Memory & Cognition
2000, 28 (7), 1213-1230

Serial causation: Occasion setting in a causal induction task

MICHAEL E. YOUNG, JANELLE L. JOHNSON, and EDWARD A. WASSERMAN
University of Iowa, Iowa City, Iowa

The temporal relations among candidate causes were studied in a causal induction task using a design that is known to produce occasion setting in animal learning preparations. For some subset of the observations, one event, the occasion setter, was accompanied by another event, the conditional cause; for another subset of the observations, the conditional cause occurred alone. The efficacy of the conditional cause depended on whether it was or was not accompanied by the occasion setter. Participants used the occasion setter to modulate their effect expectancy to the conditional cause when the events were presented serially, but not simultaneously. Current causal induction models are unable to account for the full range of effects that we observed; the relative roles of time, attention, and cue distinctiveness are discussed.

> These principles of association are reduced to three, viz., "resemblance"... "contiguity"... [and] "causation"... they are *to us* the cement of the universe, and all the operations of the mind must, in a great measure, depend on them. (Hume, 1748/1955, p. 198)

Throughout recorded history, people have been engaged in the search for causes. Medical scientists explore the human body searching for the causes of a variety of diseases; they comb through blood samples, urine specimens, and DNA in their search to improve human health. Business managers examine the circumstances affecting their companies' profitability, investigating factors that improve productivity, decrease cost, and increase market share. Even young children engage in this pervasive exploration of the causal structure of the environment; they discover the actions that lead to parental approval or disapproval of a request, those that produce punishment or reward, and those that turn the television on or off. As researchers studying causal induction, we are also occupied with this enterprise, but at a higher level; we are searching for the factors that affect the induction of causality.

Centuries of study since the famous speculations of David Hume have identified the primary factors that influence the likelihood of an observer's identifying the causes of an event (Young, 1995). We refer to these factors as *causal heuristics*, to encourage their perception as indicators of the possible presence of a cause; some of these heuristics may be violated while still allowing a causal induction. A list of the most prominent heuristics usually includes temporal priority (a cause should precede its effect), temporal contiguity (causes will be found among those events that immediately precede the effect), spatial contiguity (causes should occur near their effects), and covariation or contingency (the effect should consistently follow its cause).

In studies of causal induction, several trials are presented that correspond to a sampling of event–effect observations. The inducer must determine which of the preceding events reliably produce(s) the effect. Prior research has centered on cause–effect contingency differences or on an observer's prior expectations about cause–effect relations, whether these expectations are preexperimental (Alloy & Tabachnik, 1984; Cheng, 1993; Waldmann & Holyoak, 1992) or are produced within the experimental session (e.g., in demonstrations of blocking; Dickinson & Burke, 1996; Shanks, 1985; Wasserman & Berglan, 1998; Williams, Sagness, & McPhee, 1994). In most of these studies, the participant is told that for any single event–effect observation, all of the causal candidates appearing on a trial were immediately antecedent to the observed effect (Shanks, Pearson, & Dickinson, 1989; Wasserman & Neunaber, 1986). Researchers have demonstrated that the contiguity between a response and its consequence is an important factor in learning a response–outcome relation (Shanks et al., 1989; Wasserman & Neunaber, 1986), but little is known about how temporal relations among multiple causal candidates affect the induction process (Young, 1995).

SERIAL CAUSATION

In the present set of experiments, we used the causal induction paradigm to investigate the role of temporal contiguity among multiple causes. These experiments begin an exploration of *serial causation*, entailing situa-

The data in Experiment 1 were presented at the 1997 meeting of the Psychonomic Society, and the data in Experiments 1 and 2 were presented at the 1999 Special Interest Meeting on Cue Competition in Associative Learning, Ardennes, Belgium. The authors thank Nicole Hill for assistance in running Experiment 1, Ralph Miller for suggesting the extinction test, and David Shanks, Bill Whitlow, and anonymous reviewers for comments on earlier versions of this manuscript. Correspondence should be addressed to M. E. Young, 271F Life Sciences II, Department of Psychology, Southern Illinois University, Carbondale, IL 62901 (e-mail: meyoung@siu.edu).

tions in which a series of candidate causes is followed by an effect. In serial causation, the causal candidates occur at different times relative to the effect, thus placing some of the candidates at a temporal contiguity disadvantage relative to others. Our work focuses on a particular type of serial causation, *occasion setting*, in which the two events of interest are the occasion setter (OS) and the conditional cause (CC). For some subset of the observations, the OS is followed by the CC; for another subset of the observations, the CC occurs alone. The efficacy of the CC depends on whether or not it was preceded by the OS. The second event is called the conditional cause because its effects are conditional on the occurrence of another event (the OS) that is said to "set the occasion" for its operability.

As a concrete example, consider a vending machine. On its own, pressing the button for your favorite beverage does not produce the drink. When another event precedes the buttonpress (putting sufficient money into the machine), the efficacy of the buttonpress is changed such that it then produces the desired outcome. The insertion of money sets the occasion for the buttonpress to produce the beverage.

