— JAMES J. JENKINS, JERRY WALD, AND JOHN B. PITTENGER —

Apprehending Pictorial Events: An Instance of Psychological Cohesion

Much of contemporary experimental psychology concerns the modeling of how people process information, whether it be information contained in a text or on a Go board. For modeling to be successful, a complementary task must be accomplished: the specification of the units of analysis. Specification of the units of analysis and the processes that operate on them are interdependent endeavors, and both depend in turn on the problem area under study.

In some psychological research, the specification of appropriate units does not seem to be a problem. For example, if one models the earlier stages of reading, distinctive features of letters, individual letters, and words suggest themselves as units (or data) to be processed (e.g., Gough, 1972). When one models a particular reading task or the acquisition of reading as a skill, one may begin to determine which units are functionally important, and when and how they are used (e.g., LaBerge & Samuels, 1974. For other examples, see several of the papers in Kavanagh & Mattingly, 1972).

Note: The research reported in this paper was supported by grants to the Center for Research in Human Learning, University of Minnesota from the National Science Foundation (GB 17590), The National Insititue for Child Health and Human Development (HD 01136), and the Graduate School, University of Minnesota. The second author is a predoctoral trainee of the National Institute for Child Health and Human Development under Training Grant, HD 00098. The third author expresses thanks to the University of Arkansas for a travel grant that made collaboration possible. The authors are happy to express their appreciation to Martin Wurthman and Vincent Berg, who helped with the photographic work involved in these studies. The debt the authors owe to Robert Shaw for intellectual stimulation and inspiration is evident throughout the paper; it is gratefully acknowledged. We are also grateful to the following colleagues for criticism and comments: C. Bremer, L. Brownston, D. LaBerge, L. Larkey, K. McGovern, S. Soli, W. Strange, and G. Widin.

In other research domains, however, the units of analysis are more obscure. For example, Jenkins (1974a, 1974b) suggests that both the kinds of units and the kinds of processes involved in verbal recall and recognition change with variations in the nature of the material to be remembered and the orienting task being performed by the subject. The result of this state of affairs is that processing models are harder to specify. Visual perception, with which we shall be concerned in this paper, is another such domain of uncertainty. In the past, several kinds of units have been proposed, ranging from punctate stimuli to the holistic visual Gestalt. Associated with these units are pattern recognition devices and processing models that rely on analysis-by-synthesis routines, template matches, or the like (see Neisser, 1967). However, neither these units nor the processing models based on them seem adequate. As Neisser (1967) points out, they encounter difficult problems (such as orienting or normalizing the input stimuli) and are not always compelling as psychological models of visual processing or visual perception.

It is our aim to suggest another way to proceed. Rather than relying on neutral stimuli or forms as our basic units, we propose to adopt an ecological approach to the problem. In what follows we shall regard the *event* as primary. At the outset, we shall use the term only in its intuitive sense. We cannot at this time give a satisfactory definition of what an event is. We believe that with sufficient experimental and theoretical work, however, the concept will find an increasingly adequate definition. (Shaw and Pittenger, in press, have made first steps in this direction.)

We have taken the intuitive approach for two reasons. First, we acknowledge that events exist at several levels. One might point to microscopic events (such as chemical changes), to ecological events (such as the sun's rising), or to astronomical events (such as the expansion of the universe). Furthermore, events are usually embedded in other events. Consider one's taking one's seat to listen to the first movement of a Brahms symphony during a Brahms-Beethoven-Bach Festival. Because events are nested within events (as one's sitting down is a subevent within the overall event of the festival) and because of the related problems posed by multiple levels of analysis (e.g., the detailed analysis of the first movement of a Brahms symphony versus the more general analysis of the symphonic structure itself), finding one set of defining properties that cuts across all manner of events presents formidable difficulties.

Second, we believe that the intuitive notion of an event can be specified more closely only when we know what supports (or specifies) an event to a perceiving organism. In this paper we are concerned with visual events, so we shall focus on those characteristics that support the perception of events in the visual world. We shall report some beginning studies of the apprehension of visual events. As we go along, certain psychological consequences of having perceived an event will become apparent and certain characteristics of the stimuli that make event perception possible will begin to be seen.

This paper, then, is a first attempt to lift ourselves by our empirical bootstraps to a position from which we may see the nature of various ecological units more clearly. With such information in hand, the complementary task of modeling the apprehension of visual events should also be clarified.

Events: Some Initial Considerations

For our present purposes we may regard events as pertaining to both objects and the changes (or transformations) defined over objects. We must note, however, that although objects and transformations are conceptually distinct, they are not independent of one another. Objects, for example, may be in part defined as those "things" or properties left invariant under certain transformations, such as translation. It is best, perhaps, to regard objects and transformations as two aspects of an event.

To help clarify this view we may recall some properties of objects that are said to be conserved under certain transformations, properties that contribute to our meaning of *object*. Consider, for example, the Piagetian experiments on "conservation." The point of the experiments for our purposes is not that properties are invariant (or conserved) under all possible changes that may occur; indeed, the point is just the opposite. Some properties are conserved under some types of changes or alterations, and those properties left invariant contribute to our meaning of the object within the situation. For example, *number* of marbles is conserved over various spatial arrangements of the marbles, but it is not conserved over the operation of running the marbles through a grinder or smashing them with a sledgehammer. *Volume* of a liquid is conserved over various transformations of shapes of containers, but it is not conserved over evaporation, loss through spilling, addition of more liquid, drinking some of the liquid, etc. *Amount* of matter is conserved over translation, reshaping, amalgamation, and subdivision, but not over burning, addition, subtraction, eating, dissolving, etc. It is precisely the relation between the type of change and the relevant property of concern that is being investigated in such experiments. For the sophisticated organism the crucial perception is that the operations performed on the material in question do or do not constitute a type of change that affects the property, and hence the object, whose invariance is at test. It makes no sense at all to talk about the conservation of number, volume, or matter without talking about the nature of the changes involved.

In the above examples it is easy to lose sight of the importance of the transformational aspect of an event; the object and its properties seem to predominate. But the importance of change cannot be overlooked. In many domains we know that the rate of change itself is a critical determiner of perception. Michotte's classic experiments demonstrate that the relative rates of movement of two objects striking one another determine whether the event is seen as "entraining," "launching," or "triggering" (Michotte, 1963). And in other domains it is the "style" of change that not only specifies the nature of the event but also the nature of the object under change. Johansson (1975) has made elegant motion pictures of "point-light people" that show the power of change in the configuration of lights over time. These films were made in the dark and show only the patterns of movement and the disappearance and appearance of point-light sources mounted at the joints of the human subjects being pictured (shoulders, elbows, wrists, hips, knees, and ankles, with a single light at the crown of the head). In a static frame of the film, an observer sees nothing but an unorganized jangle of lights. In the running film the moving patterns of lights are sufficient not only to specify the actions of walking, running, dancing, approaching, receding, and transversing, but they enable the observer to specify that the moving "object" is a human of a particular sex.[1]

What is crucial, then, is the nature of the change as well as the object and its properties left invariant under the change, for these jointly influence (if not determine) the quality of the event that is perceived. Only the two aspects taken together can make the event cohere.

Some Experiments on Pictorial Events

Many years ago, Esper (1925) published an experiment that demonstrated that people may learn more than they have experienced. He presented subjects with pictures of four forms, each of which appeared in four different colors, and required the subjects to learn the names that he had systematically paired with the colored forms. The labeling system was such that the shape of the form determined the first syllable of the name and the color of the form determined the second syllable. Instead of presenting all 16 instances of the colored forms, however, Esper withheld two particular instances. Following training on the 14 items, he tested all 16 items in a naming test. He found that the subjects correctly named the two new items when they were presented. Surprisingly, the subjects could not even say which items were old and which items were new. Thus Esper demonstrated that when there is a systematic relation between stimulus variation and response variation, subjects may learn the complete system of relations even though they do not see all the members, and that once they learn the *system*, they may not even know which instances they have seen and which they have not.

