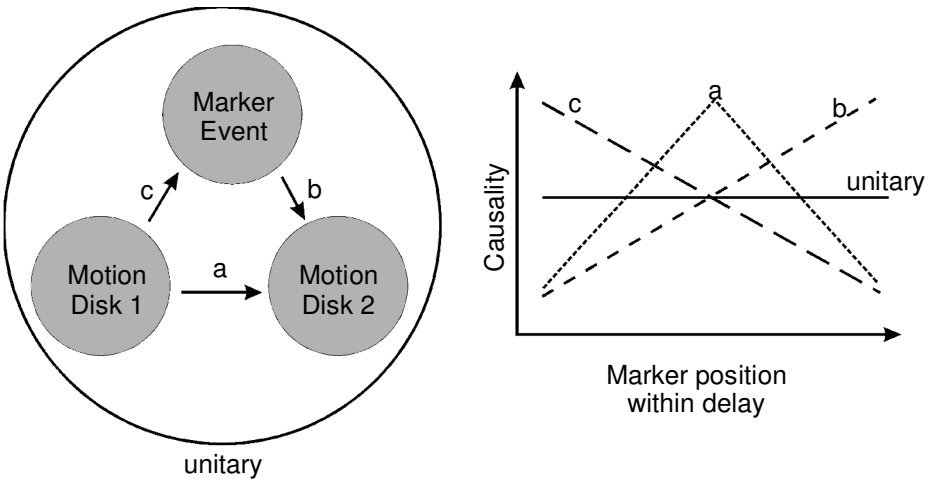

Figure 4. Models of unitary and dyadic cause–effect relations (left, a–c), together with the expected effects of the marker positions on causality ratings (right).

should be greatest with very short delays, and within the delay, it should be greatest at the beginning and smallest at the end of long delays. This hypothesis makes predictions similar to those of Hypothesis *c* in Figure 4.

Method

The experiment was performed in two parts: In Part A, delays of 100, 200, and 300 msec were used; in Part B, delays of 200, 300, and 400 msec were used.

Subjects. A total number of 21 undergraduate psychology students took part in Experiment 2 (Part A, 6 women and 5 men, mean age = 25.9 years; Part B, 5 women and 5 men, mean age = 25.8 years) as part of their course credits. None of them had participated in the prestudy or the preceding experiment, and all were naive as to the purposes and hypotheses motivating the study. They reported normal or corrected-to-normal vision and hearing.

Stimuli. The stimuli (disks, motions, and sound) were the same as those in the sound condition of Experiment 1, except for the visual delays and the variable position of the sound. Only three visual delays were selected in each part (100, 200, and 300 msec in Experiment 2A, and 200, 300, and 400 msec in Experiment 2B). The relative position of the sound within each visual delay was varied in 7 steps. For example, in the 100-msec delay condition, the clack maximum was at 0, 17, 33, 50, 67, 83, or 100 msec after the stop of the first visual movement.

Procedure and Design. The procedure was the same as that in Experiment 1. Each of the 21 conditions was measured three times, making a total of 63 trials in each part. The order of the trials was randomized individually for each subject.

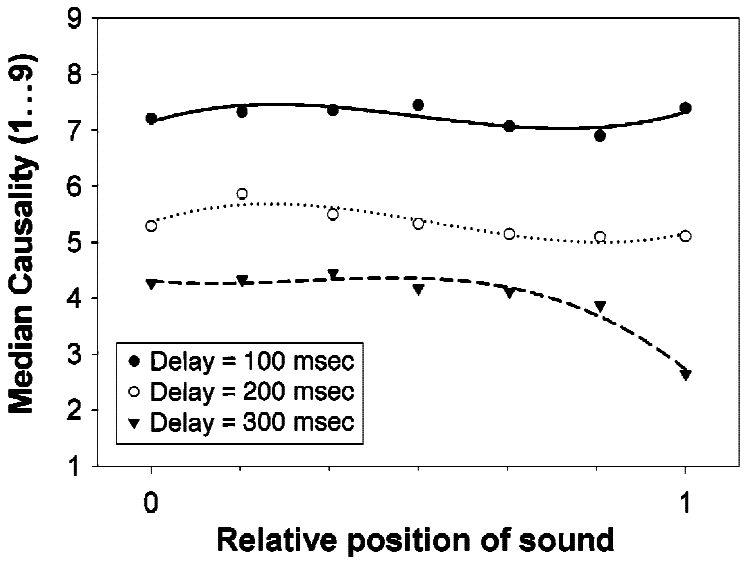

Figure 5. Effect of clack position in Experiment 2A.