The temporal order of the two events in an occasion setting design is important. If the OS were not at a temporal disadvantage, then it is more parsimonious for an observer simply to associate the OS directly with the effect and to ignore the CC (assuming that the observer had no experience with the OS in the absence of the CC). The OS perfectly predicts the effect, whereas the CC is only occasionally followed by the effect—the OS has a contingency advantage. Although this elemental learning strategy would be more parsimonious, it is possible that an observer may use the entire OS→CC compound to predict the effect. Attending to the entire constellation of cues represents a *configural learning strategy* (Shanks, Darby, & Charles, 1998; Williams et al., 1994). This strategy represents the learning of an interaction between causes.

By placing the OS at a temporal contiguity disadvantage in an occasion setting paradigm, the OS's contingency advantage has to be weighed against its contiguity disadvantage. This conflict could either prompt the selection of only one of the two events as the candidate cause, or it could prompt the formation of a unique relationship between the candidate causes so that both play a role in anticipating the effect. The OS's temporal contiguity disadvantage may increase the likelihood that this earlier event is considered a "mere condition" for the operability of the later event. Classifying the OS as a condition would obviate its consideration as a cause of the effect. Indeed, for an event to be classified as an OS (and not merely a cause in its own right), subsequent tests should reveal that the OS has little efficacy of its own when presented independently of the CC (i.e., one should not believe that inserting money into a vending machine will be sufficient to produce the beverage).

Occasion setting is thus an interaction between candidate causes. In occasion setting, the two events together have an efficacy that differs from the efficacy of each event considered alone, thus making occasion setting similar to configural learning tasks, in which the constellation of two or more events produces effects that are not reducible to those of its component events (e.g., as observed in positive and negative patterning; Shanks & Darby, 1998; Young, Wasserman, Johnson, & Jones, 2000).

OCCASION SETTING IN NONHUMAN ANIMALS

Holland (1986) pioneered the empirical investigation of occasion setting in studies of animal learning. He has identified a range of situations under which animals will use one event to modulate the effect expectation following another event.

In Holland's occasion setting studies involving serially presented events (Holland, 1983, 1989a, 1989b), the key observations are (1) high levels of responding after the OS→CC compound, (2) minimal responding following each event when it is presented alone, (3) good transfer of occasion setting when the OS is paired with another event that itself served as the CC in a separate occasion setting relation involving the same effect, but (4) poor transfer of occasion setting when the OS is paired with another event that has not served as the CC in a separate occasion setting relation. The observation of minimal responding to the OS alone is important in demonstrating that the higher level of responding on OS→CC trials than on CC-alone trials is not due to the presence of a direct OS–effect association. The specificity of transfer is important in distinguishing occasion setting from a more general modulatory mechanism in which the response threshold for any cue is modulated by the presence of the OS.

In occasion setting studies, the OS usually precedes the CC in time (cf. Holland, 1989c). Holland and colleagues have performed a number of studies (involving rats) in which performance following the *serial* presentation of the events is contrasted with performance following the *simultaneous* presentation of events (e.g., Holland, 1986, 1991a; Holland & Reeve, 1991). These researchers have found that there is greater elemental learning when the events are presented simultaneously (a direct OS–effect relation is learned) and more modulatory learning when the events are presented serially (the OS modulates responding to the CC).

The distinction between elemental and modulatory learning is documented in two ways. First, when learning is elemental, later extinction of the OS–effect relation should have a large impact on the efficacy of the OS–CC compound. This extinction effect is produced because an elemental strategy results in a direct association between the OS and the effect during the acquisition stage (the OS is a perfect predictor and will overshadow the CC); the compound's efficacy is merely the sum of the efficacy of each of its elements (OS = 100% efficacy, CC = 0% efficacy). When the efficacy of the OS is extin-

guished (driven to 0%), the efficacy of the compound is also extinguished. An elemental model of learning (e.g., Rescorla & Wagner, 1972) anticipates this extinction effect. When learning is modulatory, however, later extinction has a relatively small impact on the efficacy of the OS–CC compound. The efficacy of the compound is determined not by a direct OS–effect association but by the OS modulating the efficacy of the CC; extinguishing the direct OS–effect association would have no effect on the OS's modulatory properties.

Second, when learning is elemental, the efficacy of the OS transfers well when it is paired with events other than the CC, whereas when learning is modulatory, the efficacy of the OS transfers well only to a CC trained in another occasion setting relation. In elemental learning, the compound's efficacy is solely determined by the efficacy of the OS; thus, pairing it with other (noninhibitory) events will have no effect on the OS's efficacy. In modulatory learning, however, the OS's modulation is presumed to be a function of the history of the event being modulated; pairing the OS with an event that has no history of being modulated will undermine the efficacy of the OS.

Although nonhuman animals tend to learn elementally when the events are presented simultaneously, there is reason to believe that people may not. Williams et al. (1994) have shown that people tend to respond configurally to simultaneous compounds. Therefore, human observers may not show the differential effect of OS extinction under simultaneous and serial presentation that has been observed in nonhuman animals. If the OS–CC compound acquires its own efficacy, then the compound should be affected by extinction only to the extent that the extinction of the OS generalizes to the OS–CC compound. Thus, the retention of the OS–CC compound's efficacy after extinction could be the result of configural learning and thus observed after either simultaneous or serial presentation.