Esper's study is not just a selected curiosity; many studies of the Esper sort have been conducted (see Esper, 1975, for an extensive account), and this result is frequently found. In attempting to explain the Esper results, one is forced to conclude that subjects go beyond the learning of the finite set of stimuli with which they have been presented. That is, the stimuli that are presented to the subjects appear to specify a system of relations within which the individual stimuli cohere. Furthermore, these relations apparently specify additional stimuli that were not in fact shown. In some experiments (e.g., Foss, 1968) the additional stimuli were recognized as belonging to the system of relations and were correctly named, but the additional stimuli were also usually recognized as being new. In other cases (e.g., Segal, 1962) the new stimuli were recognized as belonging and correctly named but were not distinguished from the training stimuli; i.e., the subjects did not know that the stimuli were new. The first case (the Foss experiment) may be called *coherence*, and the second case (the Segal experiment) may be called *fusion*. It will be seen that coherence is the general case and fusion is the special case observed when specific memory for exemplars is lost or overwhelmed by the constraints of the system.

Although Foss's and Segal's data appear inconsistent, the inconsistency is more apparent than real. The difference in their findings points instead to a critical distinction, a distinction to which we shall repeatedly return. The distinction is this: when presented with any set of related stimuli (i.e., coherent stimuli), subjects may acquire two forms of knowledge—knowledge of the particular stimuli experienced and knowledge of the underlying coherent system of relations. The first may be said to be concerned with the experimental setting itself as an event (e.g., with the particular materials and tasks at hand) and the second with the system as an event. Consequently any experimental outcome will be some product of both of these bases of knowledge, the contribution of each to any particular situation being determined by a host of factors.

We should also note that the characteristic findings of Esper, Foss, and Segal are not limited to their particular experimental paradigm; analogous findings obtain with other procedures and materials. Several experimental examples might be cited (Franks & Bransford, 1971; Posner & Keele, 1968, 1970; Strange, Keeney, Kessel, & Jenkins, 1970), but for our present purposes the conclusions from one such study may suffice. Wilson, Wellman, and Shaw (reported in Shaw & Wilson, 1975) used a system in which four simple geometric forms were orbited around the four corners of a square. From the total set of 16 stimuli two subsets were selected: one subset constituted a generator set; when presented to subjects these stimuli and their transformational relations were sufficient to specify the total set of 16 stimuli. The second subset was not a generator—i.e., it could not specify the complete system. One group of subjects saw the generator subset; another group saw the nongenerator subset. The subsets were presented only once. Following this presentation, each group was given a recognition test consisting of all 16 stimuli plus 9 control items. Only the group seeing the generator set falsely recognized the appropriate new items, items that belonged to the coherent system specified by the original generator set. Shaw and Wilson (1975) concluded that what one remembers depends upon the generative power of the set of instances to which the subject has been exposed. In short, as Garner (1974) has noted, experimental subjects do not deal in some simple fashion with only the stimuli that the experimenter presents. Clearly they respond on the basis of a set of possibilities that the presented stimuli may be said to define or generate.

These findings seem at odds with much recent work that implies that visual memory is precise and virtually unlimited. Shepard (1967) and Standing and his colleagues (Standing, 1973; Standing, Conezio, & Haber, 1970; also see Haber, 1970) have demonstrated that if subjects are shown large numbers of slides (up to 2,000), they can identify them with high accuracy on a forced-choice recognition test. (Subjects are asked to choose the slide they have seen before when an old slide is paired with a new slide). Standing (1973) has carried this demonstration to the extreme of 10,000 slides presented over a period of five days. He found high accuracy of recognition for sample sets of these slides and estimated total retention as 6,600 items. Postman (1975) was so impressed by this demonstration that he offers it as evidence to "anyone who doubts that pictures are easy to remember" (p. 322).

Obviously these results are in striking contrast to the findings of Esper and the data obtained by Shaw and Wilson for their visual materials. With unrelated visual materials, memory for pictures seems virtually unlimited and precise. With just a few systematically related slides, on the other hand, memory for particular instances of visual displays is poor. One may, of course, attribute the conflicting results to differences in experimental paradigms (forced-choice recognition versus a variety of other operations); indeed, some of the difference may reside in the sensitivity of the forced-choice technique, but we believe that a major difference in outcomes is attributable to a radical difference in stimulus materials. In the experiments with arrays of forms, colors, patterns, etc., the stimuli form a system, some kind of coherent whole that the subject apprehends. In the experiments with massive numbers of slides, every effort is made to keep the pictures unrelated; i.e., each slide is a slice of a separate event, unassimilated and unassimilable except as a discrete event in itself. Such a slide, then, constitutes an event with a frequency of one exposure that is later to be compared with some other slide, another unique event, that has a frequency of zero in the subject's experience. Put in this way, the recognition of large numbers of slides may not seem to be such a dramatic feat as we had first supposed. Perhaps we should have suspected some such performances, given the older literature on frequency effects. In many identification and recognition tasks there is a sharp discontinuity in familiarity between a stimulus that has been experienced once and a stimulus that has never been experienced. As long as the events stay separate and unique and have frequencies of zero or one, there is little remarkable about knowing which is which.

Shepard's work, in fact, suggests that frequency of exposure was important in achieving the results mentioned earlier. In another experiment, he discovered that the high accuracy of recognition of selected pictures was lost if *pairs* of slides were originally presented to subjects who were asked to remember only one of the two slides. In this case both of the slides now had an exposure frequency of one, and the subjects no longer performed with high accuracy on a forced-choice recognition test that asked them to pick from the two slides the one they were supposed to learn. One possible interpretation is that picture memory itself is not being tested in these experiments; the experiments may be testing for the recognition of events one has formerly experienced (even though in impoverished form) as opposed to events one has not previously experienced. (But, as we shall see, this cannot be the whole story.)

Striking evidence that effects are different when the pictures to be remembered bear some relation to each other is found in a study by Goldstein and Chance (1970). These investigators discovered that memory for pictures was seriously impaired when the pictures were all of the same genre: all ink blots, all faces, or all snowflakes. In sharp contrast to the high accuracy of recognition usually observed for unrelated pictures, and in spite of the very modest number presented to a subject (only 14 exemplars of one of the three classes), the results in a "yes-no" recognition test showed 28% errors on faces, 49% errors on ink blots, and (despite their legendary differences in form) a whopping 64% errors on snowflakes. Thus, when the "experimental event" becomes one of viewing related, though highly differentiated and discriminable, members of a class of objects, a decrement of absolute recognition is observed.

Perhaps a *gedanken* experiment is appropriate at this point. Suppose we take a motion picture film as our visual display. If we show a portion of the film consisting of some simple event, we can assume that observers will report having witnessed that event. Now suppose we take all the odd-numbered frames of the film and splice them together, and all the even-numbered frames and splice them together. If we show the odd-numbered sequence to an observer and then show him the even-numbered sequence, he will almost certainly report that he is seeing the same thing, namely, a somewhat "jumpy" movie of a single event. We would be surprised if he reported that he had never seen the second film before, although in the technical, physical sense, he has not.

The *gedanken* experiment suggests a continuum of relatedness running from the intact movie on one end, through a series of frames with an increasing number of intervening frames removed, to the presentation of a series of scenes of isolated events of the Shepard, Haber, and Standing variety on the other end. Reflecting on this potential continuum enticed us to approach the event perception problem through a series of still pictures which were in themselves separate and distinct enough that the question of their confusability would not arise but that, taken together, presented a dynamic event: something like a picture story or a slide show that tells a story.

The question we posed was the following: If a subject saw an appropriately ordered sequence of pictures that was sufficient to give him all the necessary information for an event, would he give us evidence that he had experienced that event in its entirety? Would he, for example, falsely recognize pictures of the event that he had not seen before? Would he be able to reject pictures that were highly similar to the pictures he had seen but that violated some invariant of the event or some detail of the observation?[2]

We decided to begin with some natural but simple, everyday events: a woman making a cup of tea, a teenage girl answering the telephone, and, as one kind of control event, some pictures of people at a party. The first two cases clearly told a story. In the first, the woman standing beside a table in a dining room unwrapped a tea bag and put it in a cup on the table. She left the room and returned with a sugar bowl, which she put on the table. She left again and returned with a tea kettle from which she poured water into the cup. She then returned the kettle to the kitchen, came back into the room, sat down at the table, removed the tea bag from the cup, added sugar, and took a cautious sip of tea. Pictures were taken from a fixed station point, with the camera oriented so that the woman was always near the center of the picture.