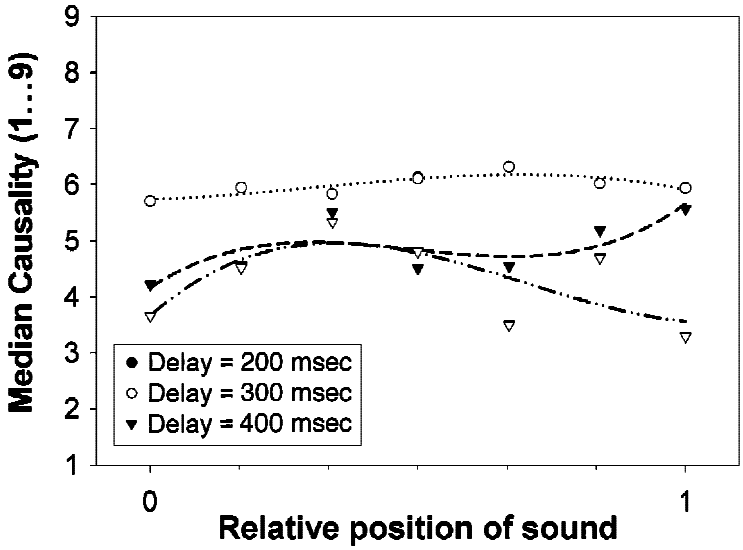

Figure 6. Effect of clack position in Experiment 2B.

Results

As is shown in Figures 5 and 6, the causality ratings at each of the four visual delays were similar to those of the comparable conditions in Experiment 1. In Experiment 2A, the clack position had no clear effect on the causality ratings. There was no significant linear, quadratic, or cubic regression with a 100-msec delay (r^2 = .05, .05, and .62, respectively; $p > .05$ each) and no significant linear or quadratic regression with a 200-msec delay (r^2 = .46 and .51, respectively), but there was a significant cubic regression ($r^2 = .83, p < .05$). With a 300-msec delay, there was no significant linear regression ($r^2 = .59$, $p > .05$), but there were significant quadratic and cubic regressions ($r^2 = .86$ and .95, respectively; $p < .01$ each). However, in single-position tests, even the decrease of the causality ratings at the end of the 300-msec delay was not significantly different from that at the next earlier position (250 msec; Wilcoxon $Z = 1.9, p = .058$). Because we still felt that the position of the clack sound might have an effect in combination with long visual delays, Experiment 2B was conducted with 200-, 300-, and 400-msec delays. This time, there was a weak cubic regression with the 200-msec delay ($r^2 = .71, p = .05$), but the significant nonlinear trends with a 300-msec delay, which were shown in Experiment 2A, were not confirmed in Experiment 2B (r^2 = .39 and .65, respectively; $p > .05$ each). Also, the causality judgments at the 400-msec delay did not show any systematic covariation with the clack position at all (r^2 = .06, .46, and .54 for linear, quadratic, and cubic regression, respectively; $p > .05$ each; see Figure 6).