The differential transfer of occasion setting to new targets under simultaneous and serial presentation also may or may not be revealed in human analogs. Given that differential transfer might be the product of configural learning under both serial and simultaneous presentation, we might see equivalent transfer under both conditions; transfer would be determined by the surface similarity of the training compound to the novel, transfer compound. However, if the differential transfer were due to some other factor, then the training history of a new target might be more relevant or salient when the events are presented serially, but not when they are presented simultaneously. Candidates for these other factors will be considered in the General Discussion section.

Both OS extinction and differential transfer were examined in Experiment 1. The effect of extinction was equivalent under serial and simultaneous presentation of a compound's events, but transfer to an event with an occasion setting history was stronger when the events were serial than when they were simultaneous.

In Experiment 1, we also observed that participants learned the efficacy of the OS–CC compound more easily when the events were presented serially. This finding was counterintuitive given that configuring seemed more likely when the events were presented simultaneously. Experiments 2A and 2B focused on this important difference between serial and simultaneous presentation. In Experiment 2A, we taught participants an ambiguous occasion setting relation, to rule out a general modulatory explanation of our results (Rescorla, 1985) and to determine whether the OS's absence at the time of CC's occurrence (when presented in a serial relation) was important to the superior learning of a compound's efficacy. Finally, in Experiment 2B, we used a series of controls that ruled out the effect of other temporal variables that might have produced the slower learning of a compound's efficacy under conditions of simultaneous presentation.

EXPERIMENT 1

In Experiment 1, we examined some characteristics of the occasion setting properties of the first event in a serial pair of putative causes; the participant's task was to determine the causes of a single observed effect—a chemical reaction. The chemistry domain was chosen for two reasons: (1) people should readily apprehend that interactions between chemicals are possible and (2) the timing of events on the computer screen (on the order of seconds) maps to the timing of events in the domain of chemistry.

After initial training in which the OS→CC compound (to avoid verbosity, we will henceforth adopt the temporal "→" notation for both serial and simultaneous OS– CC presentation) was followed by the effect but the CC alone was not, the participant received additional trials in which the OS alone was explicitly given without being followed by the effect (while receiving no feedback regarding the continued efficacy of the OS→CC compound). If a direct OS–effect association largely determined the effect expectancy following the OS→CC compound (as anticipated by elemental learning), then the extinction of that association should abolish any expectancy of the effect following the compound. If instead the OS served as a modulator of the CC's efficacy (as anticipated in occasion setting) or if there were configural learning, then extinction of the OS should not abolish the compound's efficacy.

We also examined the transfer of occasion setting by concurrently training participants with a second occasion setting relation and later testing participants' effect expectancy when the OS from one relation was paired with the CC from the other. Furthermore, we explored the generality of OS transfer by examining transfer to an event whose relation with the effect had been trained and extinguished (cf. Holland, 1991a); this training history creates an event that, like the CC, has been associated with the effect on half of the trials but that has no direct predictive efficacy of its own by the end of the experiment. The transfer of occasion setting properties to another event may not be

specific to events trained as CCs in other occasion setting relations. Although studies of OS specificity in animal learning have demonstrated much weaker transfer to events not trained as CCs, we cannot assume that the same will be true in human causal induction. Given that the events used in our task were chosen from the same category of causes (chemicals), participants may transfer the training history of an event to other events with similar training histories. Some histories may, however, transfer more readily than others; rats, for example, are more likely to transfer the properties of one OS to another (or, conversely, from one CC to another; Swartzentruber, 1998) than from one simple predictor to another (Holland, 1989a).

The most interesting aspect of a demonstration of occasion setting is the potential effect of time on causal induction; ordered cues may produce fundamentally different patterns of behavior than do simultaneous cues. To document the effect of time on learning involving compounds, animal cognition researchers who study occasion setting have used controls in which the predictors are presented simultaneously (i.e., the putative OS and the CC are contemporaneous; Holland, 1989a, 1991a). When the OS and the CC are presented serially (an occasion setting discrimination), Holland has observed good (though incomplete) transfer of the OS when it was paired with the CC from another occasion setting pair but weaker transfer of the OS when it was paired with a trained and extinguished event. When the OS and CC are presented simultaneously, Holland has observed good (although incomplete) transfer of the OS when it was paired with the CC from another occasion setting pair and when it was paired with a trained and extinguished event. In Experiment 1, we examined the effects of manipulating the temporal relation between the predictors, in order to determine the effect of time on the transfer specificity of an OS by presenting the OS→CC compound either serially or simultaneously.

To assess the changing efficacy of cues as learning progressed, we measured participants' predictions throughout the learning phase. However, for some of the cues (e.g., the OS alone), we did not want to provide explicit feedback about the occurrence of the effect. For these cues, we used "test trials," in which the cues were presented and the predictions were solicited, but no feedback was provided. We confined the use of these test trials to the end of the experiment; their absence during learning prevented the test trials from affecting the learning process.

Method

Participants. Forty students enrolled in an introductory psychology course at the University of Iowa served as voluntary participants. They received course credit for their participation.

Materials. The experiment was programmed using PsyScope v. 1.0.2 (Cohen, MacWhinney, Flatt, & Provost, 1993) on four Power Macintosh 7100/80 computers.