In the second event, pictures were again taken from a fixed station point. A girl appeared in the doorway, crossed the room, and picked up the phone. She talked for a few moments while standing, then sat at the desk on which the phone rested, put her feet up on the desk, smiled and laughed, put her feet down, and hung up the phone.

The third sequence of slides *could* have been construed to make a loose story, but the pictures were taken from two different station points and no particular story was apparent. A graduate student was seen arriving at a party, walking across a room, sitting on a couch with other students, talking to a visitor (who was also shown alone). Several new people came and went from subsequent pictures, which were mainly of a single corner of the room.

In each situation, "control" pictures were taken. For the Tea Sequence, additional pictures were taken with a new brightly-colored object on the table with the tea things, with the woman wearing glasses, with the woman pouring water with her left hand instead of her right, with the camera very close to the table, and with the camera at a new station point across the table. Control pictures for the Telephone Sequence involved changes in distance of the camera from the girl, a station point diagonally across the room so that the direction of the girl's walk was from right to left instead of from left to right, and different postures at the desk with the phone. Controls for the Party Sequence were other pictures of the party from the same station points. The pictures involved the same people but they were in different postures and different combinations.

These pictures were presented to subjects in the following fashion: For each sequence the original series of pictures was shown except that every third picture in the sequence was removed. For example, in the Tea Sequence, 26 pictures had been taken. In the initial presentation the subjects were shown pictures 1, 2, 4, 5, 7, 8, 10, 11, etc. up to 26. Each slide was shown for about five seconds. This presentation was then repeated to ensure that subjects were familiar with the pictures. At this point subjects were told they would see a test series of pictures. They were asked to indicate which pictures they had seen before. Subjects were then shown a random series of slides consisting of 8 of the original pictures (Originals), the 8 pictures that belonged in the series but that had not been shown (Belonging slides), and 8 slides that did not fit the sequence (Controls). For the Telephone Sequence, 10 slides were presented initially. The test series consisted of 4 Originals, 4 Belonging slides, and 4 Controls. For the Party Series, 10 slides were presented for learning. The test series consisted of 4 Originals, 5 slides that could have been used to make a loose story (Belonging slides?), and 3 Controls.

It should be noted that this experiment is a very strong test of our hypothesis. Our fundamental assumption is that if the pictures show an event taking place over time, the subjects will apprehend the event. For our first test to work successfully a further assertion is necessary, namely, that having apprehended the event, the subject will be unable to reject a picture that fits the specifications of the event he has experienced. This is in spite of the fact that the two presentations in original learning ought to ensure specific memory of the slides. Thus we must argue that specific memory for individual pictures will be outweighed by the abstract or general memory for the event experienced. At the same time we shall argue that some aspects of memory will be enhanced; specifically, that any picture that violates the constraints or invariants of the experienced event will be detected as new, no matter how much it resembles the original pictures in terms of its elements.

The results of the experiments were very gratifying. For the Tea Sequence 80% of the Originals were recognized as originals, 50% of the Belonging slides were falsely called originals, and only 10% of the Controls were falsely called originals. This clearly demonstrates that false positive recognitions can be obtained for new slides when they fit the overall constraints of the experienced event. The fact that the subjects detected the Controls as being new is evidence both that they were attending to the event and that they were sensitive to the particular details of the event they had experienced.

Results for the Telephone Sequence convincingly corroborated the findings of the Tea Sequence. Ninety-four percent of the Originals were recognized as being originals and 42% of the Belonging slides were falsely called originals, but only 3% of the Controls were called originals.

As we had expected, the Party scenes yielded results different from those of the Tea Sequence and Telephone Sequence events. The Party results were much more in agreement with the traditional picture memory studies discussed above; Originals were correctly recognized as originals 83% of the time and Belonging and Control slides were falsely called originals less than 10% of the time. There were very few false positive recognitions.

A reasonable interpretation of these data seems to be that events can be (and are) apprehended when they are available in pictorial stimuli and that apprehension of the event has a marked effect on subsequent recognition behavior. Further, it seems appropriate to talk about "experiencing the event." The Control slides showed us that subjects were quite sensitive to the invariant details of an event. As one example, consider the station point of the camera. Subjects correctly identified as *new* any slide portraying the event from a station point or distance other than the one shown during the original presentation of the event, even though the slide was accurate with respect to all other details. A second example occurred in the Telephone Sequence, in which the only picture of the girl smiling happened to fall into the Belonging group of slides. Almost all subjects rejected this slide even though it was like the original sequence in all other respects. That is, they were extremely sensitive to the display of emotion on the face and correctly knew that they had not seen any such display during the course of the event they had witnessed.

As we went on to further studies, we were impressed with the sharp contrast between the effects seen in these experiments as compared to the studies of isolated pictures. First, recall that the traditional studies that used hundreds or thousands of isolated, unrelated pictures obtained their high levels of picture recognition with only one presentation of the stimulus set. In our studies of coherent events, in spite of the fact that the 10 to 20 slides were shown twice, we consistently obtained high levels of *false* recognition for the Belonging slides. Second, as mentioned above, the studies of isolated pictures are sensitive to frequency effects; once a subject has seen the pictures used as "lures" in the recognition test, his ability to discriminate original pictures from such lures is greatly reduced. This appears not to be the case with pictorial events, as our next experiments showed.

In one classroom experiment we presented the Tea Sequence twice at the beginning of class and then tested for recognition at the end of class. On the next class day we explained that we were going to do exactly the same experiment again, and we performed the experiment in the identical manner. Again on the third class day we performed another exact repetition. The results on all three days were virtually identical with the results we obtained on the single administration. Subjects got neither better nor worse. Apparently the information from the event was virtually complete with the original two showings; repeated exposures of the Originals, Belonging, and Control slides made little or no difference. The Belonging pictures that were originally judged to belong to the class of original pictures continued to be so assigned, and those that were not so judged at the start were not later on. Whatever kept these latter slides from being perceived as belonging to the event continued to prevent their false positive recognition. Similarly, reexposure of the Controls did not contribute to their false recognition; they were still clearly seen as *not* belonging to the event, and their repeated exposure did not change that categorization.

Anticipating the criticism that perhaps the waiting period (the 45-minute class period) disadvantaged the subjects in this experiment and prevented them from using "fresh" visual images even after they truly understood the task, we performed a replication of the repeated trials experiment in another class. In this case the procedure was the same except that the test series immediately followed the original learning series on each of the three days. The results were the same as in the earlier experiment; Belonging slides that were initially falsely recognized continued to be so, but Controls were consistently rejected. When we analyzed the responses to each slide over the three days, however, we did see an interesting trend in the Belonging slides. The eight slides in this category were divided evenly into four slides that the majority of the subjects thought had been in the original series, and four slides that the majority thought had not been in the original series. On successive trials these slides tended to polarize; i.e., the accepted slides became even more widely accepted and the unaccepted slides became even less well accepted.[3] This suggests that even though the event is well defined on the first occasion, it may become even better specified with repeated exposures.

Further Explorations

The major conclusion we draw from the studies we have just described is that events can be primary units of analysis. We can specify events with a sequence of slides and influence subjects' recognition responses when they perceive the coherence of the events. In this way our results resemble those of Esper (1925), and Shaw and Wilson (1976). When the experimenter specifies the structure of a system with a set of systematically related stimuli, subjects learn (or "pick up") that system. However, in the case of pictorial events, what is the nature of the relations that specify the event? An obvious, but overly simple, suggestion is that it is the raw physical similarity of the slides in the original set that makes them cohere. We have already seen that this cannot be the whole story, given the way we constructed our Control slides and the data we obtained from them. The next studies shed additional light on this suggestion.