Discussion

The similarity of causality ratings at the four visual delays used in Experiment 2 with the comparable sound conditions of Experiment 1 showed that the subjects were sensitive to the clack sound. However, no consistent effect of the clack position within the visual delays of 100–400 msec was found. The cubic trend of the clack position within the 200-msec delay was almost statistically consistent, but the form of the trend differed between Experiments 2A and 2B. We conclude that the precise forms of the curves in Figures 5 and 6 are not reliable and that the subjects were not sensible to the position of the sound during the visual delay. The pattern of results falsifies the presumption that the effect of the sound on the causality ratings in Experiment 1 was due mainly to the fact that the sound occurred in the middle of the visual delay, thereby "stretching" the *causality time span* by a certain amount of time. The results also cast doubt on the presumption that the effect of the sound on the causality ratings is due mainly to attention distraction, because we would expect the greatest distraction effect at the beginning of short delays, which did not occur. It seems that the effect of the sound is due mainly to the addition of sudden information within the delay, thereby increasing the causality impression by a certain quantity, as long as the additional information occurs within the visual delay. However, this does not mean that the position of the additional information is completely irrelevant, since previous studies have shown that synchrony judgments between vision and hearing are highest when the sound occurs a few milliseconds later than the respective visual information (see the introduction above). But within the range of visual delays tested here, the position of the sound does not seem to matter, arguing for an impression of a unitary collision event, rather than a predominance of some dyadic cause–effect relation.

EXPERIMENT 3

It is still an open question why the clack sound increases the causality judgments in Michotte-type launching dis-

plays using visual delays. We hypothesized that the sound shortened the perceived duration of the delay, an effect that has been reported several times in the literature for short durations (e.g., the so-called Oppel–Kundt illusion in the temporal domain; see also Boltz, 1991; Ihle & Wilsoncroft, 1983; Underwood & Swain, 1973; Wearden, Edwards, Fakhri, & Percival, 1998). However, it should be noted that the opposite effect also occurs under certain conditions (e.g., rather long time periods). We designed an experiment that compared no-sound delays with sound delays. With a 2AFC procedure, the subjects had to decide which of the two launching stimuli had the longer delay.

Method

Subjects. Ten psychology students (5 women and 5 men, mean age = 25.5 years) took part in Experiment 3 for course credit. None of them had participated in one of the previous experiments, and all were naive as to the purposes and hypotheses motivating the study. They reported normal or corrected-to-normal vision and hearing.

Stimuli. The Michotte-type no-sound launching display of Experiment 1, with a 200-msec delay between the stop of the first and the start of the second movement of the disks, was used as the reference stimulus, and displays with variable delays (40–440 msec in 11 steps) containing the clack sound in the middle of the delay period were used as comparisons.

Procedure and Design. Each trial consisted of the successive presentation of two launching stimuli separated by an interstimulus interval of 1 sec. One of the displays was the reference stimulus; the other was the comparison stimulus. Each of the 11 possible comparisons was presented five times, resulting in a total of 55 trials. The order of reference and comparison was counterbalanced. The order of the trials was randomized. The subjects were asked to indicate which of the two delays between the respective stop of the first and the start of the second disk movements appeared longer. The clack was not mentioned in the instructions. The verbal responses were noted down by the experimenter. All other procedural details were the same as those in Experiment 1.

Results

As was expected, the overall effect of the filled durations on the judgments was highly significant [within subjects, $F(1,10) = 29.35, p < .001$], and the linear model fits were best [$F(1) = 59, p < .001$]. The cubic term was also significant [$F(1) = 14.2, p < .005$]. The point of subjective equality for display durations containing the clack sound was 259.5 msec, as compared with 200 msec without the sound (Figure 7). In other words, the durations filled with clack sound appeared to be about 60 msec shorter than the unfilled durations. This difference is outside the 95% confidence interval of 200-msec duration judgments.

Discussion

The results were consistent with previous studies, showing a subjective shortening of short filled delays, in comparison with unfilled ones. Our results showed a difference of 60 msec between visually defined intervals of 200-msec duration containing no event and visually defined intervals containing an auditory event. It is possible that cross-modal effects are less dramatic, with respect to duration judgments, than are unimodal auditory effects. Differences of 60 msec in cross-modal judgments do not explain the difference between the causality ratings of sound and no-sound launching displays. As is shown in Figures 3 and 4, the tolerated delay in the no-sound condition of Experiment 1 was 110 msec, and in the sound