Six different fictional chemicals—Adelpine, Furval, Glexus, Morphid, Rezitak, and Sopatonin—were used as cues to aid the participants in their prediction of positive and negative outcomes. The cues were presented in 18-point, bold, black New York type in the center of the screen. Outcomes ("REACTION" or "no reaction") were presented directly below the cues in 18-point, bold, red New York type. The reaction outcomes were emphasized through capitalization to approximate the conditions present in a conditioning experiment where outcome occurrence is more salient than outcome nonoccurrence (e.g., food vs. no food, or shock vs. no shock). This asymmetry is important in distinguishing causal induction from categorization (Gilovich, 1991; Nisbett & Ross, 1980).

Procedure. Between 1 and 4 participants were studied concurrently on four identically configured computer workstations. Each participant sat in front of a workstation and listened to the experimenter's recitation of a series of general instructions. The participants then read the following instructions on the computer monitor:

> You are about to be placed in the role of a chemist. One task of a chemist is to determine how and when certain chemicals produce chemical reactions. Here, you will examine a series of experimental results recreated for your convenience. You will be asked to give your best guess as to whether a particular chemical, or series of chemicals, will produce the reaction of interest. The reaction of interest is a color change, and it may take time for the reaction to occur. The time of presentation of the chemicals reflects when the chemicals were added in the real experiment. You will then be provided with the actual outcome for the experiment. This information will assist you in making subsequent predictions. Note: Only chemicals listed on the screen were added; if a chemical is not mentioned, then it was not used in the experiment.
>
> Your goal is to accurately predict whether chemicals will produce a color change reaction. You will observe a series of screens. For each experiment, one or more chemicals will be displayed on the upper half of the screen. When you are prompted with the word "RESPOND," please press <1> if you believe there will be a reaction, and <3> if you believe there will be no reaction. Please answer **as quickly as possible while still being accurate**. A running score of your accuracy will be displayed on the bottom of the screen in terms of the percentage of trials on which you made the correct prediction.
>
> After your prediction, the actual outcome will be displayed in red on the center of the computer monitor. Press the <2> to move on to the next experiment's data. Obviously, at first you will have to guess the outcome because you will not know anything about the possible causes of the reaction; but, as you see more cases, you will begin to learn which chemicals cause the reaction and whether the time or order of addition affects the occurrence of the reaction.
>
> After your prediction, the actual outcome will be displayed and you will hear a tone. You will hear a high pitched "beep" tone if you correctly predicted the outcome and a lower pitched "buzz" tone if your prediction was incorrect. Press the <2> key to move on to the next set of data. Obviously, at first you will have to guess the outcome because you will not know anything about the possible causes of a reaction; but, as you see more cases, you will begin to learn which chemicals cause a positive reaction, whether the time or order of their addition affects the reaction, and whether combinations of certain chemicals are especially effective or ineffective. Towards the end of the study, outcome and auditory feedback will cease while we test your learning in more depth - just keep making predictions to the best of your ability.
>
> NOTE: It is very important for you to identify potential positive reactions, therefore be especially diligent to identify them while avoiding false positives as much as possible.

The participants pressed the space bar to confirm their understanding of the instructions and to begin the experiment. Each chemical was presented in the center of the screen. Only those chemicals present on the trial were displayed. Each chemical was programmed to appear on the screen at a specified time and duration. After the trial's last chemical appeared for its full duration, the word "RESPOND" was displayed (just below the middle of the screen) to prompt the participant to make a prediction. The last chemicals

displayed remained on the screen until a response was made. The program did not register any responses made before presentation of the prompt.

The participants pressed the "1" key on the numeric keypad to signify their expectation of a chemical reaction and the "3" key for their expectation of no chemical reaction. Immediately following the keypress, the actual outcome was presented directly below the cue. A reminder to press the "2" key to begin the next trial was displayed. Following the "2" response, the percent correct score was updated (except on test trials, see below) and was followed by an intertrial interval of 2,000 msec, after which the chemical(s) for the next trial was (were) displayed.

The percentages of trials in which a chemical reaction followed the cues are shown in Table 1. The prediction portion of the experiment comprised three phases—the initial phase, the extinction phase, and the testing phase—with no pause between the phases; the change between phases was not signaled in any way (although the change to the testing phase was obvious, given that none of the predictions provided feedback). In the initial phase, we included training to produce two distinct occasion setting relations: $OS_1 \rightarrow CC_1$ and $OS_2 \rightarrow CC_2$. In the extinction phase, training was continued with the $OS_2 \rightarrow CC_2$ relation, but there was no further training with the $OS_1 \rightarrow CC_1$ relation; instead, any residual association between OS_1 and the effect was extinguished by presenting trials in which OS_1 was not followed by the chemical reaction. An additional cue, EXT, was trained as a cause of the effect in the initial phase, but the relation between EXT and the effect was extinguished during the extinction phase.

Henceforth, the use of a superscript will indicate the percentage of trials on which a cue or cue compound was followed by the effect; an asterisk (*) will be used to designate a cue or cue compound that appeared as a test trial (i.e., with no explicit feedback).