In the next set of studies, we randomized the original order of presentation for each of the three picture sequences we discussed above. We predicted, of course, that the randomization would do nothing whatsoever to the recognition performance of our subjects in the Party Sequence. Since no event was picked up when the pictures were shown in their original sequence, there was no reason to suppose that an event would be created by their randomization. We predicted specific picture recognition as before, and that is exactly what we observed. The results for the Party pictures duplicated those of the first experiment.

We did not know what to predict for the organized events. One might suppose that some events are so intrinsically ordered that any presentation of details can be correctly ordered by the observer. If this is the case, the event will be apprehended in spite of random ordering (especially because the slides are presented twice prior to the recognition test.) Further, Garner (1974) argues that any subset of slides will specify a set of alternatives. Thus, while original order is lost, rather considerable constraints will remain as to the set of possible slides that the subject might have seen. If the information is coherent enough to specify the nature of the event, that may be sufficient to determine the same pattern of false positives for the Belonging slides that we had seen on the ordered presentation. On the other hand, if the event is intrinsically only weakly ordered or if it is time-dependent, perhaps the specific memory for individual pictures will be manifest.

Fortunately for the stimulation of future research, both of these outcomes were observed. The Tea Sequence pictures, even though randomly presented, yielded the same results observed above; about half the responses to Belonging slides were false positives, but Control slides were rejected. When the Telephone Sequence was presented randomly, however, the test series yielded high recognition responses only to the Original slides, with almost no false positives for the Belonging slides or the Control slides. Although we cannot at this time specify the source of the coherence of the Tea Sequence as opposed to the Telephone Sequence, we see such specification as an attractive research possibility.

The fact that the Telephone series breaks down when it is presented in random order is useful in that it furnishes valuable information about the role of picture similarity. Obviously similarity between individual pictures could not be the source of the false positives that were originally observed for the Belonging slides in this sequence. If the false positives had been due simply to picture similarity, there is no reason for the order of presentation to make any difference at all. It is tempting to think that the Telephone Sequence is close to some critical point on the continuum between the split movie and the array of isolated events discussed above. The set of pictures is apprehended as a coherent event when the appropriate order of presentation is followed, but the individual pictures are so dissimilar that they are perceived as a set of unrelated pictures when the order is scrambled. Thus order in time is an important source of information in specifying possible alternatives, at least for some events.[4]

These early experiments were crudely done with a hand-held household camera and with very little precision as to timing of the photographs, lighting, specification of the station point, etc. Given their promising outcomes, we decided to simplify our events and improve the technical quality of the materials. From these new experiments a few examples will show that the phenomena can be enhanced impressively.

One of the best series we have done to date is called Orbiting. This series shows an octagonal tray sitting on a black background. In the center of the tray is a large jam jar. The various pictures show the tray and jar immobile while a small saltcellar moves from one position to another around the rim of the tray through each vertex and each midpoint between vertices. Sixteen pictures make up the series and complete the orbit. The learning series was prepared by drawing randomly two slides from every subset of three ordered slides (so that the missing slides would not be periodic). As before, the sequence was shown twice. The subjects were then tested on five slides from the Original series, five Belonging slides, and five Control slides that violated some aspect of the experience of the event (distance, perspective, relation of the saltcellar to the rim, reversal of jar and saltcellar, missing objects). The effect was very striking. Subjects correctly recognized Original slides as originals 89% of the time. They incorrectly identified Belonging slides as originals 73% of the time (false positives) and *never* identified Control slides as originals (zero false positives). It is interesting to note that four of the Belonging slides are as well accepted as the actual originals. Most of the detection of a Belonging slide as a new slide occurred on one slide that showed the saltcellar emerging from behind the jam jar. This was the only picture in the series that showed partial occlusion of one object by another; this was apparently sufficiently important as a "subevent" that its distinctiveness was noted by most of the subjects.

This series illustrates how compelling the fusion process can be. The details that supported the apprehension of the event are almost completely lost or merged in the quality of overall event. One simply cannot believe that the Belonging slides are new because they are so much a part of the completely apprehended event. At the same time we can see that the test of event apprehension employed in these studies is too strong. All of the subjects knew that the slide showing occlusion belonged in the series, but they also knew they had not seen that particular instance *just because* it was a particularly distinctive portion of the event. It is in such cases that the difference between subjects' knowledge of the event and specific knowledge of what they have seen becomes apparent.

Fusion and the loss of memory for particular pictures is a phenomenon that serves to call our attention to the dominance of the event, but it is not a necessary phenomenon in that some specific knowledge of the specifying stimuli also occurs. Obviously we should be able to develop a variety of effective ways to explore the coherence of perceptual events. For example, if we were not interested in the subjects' knowledge of what was seen but were interested only in the coherence of the event and the subjects' knowledge of the quality of the event, a sufficient test would be to ask, "Does this picture belong in the event which you have just witnessed?" (rather than asking which slides had been seen before). (See Baggett, 1975 for some related research.)

Questions regarding both kinds of knowledge, particularly when asked of carefully constructed sequences and ingeniously selected Controls, should greatly enhance our understanding of the perception of events, the kinds of information that specify events, and the kinds of information that support the various qualities of the events we are interested in. The following incident is instructive. In examining the Control slides of the Orbiting series, we learned another aspect of what-is-perceived. Our camera man made one slide that fitted the series perfectly well but used a different position for the light sources. This slide is impressive because an observer knows immediately that it is not one of the series, but the source of the difference is not apparent for some time. Then one suddenly becomes aware that the shadows are wrong, something that almost no one would specify if asked to describe the picture. This points out again to us that any invariant in the situation can become an important cue to divergence. The invariants are accepted as the defining properties of the event or constraints on "what counts" in the pictures. One becomes aware of these invariants when they are violated, although they may not be given in the description of the event or even be available in consciousness (see Garner, 1974, for more on this point). It seems to us that the converse of this may also hold. If some aspect of the event varies freely in the learning series (e.g., the quality of lighting in the original Tea series), it is ruled out as a defining property of the event, and unless this variable reaches extreme values in the test series, it will be ignored. What is important here is that what is taken to be invariant or deviant for any event will be defined over the course of the event itself. In this sense events are self-defining, and they may be studied as such.

Studies bearing on the power of events to specify their own important characteristics have been carried out by Robert Kraft, who was interested in a special aspect of picture memory. Kraft pointed out to us that picture memory could hardly be images because left-right orientation was often not preserved. When Standing, Conezio, and Haber (1970) tested subjects for their knowledge of whether a slide was reversed or not, they found a marked drop in the accuracy of orientation information over 24 hours even though subjects were still highly accurate in identifying pictures they had seen against new pictures. In Kraft's own work on memory for orientation of human profiles he found virtually chance identification of the original left-right orientation, even when subjects were warned that they were going to be tested on orientation.

The point of view espoused in this paper suggests that it should be easy to enhance memory for orientation. If the event portrayed in a series of photographs had a natural movement through space that was an intrinsic part of the event, subjects should be able to remember orientation far above chance because orientation would be defined over (and hence a defining property of) the event. Kraft and Jenkins (1977) developed three picture sequences that portrayed events flowing to the left and to the right. Each event had both left-going and right-going actions, but these were part of the overall event in a natural way that made the orientation of objects and movements an integral part of the story. One story involved a boy and a girl. The girl dumped snow on the boy, the boy chased her and was about to hit her with a snowball. A second story showed a woman going out of the house to a shed, getting a box from the shed, and loading it in her car. The third story followed a girl to a skating rink and watched her put on her skates and then skate off across the pond into the distance.

One group of subjects saw these slides in their correct order and correct left-right orientation. They were then tested on a randomized set, half of which had been reversed in left-right orientation. The subjects were 94% correct in assigning the proper orientation to the slides. One group of subjects saw the slides in random order but with the proper orientation. These subjects apprehended the stories, despite the random order, and were 91% correct in assigning the correct orientation to the slides. A third group of subjects saw the slides in random order and with random orientation. (This group serves as a control for the memory for orientation of individual slides.) These subjects performed poorly when asked to designate original orientations; they were correct in only 67% of the cases. Thus, when orientation is an integral property of the event being portrayed, left-right orientation of test slides can be correctly designated. But when orientation is simply an arbitrary property of an individual slide, subjects are not very successful in remembering it.