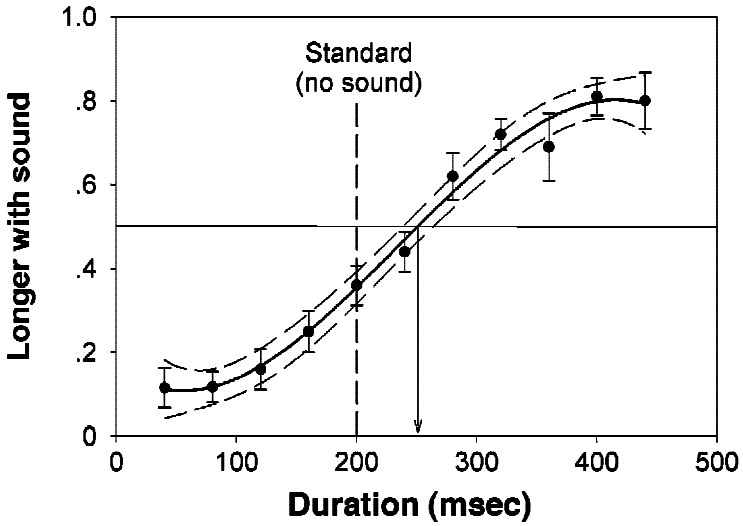

Figure 7. Comparisons of durations of constant 200-msec delays containing no sound with variable delays (40–440 msec) containing a clack sound in the middle of the delay. The *y*-axis shows the probability with which subjects report that the duration seems to be longer when it contains a sound as opposed to when it does not contain a sound. Depicted are means (filled circles), standard errors, and 95% confidence intervals.

condition, it was 280 msec. Although the subjects probably had experienced shorter visual delays when the delays were filled with a sound, this effect seems too small to fully explain the long delays tolerated.

GENERAL DISCUSSION

Our experiments used the classic launching-type display with varying delays between the stop of the first disk and the start of the second one. Some of the delays were filled with a clack sound, some with a visual blink. A number of results occurred that have not been reported before: (1) There was a continuous decrease of causality ratings with an increasing delay; (2) part of the delay could be compensated for by additional sudden marker events (e.g., clack sounds or blinks) within the delay; (3) the nature of the marker event did not seem to play a large role—clack sounds that could be easily interpreted as part of the collision event had the same effect as blink events that had no obvious semantic relation to a collision event; (4) the exact temporal location of the marker event within the delay did not affect the causality ratings; and (5) although the additional marker event had a small effect on the perceived duration of the delay period, this apparent extension could not account for the effects observed in Experiment 1.

Several hypotheses have been discussed and tested in order to explain the increase of causality ratings by means of additional marker events that occur during the delay between the stop of the first disk and the start of the second one (see the introduction section above).

1. *The increase in causality judgments is due to the temporal proximity of visual and auditory events and is no true cross-modal phenomenon.* This *temporal link* hypothesis has been supported by showing that the same increase occurs when blinks are used, instead of clacks, in close temporal proximity to the visual motions.

2. *The increase is due to an increase of the total available information about sudden changes in a stream of events.* This interpretation (*information link*) cannot be ruled out, because all the stimuli used in the noncontrol conditions of the experiments contained additional information about sudden changes, as compared with the control conditions.

3. *The increase is due to the increased similarity of the simulated collisions to real collisions.* Although all the stimuli in the noncontrol conditions contained information about sudden changes, this *reality link* hypothesis is not very probable, because blinks or color changes rarely occur during real collisions.

4. *The increase is due to attention distraction.* The insertion of an additional stimulus in the delay distracts attention from the delay, and observers are less likely to process the causality-contradicting information provided by the delay. Such an effect should be greatest with very short delays and with the additional marker event positioned at the beginning of the delay. The results of our experiments do not fit this hypothesis, because the contribution of an additional stimulus to the mean causality ratings was greatest with medium and long delays (Experiment 1) and the position of the clack sound within the delay did not influence causality judgments (Experiment 2).

5. *The increase is due to the subjective shortening of filled delays*, as compared with unfilled ones. Although subjective shortening was shown in Experiment 3, the size of this effect (about 60 msec for 200-msec delays) cannot explain the size of the contribution of additional stimuli to the causality ratings (about 180 msec for mean tolerated delays).