The initial phase consisted of three randomized blocks comprising feedback trials. The participants observed 24 trials in each block that included 16 occasion setting trials (4 $OS_1 \rightarrow CC_1^{100}$, 4 CC_1^0, 4 $OS_2 \rightarrow CC_2^{100}$, and 4 CC_2^0), 4 compound control trials ($X \rightarrow Y^0$) included so that the participants would not conclude that all compounds were efficacious, and 4 trials with the cue that would be later extinguished (EXT^{100}), for a total of 72 trials.

The extinction phase consisted of three randomized blocks comprising feedback trials. The participants observed 24 trials in each block that included 8 trials with the first occasion setting relation (4 OS_1^0 and 4 CC_1^0) in which OS_1 was extinguished and no information was provided regarding its continued efficacy within the compound, 8 occasion setting trials with the second occasion setting relation (4 $OS_2 \rightarrow CC_2^{100}$ and 4 CC_2^0), 4 compound control trials ($X \rightarrow Y^0$), and 4 trials with the extinguished cue (EXT^0), for a total of 72 trials.

The testing phase consisted of two randomized blocks comprising test trials only. The participants observed 12 trials in each block that included each of the cues or cue compounds experienced in training (OS_1^*, CC_1^*, CC_2^*, $OS_1 \rightarrow CC_1^*$, $OS_2 \rightarrow CC_2^*$, $X \rightarrow Y^*$, and EXT*), the never-before-presented OS_2 by itself, and a series of transfer tests ($OS_1 \rightarrow CC_2^*$, $OS_1 \rightarrow EXT^*$, $OS_2 \rightarrow CC_1^*$, and $OS_2 \rightarrow EXT^*$). Young and Wasserman (1998) observed no significant difference in the chemical reaction predictions for the OS when it was presented at two different times relative to the prompt; therefore, here we tested the OS only at the OS temporal position (i.e., the interval between the prompt and each OS when it was presented alone was identical to that between the prompt and each OS when it was presented in its occasion setting compound).

The participants were randomly assigned to one of two conditions: simultaneous or serial. In both conditions, the cues were 1,000 msec (plus the response time, or RT) in duration. In the simultaneous condition, all cues presented within a trial started and ended together with their onset occurring 1,000 msec before the "RESPOND" prompt and terminating when a response was registered. In the serial condition, cues that occurred first in a serial compound (OS_1, OS_2, or X) were presented for 1,000 msec and were followed by the second cue (CC_1, CC_2, Y, or EXT, all with a duration of 1,000 msec plus the RT) after a delay of 4,000 msec. The sequence of events was thus: first event→4,000 msec→second event, where each event terminated before the next event began.[1] Therefore, trials in the simultaneous condition were 1,000 msec in duration (plus the RT), whereas trials in the serial condition were 6,000 msec in duration (plus the RT).

Assignment of cue identity to cue type was done through a partial 7 × 7 Latin square: 7 cues (Adelpine, Bucagon, Furval, Glexus, Morphid, Rezitak, and Sopatonin) × 7 roles (OS_1, CC_1, OS_2, CC_2, X, Y, and EXT). Each of the four workstations used a consistent assignment of cue to role for a given condition. Condition was manipulated between subjects.

Table 1
Percentage of Trials in Which a Chemical Reaction Followed the Addition of Chemical(s) in Experiment 1

Chemical	Percentage (Number) of Trials With a Chemical Reaction: Initial Phase	Extinction Phase
Occasion Setting Compound 1		
$OS_1 \rightarrow CC_1$	100 (12)	—
CC_1	0 (12)	0 (12)
OS_1	—	0 (12)
Occasion Setting Compound 2		
$OS_2 \rightarrow CC_2$	100 (12)	100 (12)
CC_2	0 (12)	0 (12)
Controls		
EXT	100 (12)	0 (12)
X→Y control	0 (12)	0 (12)

Note—The number in parentheses indicates the number of trials of that type within the phase. "—" indicates that these types of trials did not occur in this phase of the experiment. The participants in the serial condition experienced the cue preceding the arrow before the cue following the arrow; the participants in the simultaneous condition experienced both of these cues simultaneously.

Results

Any participant who failed to achieve at least 50% accuracy during the final block of training (Block 6) for every one of the cues or cue compounds was excluded from further study. This criterion resulted in the elimination of only 4 of the 40 participants (2 in the simultaneous condition and 2 in the serial condition).

Assessing the participants' learning of experienced contingencies. The response profile produced during the initial and extinction phases of the experiment is shown in Figure 1. The participants clearly learned the cue contingencies (those cues that were followed by the effect produced more reaction predictions than those cues that were not), and accuracy improved after further training (scores were closer to the true values in the extinction phase than they were in the initial phase).