As a further interesting variation, Kraft presented a group of subjects with the ordered set of slides in the correct orientations and then tested them on the orientation of Belonging slides (slides that fitted the stories but had not actually been presented). The subjects assigned the correct orientation to these new pictures 90% of the time. Kraft's work is both interesting in itself and suggestive of new directions for research and new methods of determining what subjects have learned about events.

Finally, one other direction of research must be mentioned. All the materials above have been developed from the perspective of the static observer, yet this is only one kind of visual experience giving rise to events. Information is also available over time to an observer who is moving through an environment (Gibson, 1966). Accordingly, we undertook an inverse experiment: the observer moving over the still landscape. The event in this case is a walk across a campus from the student union to the psychology building and back again. The pictures were taken early one Sunday morning and show the campus empty of people. Every 20 paces or so the walker (J.P.) took a photograph looking straight ahead on his walk. Control pictures were other scenes of the same campus, other pictures of some of the same buildings taken from positions off the walk, scenes taken along other walks at the university, etc.

The experiment presented a set of slides that "took the subjects for a walk," as described above, and a set of test slides that evaluated the subjects' ability to discriminate old from new slides among the Original, Belonging, and Control slides. In all but one respect, the procedure was similar to the procedures used above: the learning set was shown twice in the appropriate order, and uniformly distributed slides of the original series were held out as Belonging slides. In addition, however, we withheld a sequence of six consecutive slides from the return walk, thus leaving a considerable gap in the temporal-spatial sampling of the walk.

On the test sequence the subjects performed very much like subjects previously studied with the events from static points of observation: 82% of the Originals were correctly identified as having been presented before, 70% of the Belonging slides (including two from the "gap") were incorrectly identified as having been in the original set (false positives), and only 11% of the Control slides were falsely recognized as having been seen before. When separated from the other Belonging slides, the slides in the "gap" were identified only 27% of the time as having been seen before. With these slides removed, the overall results are even more impressive. The Belonging slides that were simply interspersed along the walk were falsely recognized 83% of the time as having been seen before. This means they were indistinguishable from the slides that were actually presented.

In this context one may ask why the Belonging slides in the gap were so poorly recognized. Because so many slides had been excluded, either the content of the gap was not specified for these subjects or the subjects became aware of the gap in the presentation and specifically noted that such views were excluded from the original series. In an effort to shed more light on this question, we repeated the experiment under different conditions.

In the repetition of the experiment, the learning set was presented in random order. We felt that the scrambling of space-time order would provide some evidence as to the coherence of the event and would help us decide between the alternative accounts of the slides in the gap. According to one alternative, subjects might still identify the nature of the walk in the random series but never notice the existence of the gap. If this were the case, it could be argued that the Belonging slides in the gap would not be different in recognition from the other Belonging slides from the random series. On the other hand, it could be argued that the Belonging slides in the gap would be treated just like Controls because no matter what the order of their presentation, there was not enough information in the learning series to specify them as possible alternatives.

The results of the random presentation were similar to those of the ordered presentation: 76% of the Originals were correctly recognized, 58% of the Belonging slides were falsely recognized, and only 8% of the Controls were falsely recognized. When the slides from the gap were considered alone, they showed 20% false recognition, just about the same percentage as that observed in the ordered presentation.

When we exclude the slides in the gap, the Belonging slides in the correctly ordered series received 15% more false positive responses than the same slides in the unordered condition (83% versus 68%). It appears, then, that the appropriate temporal-spatial order was important in inducing the high level of responding to the Belonging slides in the original experiment.

With respect to the slides in the gap, however, there seems to be little difference between the correctly ordered series and the random series; both show very low rates of recognition. Thus we can conclude that the slides from the gap are not well specified in either presentation. There simply is not sufficient information in the original set of slides to determine or support these alternatives.

These experiments with the moving observer, unlike our earlier experiments, offer the further possibility of determining how much of the false recognition effect was attributable to the general knowledge subjects had of the physical campus and how much was attributable to the visual information present in the slide series alone. To exploit this circumstance, we performed the same random and ordered experiments on a similar population of students at another university. These students, of course, could be expected neither to identify any of the buildings nor to have any knowledge of the general campus layout. If the false positives in the original experiments are attributable to extensive knowledge of the constraints of the campus and to the subjects' awareness of the nature of the walk, then the naïve subjects should show little or no false recognition of Belonging slides. If, on the other hand, the walk is specified as a coherent visual event in itself, then subjects from another campus may be expected to show the same phenomenon of false recognition.

The repetitions of the experiments with the second population of subjects yielded several interesting comparisons. The subjects who saw the ordered series responded in the usual fashion of subjects viewing some coherent event. They recognized Original slides 85% of the time, Belonging slides (excluding the gap slides) 54% of the time, and Control slides 4% of the time. The subjects who saw the randomized series responded somewhat more profusely to all cases: Originals 89%, Belonging slides (again without the gap) 66%, and Controls 9%.

When the data for these groups are compared with those for the students who knew the campus, the parallel is remarkable. The data for the randomized presentations are almost exactly the same (except that students who did not know the campus were somewhat more likely to recognize correctly the Original slides than students who did know the campus). The data for the *ordered* presentation, however, show a striking difference in the recognition rate for the Belonging slides. The subjects who knew the campus believed 83% of the time that they had seen the Belonging slides, whereas those who did not know the campus believed only 54% of the time that they had seen the same slides, a difference of 29%.

It appears that this series of pictures is fortuitously chosen to reveal both the nature of coherence of a new visual event and the contribution of personal knowledge to that event. The series is sufficient to specify the event in enough detail to make the interpolated slides "familiar" even to an outsider or even when presented in random order; but at the same time personal knowledge and correct temporal-spatial order specify the total event even more fully and result in almost complete fusion.

Frankly, we had not expected so strong an outcome. Even with two viewings, the slides leave the naïve observer with the impression that he knows very little about the walk. Yet one of the things he does know is that a walk is specified. Almost always one sees the path itself in a relatively constant position on the screen. This invariant alone is sufficient to reject some of the Control slides, but it will not, of course, reject any Belonging slide. That such a single cue is not sufficient to account for all the data is shown by the fact that some Belonging slides with this detail are rejected and the fact that some Control slides that show this invariant are likewise rejected. These results have sensitized us to the fact that real events may have many more sources of coherence than those of which we are typically aware. They challenge us to specify such sources sufficiently well that we can design (or synthesize) new slides that will behave like Control or Belonging slides as we manipulate the variables which specify the event.

A final study may be mentioned. Although it is a study of number recognition, it is included here because it is an extreme case in which a "hidden" invariant can be apprehended by the subject. When the invariant is apprehended, it makes a difference in his construal of the experimental event and results in false positive recognitions like those in the experiments described above.

In this study 30 numbers between 0 and 200 were presented to the subjects. They were told to study the numbers as they were presented one at a time and to attempt to remember them so that they could recognize them later. The numbers were shown for five seconds each; there was only one presentation. One group of subjects saw the numbers in ordinal sequence (e.g., 2, 12, 18, 22, 30, 34, . . . , 190); the other group was given the numbers in random order. The numbers were all multiples of two. Subjects were given a recognition test of 20 items; 6 Original, 6 Belonging (e.g., 8, 16, 24, . . .) and 8 Controls (e.g., 15, 23, 31, . . .).

The group receiving the ordered presentation responded correctly to 77% of the Originals, 46% of the Belonging numbers, and 8% of the Controls. The group receiving the unordered set responded correctly to 74% of the Originals, 49% of the Belonging, and 20% of the Controls. Except for the high rate of responding to the Controls in the unordered group, this result looks very much like the results obtained with the perception of simple events reported earlier. It appears that the invariant in this event (all numbers being even numbers) is a strong determiner of the recognition response and influences responses to about half of the possible numbers that fit this category even though they have not been exposed.