We conclude that the subjects in our experiments processed most of the audiovisual and visual–visual stimuli as being unitary events. Intuitive logical reasoning may have contributed to the subjects' judgments: Accidental temporal and spatial co-occurrence of two events (e.g., the stop of the first disk in the vicinity of the second disk and the start of the second disk) is rather unlikely, but accidental co-occurrence of three events (e.g., the two events described above, together with a marker event) is even more unlikely. It seems that the probability of spatial and temporal proximity of the events is sufficient to induce causality impressions, irrespective of the perceptual modality.

Until now, it has been an open question as to which of the many informational aspects of a clack sound (or a blink) in the context of visual launching events contribute most to the increase of causality ratings. Possible candidates are (1) the sudden change of the stimulation as such, irrespective of the dimension and modality of change, (2) the specific sudden change of intensity at the beginning of the sound/blink, (3) the maximum intensity of the changed stimulation (sound or light), (4) the sudden return of the intensity (sound or light) to initial levels, (5) the spatial closeness of the events, and (6) the content of the events, which may be associated with events in the real world. In the case of the clack sound, the informational aspects (1–6) covaried in our experiments, and they do so in most real-world collisions. In the case of the blink, this covariation is rarely found in the real world, although one of the subjects spontaneously mentioned "explosion" as an association to the blink. Future experiments will have to test the information aspects necessary and sufficient to enhance multimodal causality judgments.

Although the causality judgments of the sound and the blink conditions in Experiment 1 produced comparable results, we would like to stress qualitative aspects of the launching display, because 80% of our subjects judged the disks to be "real" objects—for example, colliding balls. This may have contributed to the *assumption of unity*, as proposed by Welch and Warren (1980)—that is, a cognitive factor supporting the assumption of a common cause of the separate stimulus events. In the terms used by Jones and Boltz (1989) and Boltz (1991), the scenes containing colliding balls may be seen as a "higher order" coherent structured information sequence. This allows observers to switch between attending to the global struc-

ture (e.g., collision events) and details of the event's structure (e.g., the velocity of motions, the sound source, etc.). When observers are attending to the global structure, certain violations of physical principles of collisions (e.g., a short delay or a sudden color change) will not be detrimental for causal judgments, especially in the case of redundant information (e.g., with sound + blink). This perspective is compatible with our results in Experiment 1 and with the general observation that causal judgments are not necessarily qualitative yes/no responses; they can be ordered on a continuous scale, ranging from *not at all* to *extremely*. The more coherent information that fits a physical causation principle is available, the higher the causal judgment will be.

REFERENCES

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Erlbaum.

Boltz, M. (1991). Time estimation and attentional perspective. *Perception & Psychophysics*, **49**, 422-433.

Crowder, R. G. (1993). Auditory memory. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 113-145). Oxford: Oxford University Press, Clarendon Press.

Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception*, **9**, 719-721.

Gruber, H. E., Fink, C. D., & Damm, V. (1957). Effects of experience on perception of causality. *Journal of Experimental Psychology*, **53**, 89-93.

Ihle, R. C., & Wilsoncroft, W. E. (1983). The filled-duration illusion: Limits of duration of interval and auditory fillers. *Perceptual & Motor Skills*, **56**, 655-660.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, **96**, 459-491.

Kohlrausch, A., & van de Par, S. (2000). Experimente zur Wahrnehmbarkeit von Asynchronie in audio-visuellen Stimuli [Experiments on the perception of asynchrony with audio-visual stimuli]. In *Fortschritte der Akustik* (DAGA 2000, pp. 316-317). Oldenburg: DEGA Geschäftsstelle.

Lewald, J., Ehrenstein, W. H., & Guski, R. (2001). Spatio-temporal constraints for auditory-visual integration. *Behavioural Brain Research*, **121**, 69-79.

Lewkowicz, D. J. (1996). Perception of auditory–visual temporal synchrony in human infants. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 1094-1106.

Maass, H. (1938). Über den Einfluss akustischer Rhythmen auf optische Bewegungsgestalten [On the influence of acoustic rhythms on visual movement gestalts]. *Archiv für die Gesamte Psychologie*, **100**, 424-464.