To determine the degree of final learning of the assigned contingencies and to assess any differences in learning between the conditions, we examined the percentage of effect predictions during the testing phase for each individual stimulus and stimulus compound that appeared in the

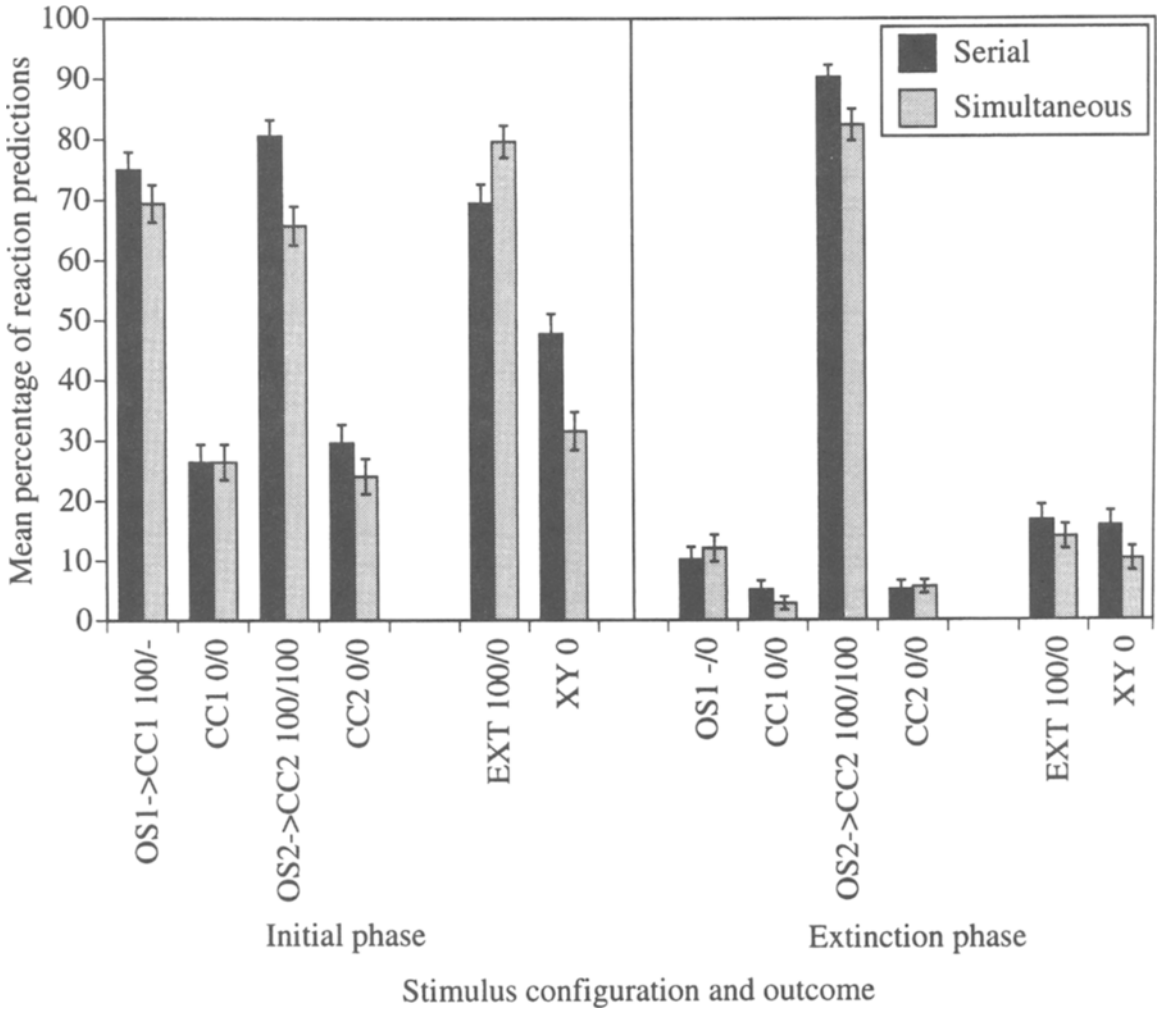

Figure 1. Mean percentage of reaction predictions for each stimulus or stimulus compound during the training phase of Experiment 1. The *a/b* notation used in labeling values on the *x*-axis indicates the type of feedback that followed the cue during the initial phase (*a*) and the extinction phase (*b*).

extinction phase of training (OS_1, CC_1, CC_2, $OS_2 \rightarrow CC_2$, X→Y, and EXT); because the causal efficacy of those stimuli that appeared only in the initial phase may have changed during the extinction phase, they were not included as measures of the participants' final learning. The results are included in Figure 2. Learning was very good across both phases in the serial and simultaneous conditions (accuracy was high in both; *M*s = 92.6% and 91.4% correct for serial and simultaneous conditions, respectively), closely paralleling the prevailing contingencies.

Although only a subset of the stimuli that appeared in the testing phase was of interest for the analysis of differences in learning, we included the entire set of stimuli in our analysis of variance (ANOVA) of these results with the intent of using planned comparisons for specific purposes. Thus, we will first present the overall ANOVA and follow it with the planned comparisons that were relevant in the investigation of any between-condition differences in the participants' learning of the stimulus–effect relations. A later analysis will use a different set of planned comparisons between the conditions to consider the participants' responses during the transfer tests.