Limiting Factors

Lest we leave the reader with the feeling that everything we choose to display turns out to be an event, we would like to present some experimental "failures" that have taught us something more about the nature of events. Three experiments in particular are relevant: one an attempt to simplify the experiment, one an attempt to capture a "formless invariant" (Gibson, 1966), and one an adventure in the undergraduate's understanding of arithmetic.

The first experiment arose when we attempted to simplify our experiment and increase its analytic power by going to cartoons of movements. Four-panel cartoon sequences depicting particular events of motion were borrowed from Robert Verbrugge's (1974) studies of metaphors of movement. These sequences showed simple events such as an object falling from a support and smashing on the floor, an object being enclosed or entrapped by a structure with a door or gate, an object running into another object, etc. Figure 1 illustrates two variants of each of two sequences.

Our first attempt was to see whether subjects would apprehend an event if they were merely shown several overlapping portions of it. This experiment was designed as a visual analogue to the experiments of Bransford and Franks (1971), who showed that subjects presented with portions of a complex sentence will falsely recognize the complete sentence. For example, subjects who hear "The rock rolled down the mountain," "The rock crushed the hut at the edge of the forest," "The hut was tiny," "The rock that rolled down the mountain crushed the tiny hut," "The tiny hut was at the edge of the forest," etc., will falsely recognize a sentence that they have never heard: "The rock that rolled down the mountain crushed the tiny hut at the edge of the forest."

In this experiment, slide sequences instead of single pictures were presented to the subjects. Four events were chosen. Subjects saw pairs of slides or triples of slides from the separate events but did not see the critical events in their entirety. Thus of the four slides in one event, a subject might see slides 1 and 2, then later slides 1, 3, and 4, then later slides 2 and 4, and still later slides 2, 3, and 4. Proper sequence order of the slides was always preserved. Subjects were tested for recogntition of doubles and triples and full sequences. Almost no subjects showed the predicted false recognition of the 1, 2, 3, 4 sequences.

Reflecting on this study, we could see a number of possible shortcomings. Some were merely experimental: each sequence was very short and the subjects might have been aware of sequence length.

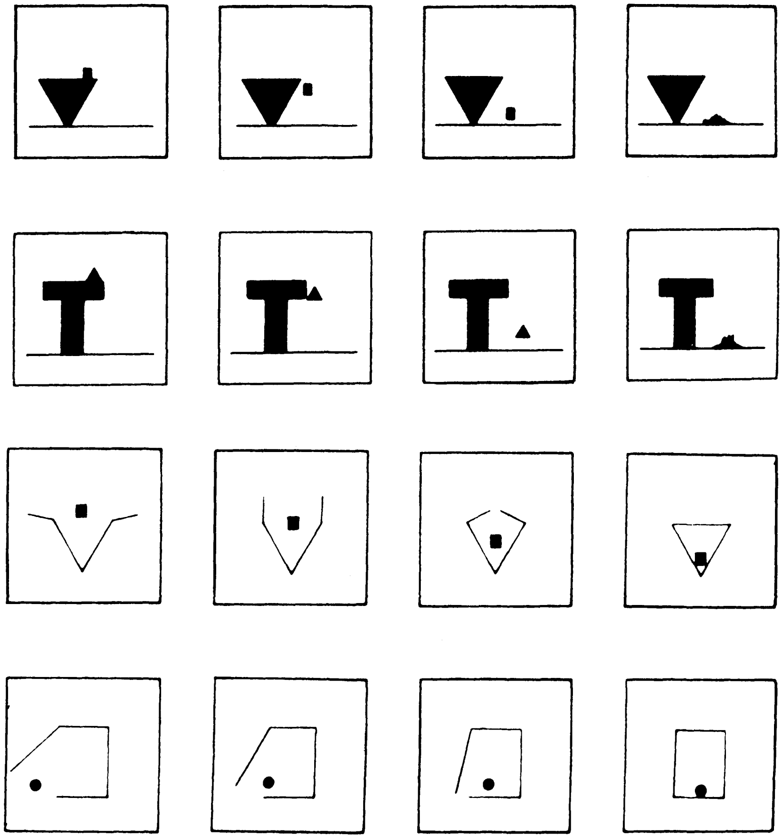

Figure 1. Examples of cartoon events of motion. Two different renderings of each of two events. From Verbrugge, 1974.

For the most part the events had very clear end points, and subjects seemed sensitive to the fact that they had not seen the full sequence from end point to end point. A major flaw, however, was conceptual; we ourselves had neglected the rate of change portrayed. Nonconsecutive samplings of the sequences of these slides automatically accelerated or decelerated the speed of the event being portrayed! Instead of providing samplings of the *same* event, we were presenting the subjects with sequences that differed in important properties.

Rather than presenting four events, we had been presenting many independent events, some of which merely happened to overlap in the objects portrayed. Consider, for example, the portrayal of the event we called *falling*, shown at the top of Figure 1. When we presented slides 1, 2, and 4, the subject saw an accelerating movement at the end. When we presented slides 1, 3, and 4, it was not the same event but rather one that decelerated at the end. We decided that our original approach had been naïve and that we had not recognized the special nature of the simplified materials we were dealing with.

We were able to bring about a successful demonstration of the abstract nature of the perceptual event by capitalizing on the fact that we had two instantiations for each of the "same" events. As is apparent in Figure 1, there are two separate renderings of each qualitatively different motion. Particulars of each rendering differ from the specific features of its counterpart; i.e., nothing is shared but the abstract nature of the event itself. In our modified study, the stimulus materials were arranged so that the subjects would become familiar with the specifics of each of two renderings for each of three events. In the course of the learning sequence one of each of the paired events was seen from beginning to end. Then in the recognition test the crucial question was whether the observers thought they had seen the other rendering of the same event from beginning to end. The results were gratifying. In general, subjects believed they had seen both sequences through completely. That is, they falsely recognized the full sequences (1, 2, 3, 4) that they had not in fact seen.

This experiment is instructive in several ways. It cautions against our identifying events with the objects or forms involved in the events or with a simplistic view of what constitutes a transformation. It argues that if the depicted transformations are different, the presence of "the same" objects may not be sufficient to define a common event. Furthermore, it argues that if the elements are familiar and the transformations are the same, the lack of common objects need not prevent two sequences as being seen as the same event. If only one illustration of a moving event is given, the knowledge of the event and the knowledge of the particulars are perfectly confounded. If there is more than one exemplar of the event, the abstract nature of the event itself can be separated from the particulars used to specify it, and the event may dominate its particular instantiations.

The second lesson we learned through experimental failure is almost the inverse of the conclusions just stated. We tried to construct an event that would be object-independent—i.e., defined solely in terms of its transformation. This experiment was arranged like the Orbiting experiment presented earlier. It differed only in one respect: the object that appeared at each position about the jam jar was different in every slide. Thus the subject saw a saltcellar in the first position, a pen in the second position, a screwdriver in the third position, an inkwell in the fourth, etc. Quite clearly, orbiting was taking place, but no specific object was doing the orbiting. Test slides used Original pictures, new objects at unseen positions of the orbit (Belonging slides), and various Control slides as before. Subjects were virtually perfect at identifying the slides they had seen before and in rejecting the new slides, whether Belonging or Control. In other words, our unnatural conditions had simply reduced the series to a set of independent events, not one overall event with object changes embedded in it. The data are entirely consistent with what one would expect from a study using unrelated slides.

The moral of this experimental failure is that abstraction may require a reasonable or coherent base. When the orbiting transformation is accompanied by object constancy, as in our earlier experiment, it is one of our strongest demonstrations of fusion. When the transformation is imposed on different, random objects, the effect disappears.

In this regard, our first two failures are perfectly consistent with one another. As we remarked earlier, transformations and objects are best viewed as two aspects of one event. When the stimulus materials are unrelated—whether by their depiction of different transformations (as in our cartoon study) or of different objects (as in the current Orbiting study)—the knowledge of the event is not separable from the knowledge of the materials per se. Only when the subject is able to perceive the invariants across instances—whether by being presented with different instantiations of the same event or by an invariant object's participation in the event—will the subject's knowledge of the event be different from his knowledge of the particulars. Only in such cases might we expect to find evidence supportive of fusion.