Michotte, A. (1946). *La perception de la causalité.* Louvain: Publications Universitaires. [English translation: *The perception of causality*, London: Methuen, 1963]

Michotte, A., & Thinès, G. (1963). La causalité perceptive. *Journal de Psychologie Normale et Pathologique*, **60**, 9-36. [English translation: "Perceived causality," in G. Thinès, A. Costall, & G. Butterworth (Eds.), *Michotte's experimental phenomenology of perception* (pp. 66-87). Hillsdale, NJ: Erlbaum, 1991]

Oestermeier, U., & Hesse, F. W. (2000). Verbal and visual causal arguments. *Cognition*, **14**, 65-104.

Parducci, A. (1974). Context effects: A range-frequency analysis. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2, pp. 127-141). New York: Academic Press.

Piaget, J. (1967). Le développement des perceptions en fonction de l'age. In J. Piaget, P. Fraisse, E. Vurpillot, & R. Francès (Eds.), *Traité de psychologie expérimentale: Vol. 4. La perception* (pp. 1-62). Paris: Presses Universitaires de France.

Rudloff, I. (1997). *Untersuchungen zur wahrgenommenen Synchronität von Bild und Ton bei Film und Fernsehen* [Experiments on the perceived synchrony between pictures and sound with movies and television]. Unpublished diploma thesis, Ruhr-Universität Bochum, Fakultät für Psychologie.

Sarris, V. (2000). Perception and judgment in psychophysics: An introduction into the frame of reference theories. In A. Schick, M. Meis, & C. Reckhardt (Eds.), *Contributions to psychological acoustics: Results of the 8th Oldenburg Symposium on Psychological Acoustics* (pp. 39-62). Oldenburg: Bibliotheks- und Informationssystem der Universität.

Schlottmann, A., & Shanks, D. R. (1992). Evidence for a distinction between judged and perceived causality. *Quarterly Journal of Experimental Psychology*, **44A**, 321-342.

Scholl, B. J., & Nakayama, K. (2000, November). *Contextual effects on the perception of causality.* Poster presented at the annual meeting of the Psychonomic Society, New Orleans. [Abstract published in *Abstracts of the Psychonomic Society*, **5**, 91]

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, **4**, 299-309.

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, **385**, 308.

Underwood, G., & Swain, R. A. (1973). Selectivity of attention and the perception of duration. *Perception*, **2**, 101-105.

van de Par, S., & Kohlrausch, A. (2000). Sensitivity to auditory-visual asynchrony and to jitter in auditory-visual timing. In B. E. Rogowitz & T. N. Pappas (Eds.), *Human vision and electronic imaging* (Proceedings of SPIE, Vol. 3952, pp. 234-242). Bellingham, WA: SPIE Press.

Watanabe, K. (2000). When sound affects vision: Effect of auditory saliency on visual motion perception and its relation to crossmodal attention. *Abstracts of Sloan Meeting 2000.* New York: New York University, Center for Neural Science.

Watanabe, K., & Shimojo, S. (1998). Attentional modulation in perception of visual motion events. *Perception*, **27**, 1041-1054.

Watanabe, K., & Shimojo, S. (2001a). Postcoincidence trajectory duration affects motion event perception. *Perception & Psychophysics*, **63**, 16-28.

Watanabe, K., & Shimojo, S. (2001b). When sound affects vision: Effects of auditory grouping on visual motion perception. *Psychological Science*, **12**, 109-116.

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. *Quarterly Journal of Experimental Psychology*, **51B**, 97-120.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, **88**, 638-667.

White, P. A. (1995). *The understanding of causation and the production of action: From infancy to adulthood.* Hove, U.K.: Erlbaum.

White, P. A. (1999). Toward a causal realist account of causal understanding. *American Journal of Psychology*, **112**, 605-642.

Yela, M. (1952). Phenomenal causation at a distance. *Quarterly Journal of Experimental Psychology*, **4**, 139-154.

Zietz, K., & Werner, H. (1928). Über die dynamische Struktur der Bewegung [On the dynamic structure of motion]. *Zeitschrift für Psychologie*, **105**, 226-249.

(Manuscript received August 28, 2001;
revision accepted for publication January 29, 2003.)