The reaction predictions were submitted to a repeated measures, factorial ANOVA of condition (serial vs. simultaneous) and stimulus (OS_1, CC_1, OS_2, CC_2, $OS_1 \rightarrow CC_1$, $OS_2 \rightarrow CC_2$, $OS_1 \rightarrow CC_2$, $OS_2 \rightarrow CC_1$, $OS_1 \rightarrow$EXT, $OS_2 \rightarrow$ EXT, X→Y, and EXT). The main effect for condition was not statistically significant [$F(1,34) = 3.89$, $MS_e = 0.365$, $p < .10$]. There was a statistically significant main effect for stimulus [$F(11,374) = 35.65$, $MS_e = 0.194$, $p < .0001$], and there was a statistically significant condition × stimulus interaction [$F(11,374) = 2.11$, $MS_e = 0.194$, $p < .05$]. For the purposes of determining the presence of any between-condition differences in asymptotic learning of the stimulus–effect relations, we performed planned orthogonal comparisons ($\alpha = .05$) between conditions for the OS_1, CC_1, CC_2, $OS_2 \rightarrow CC_2$, X→Y, and EXT cues. None of these comparisons were statistically significant

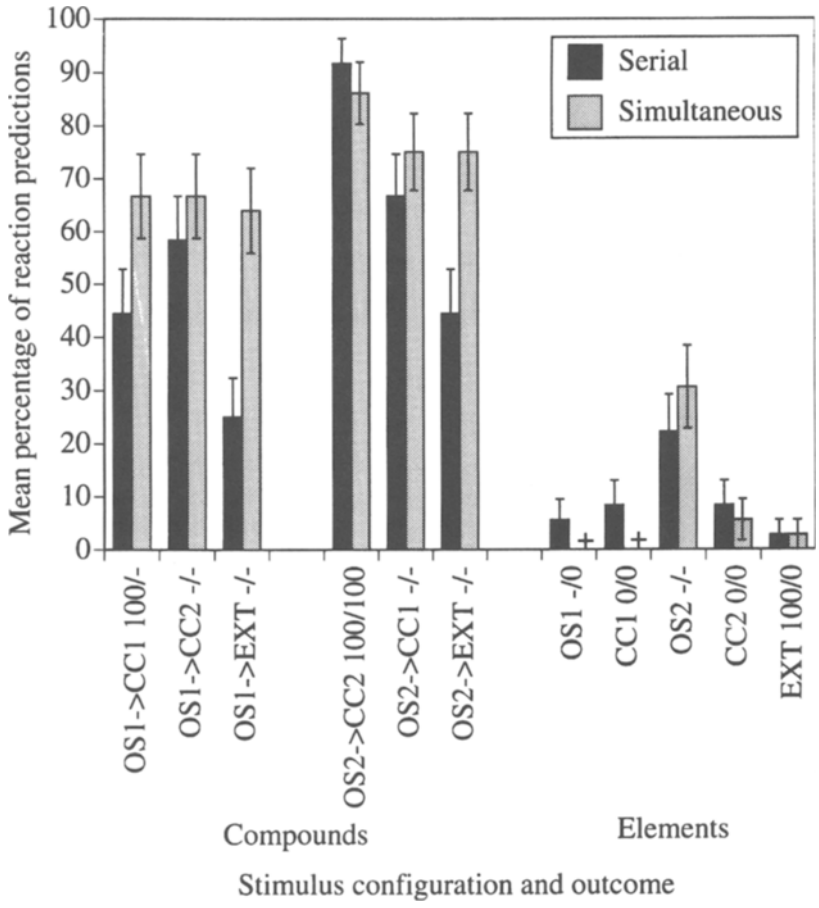

Figure 2. Mean percentage of reaction predictions for each stimulus or stimulus compound during the testing phase of Experiment 1. Note—A plus sign ("+") was used to designate those values with a mean of 0.0 and a standard error of 0.0.

(all $ps > .10$). These results suggest that the participants learned the task as directed and that there was no evidence of differential asymptotic learning between the serial and simultaneous conditions, notwithstanding the significant condition × stimulus interaction.

We found that the extinction of OS_1 resulted in it having a significantly lower predictive efficacy than its nonextinguished OS_2 counterpart, with OS_1 having a lower judged predictive efficacy ($M = 2.8\%$) than OS_2 ($M = 26.4\%$), with no statistically significant difference between the conditions (shown in the right half of Figure 2). These means demonstrate that OS→CC training may produce a modest OS–effect association unless that association is extinguished.

Transfer tests. Experiment 1 was specifically designed to determine whether there were any behavioral differences between serial and simultaneous presentation of a compound's events. Of critical interest was the effect expectancy when the OS was paired with a new event that either functioned as a CC in another OS→CC relationship or was trained as a cause in its own right and subsequently extinguished. Differential transfer between the conditions is clearly shown in the left half of Figure 2. The mean percentage of reaction predictions was greater when the OS was paired with the CC of a different occasion setting relationship ($OS_1 \rightarrow CC_2$, and $OS_2 \rightarrow CC_1$) than when it was paired with the corresponding trained and extinguished EXT cue ($OS_1 \rightarrow EXT$, and $OS_2 \rightarrow EXT$, respectively) only in the serial condition; in the simultaneous condition, high and equivalent transfer was seen in each case.