The final failure draws our attention to the subjects' knowledge as it contributes to the coherence phenomena we are discussing. We attempted to repeat the numerical experiment discussed above with multiples of three in place of the multiples of two used earlier. The experiment was designed in the same way and given to a comparable group of subjects. Subjects were presented with either an ordered series (3, 9, 12, 18, 27, . . . , 198) or a random arrangement of the same items (54, 174, 3, 183, etc.). As before, the test list contained Original, Belonging, and Control items. Subjects in both groups correctly recognized the Original items (74% and 71% respectively). They responded moderately well to the Belonging items (46% and 37%), but they responded almost as often to the Control items that should have been rejected (35% and 35%). In the Control items subjects seemed to be responding on the basis of numerical proximity. That is, a subject who had seen 378 on the learning list may have responded more to 376 (which is not divisible by 3) than he did to 375 (which *is* divisible by 3).

Inquiries by the experimenter revealed that over half of the subjects were not aware (did not perceive?) that the acquisition set consisted of multiples of three. What is more remarkable is that even those subjects who were aware of this responded at a high rate to the Control items that were not multiples of three. This apparently baffling finding has a simple explanation. The subjects do not know how to decide quickly and accurately whether a new number is divisible by three!

The implication of this experiment supports those of our other two failures and seems to return us to the beginning of our discussion; subjects can and do apprehend an event when there is a discernible invariant (e.g., even numbers), but they fail to do so when the invariant is not available to them or not detected by them. The invariant property in the case of multiples of three is not detected by the typical subject, and, even if he apprehends the property, he is unable to use it effectively because he cannot determine new instances readily. We assume this is the case with respect to many learned skills: i.e., the perception of special materials is vastly different for persons who have different backgrounds with respect to the materials. It is well known that the chess board is not the same to the master as it is to the novice, nor is the symphony the same to the amateur as it is to the virtuoso musician. Thus the example provided by the numerical case may be a simple illustration of the general case of the role of one's experience in the perception of coherent events.

Concluding Remarks

In this final section of the paper we shall briefly summarize the findings of our studies and comment on the wider implications we see in the research. Then we shall briefly mention issues we consider important for those who wish to model processes; and finally we shall engage in a little optimistic "handwaving" concerning the promise of our line of research.

Findings of our studies. The studies to date provide a set of demonstrations that are quite convincing at the phenomenal level. We have clearly shown that subjects can and do apprehend natural events portrayed over a series of slides. When subjects apprehend such events, they describe what they have perceived as an event (rather than as a collection of slides), and their behavior on subsequent recognition tasks is influenced in powerful ways by that perception: subjects often believe they have seen Belonging slides that, in fact, they have not seen. Subjects do not believe they have seen highly similar Control slides that violate either the static or dynamic invariants defined over the original set. That is, they detect discrepancies in station point, lighting, direction of action, event-specific transformations and relations, presence or absence of objects, etc.

We have also shown that analogous phenomena can be obtained under highly varied conditions. Most important, perhaps, is the demonstration that the observer moving over the still environment is just as much a natural event as the dynamic event presented to the stationary observer. Equally interesting are the demonstrations that artificial events (such as Orbiting) can be constructed and that such events may show greatly enhanced experimental effects. The promise here is of greater control and careful evaluation of specific aspects of the visual presentation and their contribution to the coherence of the event. We have also learned, however, that such artificial events may fail. To date they have failed in interesting ways that have directed our attention to constraints and considerations we neglected at first. Examples are the experiments with Verbrugge's cartoons, in which, by neglecting the effect of rate transformations, we created multiple separate events instead of several views of a single event; the Orbiting of Different Objects experiment, in which, by neglecting to specify object constancy, we failed to provide sufficient support for the apprehension of a unified event; and the experiment on "threeness," in which it appeared that subjects' lack of knowledge prevented them from apprehending the event we believed we were presenting.

Wider implications. At the general level these studies can be interpreted as adding to the evidence already available that shows that any set of slides implies some set of possible alternatives. In particular, we view these experiments as strong support for the position that coherent sets of slides, i.e., slides that relate to each other in some systematic fashion, specify other stimuli that may or may not be presented. Stimuli that are thoroughly specified are likely to be falsely recognized just because they fit all of the constraints or invariants of the system that has been apprehended. They may not be falsely recognized if they are only weakly specified or if there is some aspect of the particular stimulus that makes its very absence a salient feature of the presentation.

We see events as natural wholes that are, so to speak, perceived *through* the slides, rather than built up from the slides. The slides are windows through which the specifications of the event are glimpsed; they are not Tinker Toys that are used to construct some kind of event-like edifice. We believe events define their own invariants over time, and we now believe there are many more sources of coherence in real events than we had previously imagined. Perhaps any characteristic or change that can be specified in the visual array could become an invariant for some kind of event. Conversely, random variation in any aspect of an event may signify allowable variation and result in that aspect's being "ruled out" as a property relevant to that event.

We see a wealth of evidence that suggests that these experiments give rise to two kinds of knowledge: specific knowledge of what was seen in the experiment, and extensive knowledge of the event itself. In the extreme case of unrelated materials the two levels are the same; the independent slides presented are single representatives of the individual events experienced. As the event level comes to differ from the particular slide level through systematic relations (as in Esper's or Garner's experiments), through conceptual class relations (as in Goldstein and Chance's experiments), or through coherence in action patterns (as in our experiments), the two kinds of knowledge may coexist. At the extreme of relatedness, knowledge of the event itself may totally dominate knowledge of the specific inputs.

Several findings reflect the presence of these two levels of knowledge. Our subjects describe the event rather than the slides when we ask them what they have seen. When they apprehend the events, they show the false recognition of Belonging slides; when they do not apprehend events, they do not. Frequency effects play little or no role in the recognition data when coherent events are perceived, but they are predominant when one deals with unrelated events.

We believe we have shown that information is specified over time and that we have developed a technique for assessing the importance of the time course of the display in studying the event. The difference in natural- versus random-order displays offers us an opportunity to evaluate the contribution of time and rate information to the quality of an event. Some of the events we have studied so far (e.g., the Tea Sequence) seem to be so well specified or so constrained that random presentation does not prevent their veridical perception. Others are much less constrained (the Telephone Sequence) or are much more time-dependent (the Verbrugge cartoons) and lose their single-event quality under randomization.

Finally, we have reaffirmed the contribution of the subject's knowledge to the nature of the perceptual event. Our subjects' ability to pick up "twoness" but not "threeness" is a "poor man's demonstration" of the elegant work that has been done by others on the perceptual abilities of chess experts (e.g., Chase & Simon, 1973) and other specialists.

Implications for modeling processes. We attempt to study the perception of events and the characteristics of stimulus presentations that give rise to event perception. We are not trying to explain how the subjects apprehend these events; we are simply trying to say what supports the apprehension. We see the perception of the event as primary, for only after the event is apprehended can we appropriately analyze the "units" that contribute to its apprehension.

We regard the selection of materials to be studied as crucial and feel that the dramatic differences between the phenomena associated with unrelated pictures and the phenomena associated with related pictures are an important warning to investigators about permissible generalizations from impoverished materials and artificial laboratory settings. Furthermore, the choice of ecologically inappropriate materials will only result in the development of equally inappropriate models, since the two are complementary.

With respect to our experiments we see several thorny problems for any modeling approach that restricts itself to static displays. First, what is it that is taken from each individual slide and coded and stored? We think the evidence is fairly strong that whatever an image might be, it is not some sort of picture in the head; it is not another photographic representation. Second, what kind of device scans the stored representations? It seems to us that it must be some kind of dynamic "change detector." If there are such change detectors, perhaps they should work directly on the world, rather than on less-rich, static representations of the world. Third, given that our subjects seem to respond to event-specific invariants, how could the device get along with anything less than universal storage of all possible aspects of every visual representation? How would such a device discard the random changes from slide to slide but detect and represent as crucial the dynamic invariants? In brief, even in our simple experiments we see a set of challenges that will tax the ingenuity of model builders for some time to come.

We, along with the model builders, believe that the human being is a marvelous device, shaped by millions of years of evolution and millions of experiences to pick up the qualities of events. But we do not need to wait for a solution of the analytic processing dilemmas before we start to work on understanding event perception. We can simply move in the complementary direction, away from the machinery of the organism and toward the structure of the world that is to be processed. That is, we can try to understand what is perceived and what will provide sufficient ground for a coherent perception before we try to specify how it is done.

Directions of research. Starting with the assumption that events are primary, we see a rich set of questions within reach. We believe the line of investigation that we are pursuing can be readily exploited to explore the phenomena we have already tapped. We can ask what kinds of variables serve to increase coherence in events of a given kind without expecting to obtain transituational answers. We can begin to separate and study a subject's knowledge of the stimuli presented versus his knowledge of the event specified. We can ask under what circumstances we obtain coherence (what possible slides could belong to some event) and under what circumstances we obtain fusion (the inability to reject instances that belong to an event although they were not presented as part of the event). We can evaluate the role of past experience and knowledge of the observer in contributing to the coherence of an event. And we can explore the fashion in which a dynamic event creates its own "features" through the invariants it manifests over time.

These are exciting questions that we think will furnish new stimulation for both psychologists and philosophers of all orientations. They are questions pertaining to the perception of events that can be investigated at this very moment. We hope their answers will suggest new processing metaphors and new solutions to old problems.

Notes

1. Although all the experiments described here are in the visual domain, parallel arguments can be made for natural events in other modalities. A particularly good example is the perception of vowels as treated by Shankweiler, Strange, and Verbrugge (1977). These authors argue that the information that specifies a vowel cannot be found in any temporal cross section of the acoustic signal but can only be specified over time.

2. These studies are closely related to those of Bransford and Franks (1971), Franks and Bransford (1971) and Bransford, Barclay, and Franks (1972), as well as later studies in that tradition cited by Bransford and McCarrell (1974). The Franks and Bransford (1971) studies are especially significant because they demonstrate that subjects who were shown a small series of visual patterns, all of which were derived from a single prototype, falsely recognized the prototype figure and varied their recognition ratings of other figures in terms of transformational distance from the prototype.

3. Of the four slides that were not accepted, three showed the woman returning from the kitchen. This interesting accident in the selection of Belonging slides demonstrates two levels of knowledge of the event. The subjects knew that the slides belonged in the Tea Sequence but they also knew that they had not seen them because they had *never* seen the woman return to the room. This "sub-event", nested in the overall event, had a distinctiveness of its own. Because *all* "returning" pictures had been omitted from the learning series, this was an invariant of the presentation and could characterize the learning series itself.

4. Obviously, the similarity problem is a critical one and cannot be dismissed by the results of one experiment. We have currently developed several more tests of the effect

of slide similarity as opposed to "event-defined" similarity by designing intersecting events. For example, in contrast to the *orbiting* event discussed below, a *transversing* event can be created with the saltcellar moving across the tray and over the jam jar. The orbiting and transversing events intersect in two identical pictures at the edges of the tray. If a subject sees an orbiting sequence, we expect him to reject the adjacent, highly similar, slides that belong to the transversing sequence. If the subject sees the transversing sequence, we expect him to reject the highly similar slides that are adjacent in the orbiting sequence. Pilot studies show these effects are actually observed.

References

Bransford, J. D., Barclay, J. R., & Franks, J. J. Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, 1972, *3*, 193-209.

Bransford, J. D., & Franks, J. J. The abstraction of linguistic ideas. *Cognitive Psychology*, 1971, *2*, 331-350.

Bransford, J. D., & McCarrell, N. S. A sketch of a cognitive approach to comprehension: Some thoughts about understanding what it means to comprehend. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes*. Hillsdale, N.J.: Erlbaum: 1974.

Chase, W. G., & Simon, H. A. Perception in chess. *Cognitive Psychology*, 1973, *4*, 55-81.

Esper, E. A. *Analogy and association in linguistics and psychology*. Athens, Ga.: University of Georgia Press, 1973.

Esper, E. A. A technique for the experimental investigation of associative interference in artificial linguistic material. *Language Monographs of the Linguistic Society of America*, 1925, *1*.

Foss, D. J. An analysis of learning in a miniature linguistic system. *Journal of Experimental Psychology*, 1968, *76*, 450-459.

Franks, J. J., & Bransford, J. D. Abstraction of visual patterns. *Journal of Experimental Psychology*, 1971, *90*, 165-174.

Garner, W. R. *The processing of information and structure*. Potomac, Md.: Erlbaum, 1974.

Gibson, J. J. *The senses considered as perceptual systems*. Boston: Houghton-Mifflin, 1966.

Goldstein, A. G., & Chance, J. E. Visual recognition memory for complex configurations. *Perception and Psychophysics*, 1970, *9*, 237-240.

Gough, P. B. One second of reading. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye*. Cambridge, Ma.: The M.I.T. Press, 1972.

Haber, N. R. How we remember what we see. *Scientific American*, 1970, *222*(5), 104-112.

Jenkins, J. J. Can we have a theory of meaningful memory? In R. E. Solso (Ed.), *Memory and cognition: The second Loyola symposium*. Potomac, Md.: Erlbaum, 1974, 1-20. (a)

Jenkins, J. J. Remember that old theory of memory? Well, forget it! *American Psychologist*, 1974, *29*, 785-795. (b)

Johansson, G. Visual motion perception. *Scientific American*, 1975, *232*(6), 76-88.

Kavanagh, J. F., & Mattingly, I. B. *Language by ear and by eye*. Cambridge, Ma.: The M.I.T. Press, 1972.

Kraft, R. N., & Jenkins, J. J. Memory for lateral orientation of slides in picture stories. *Memory and Cognition*, 1977, *5*, 397-403.

LaBerge, D., & Samuels, S. J. Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 1974, *6*, 293-323.

Michotte, A. *The perception of causality*. London: Methuen & Co., Ltd., 1963.

Neisser, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.

Posner, M. I., & Keele, S. W. On the genesis of abstract ideas. *Journal of Experimental Psychology*, 1968, *77*, 353-363.

Posner, M. I., & Keele, S. W. Retention of abstract ideas. *Journal of Experimental Psychology*, 1970, *83*, 304-308.

Postman, L. Verbal learning and memory. *The Annual Review of Psychology*, 1975, *26*, 291-335.

Segal, E. M. *Stimulus perception as a function of response set*. Unpublished doctoral dissertation, University of Minnesota, 1962.

Shankweiler, D. P., Strange, W., & Verbrugge, R. R. Speech and the problem of perceptual constancy. In R. E. Shaw & J. D. Bransford, (Eds.), *Perceiving, acting, and knowing. Toward an ecological psychology*, Hillsdale, N.J.: Erlbaum, 1977.

Shaw, R. E., & Wilson, B. E. Conceptual knowledge: How we know what we know. In D. Klahr (Ed.), *Cognition and instruction*. Hillsdale, N.J.: Erlbaum, 1976.

Shaw, R. E., & Pittenger, J. B. On perceiving change. In H. L. Pick & E. Saltzman (Eds.), *Modes of perceiving and processing information*. Hillsdale, N.J.: Erlbaum, in press.

Shepard, R. N. Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behavior*, 1967, *6*, 156-163.

Standing, L. Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology*, 1973, *25*, 207-222.

Standing, L., Conezio, J., & Haber, R. N. Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science*, 1970, *19*, 73-74.

Strange, W., Keeney, T., Kessel, F. S., & Jenkins, J. J. Abstraction over time of prototypes from distortions of random dot patterns. *Journal of Experimental Psychology*, 1970, *83*, 508-510.

Verbrugge, R. R. *The comprehension of analogy*. Unpublished doctoral dissertation, University of Minnesota, 1974.