Following the repeated measures ANOVA reported earlier, we conducted a series of confirmatory planned orthogonal comparisons in which the effect predictions following $OS_1 \rightarrow CC_2$ were compared with those following $OS_1 \rightarrow EXT$ and the effect predictions following $OS_2 \rightarrow CC_1$ were compared with those following $OS_2 \rightarrow EXT$. In the serial condition, the percentage of effect predictions following $OS_1 \rightarrow CC_2$ ($M = 58.3\%$) was significantly higher than that following $OS_1 \rightarrow EXT$ ($M = 25.0\%$) [$t(187) = 3.21, p < .01$], and the effect expectancy following $OS_2 \rightarrow CC_1$ ($M = 66.7\%$) was significantly higher than that following $OS_2 \rightarrow EXT$ ($M = 44.4\%$) [$t(187) = 2.14$, $p < .05$]. In the simultaneous condition, the percentage of

effect predictions following $OS_1 \rightarrow CC_2$ ($M = 66.7\%$) was not significantly different from that following $OS_1 \rightarrow EXT$ ($M = 63.9\%$) [$t(187) = 0.27, p > .10$], and the effect expectancy following $OS_2 \rightarrow CC_1$ ($M = 75.0\%$) was identical to that following $OS_2 \rightarrow EXT$ ($M = 75.0\%$) ($t = 0.00$).

We were also interested in the effect of OS_1 extinction on the $OS_1 \rightarrow CC_1$ relationship. In the serial condition, the effect expectancy following $OS_1 \rightarrow CC_1$ during the testing phase ($M = 44.4\%$) was significantly lower than that following $OS_1 \rightarrow CC_1$ during the last block of the initial phase ($M = 79.2\%$) ($t = -5.72$, $p < .0001$) (last block data are not shown in any figure). In the simultaneous condition, the effect expectancy following $OS_1 \rightarrow CC_1$ ($M = 66.7\%$) was significantly lower than that following $OS_1 \rightarrow CC_1$ during the last block of the initial phase ($M = 80.6\%$) ($t = -2.29$, $p < .05$) (also not shown).

Differential accuracy for a compound and its elements. Thus far, our data have revealed that none of the participants demonstrated purely elemental learning; in both the serial condition and the simultaneous condition, extinction of the OS decreased the judged efficacy of the OS→CC compound but did not eradicate it. This finding suggests that the occasion setting compound has an efficacy above and beyond that predicted by the mere summation of its elements under both serial and simultaneous presentation. To determine whether there were any indications of stronger elemental learning in the simultaneous condition than in the serial condition, we examined the relative prediction accuracy on occasion setting compound trials (involving the OS and CC) and on element trials (involving the CC alone) in the two conditions throughout both training phases (the previous analyses examined only asymptotic performance). We reasoned that an observer might be predisposed to focus on one element of a simultaneous compound to the detriment of the other. This focused attention could create problems; focusing on the wrong element (e.g., the CC) would retard learning when its efficacy is determined by the other element (e.g., the OS), as is required for OS–CC compound trials.

The XY compound was not considered in this analysis because the participants could have used an elemental strategy and attended to only one of the compound's elements (X or Y) but still have learned the correct response for this compound; thus, there is no guarantee that the learning that accrues to this compound involved the learning of an interaction between cues. The EXT cue was not considered because its changing valence makes the results more difficult to interpret.

As a basis for our planned comparisons, the prediction accuracies were submitted to a repeated measures, factorial ANOVA of condition and stimulus across the two training phases. Accuracy was assessed by scoring a correct prediction as 1.0 and an incorrect prediction as 0.0. The condition × stimulus within-training-phase interaction was statistically significant [$F(10,340) = 2.99$, $MS_e = 0.265$, $p < .01$].

To identify any between-condition differences in accuracy between those trials involving an occasion setting compound (OS→CC) and those involving one of its elements (CC), we performed planned orthogonal comparisons of accuracy on OS→CC compound trials (for both occasion setting compounds) with accuracy on CC element trials in the initial phase and of accuracy on $OS_2 \rightarrow CC_2$ compound trials with accuracy on CC_2 element trials in the extinction phase (the $OS_1 \rightarrow CC_1$ compound did not occur in this phase).

In the serial condition, the predictive accuracy on compound trials was slightly higher than that on element trials during the initial phase (Ms = 77.8% and 72.0% correct, respectively), but not in the extinction phase (Ms = 90.3% and 94.9%, respectively); neither of these differences approached statistical significance. By contrast, in the simultaneous condition, the predictive accuracy on compound trials was lower than that on element trials both in the initial phase (Ms = 67.5% and 74.7% correct, respectively) [$t(170) = 1.708$, $p < .10$] and in the extinction phase (Ms = 82.4% and 94.4%, respectively) [$t(170) = 3.238$, $p < .01$]. This analysis suggests that separating the cues in time makes it easier to conditionalize an element's efficacy.

Discussion

In Experiment 1, we compared occasion setting in serial and simultaneous training conditions by using a between-subjects design. We observed no overall difference between the conditions in terms of the participants' asymptotic accuracy in judging the predictive efficacy of the cues; however, the two conditions differed in terms of the participants' responding when an OS was paired with other events in a transfer test and in terms of the relative prediction accuracy on OS→CC compound trials and CC-alone trials.

The transfer of an OS to other targets in the serial condition was distinctly different from the transfer observed in the simultaneous condition. In the serial condition, transfer was strong for the CC of another compound but weaker for a trained and extinguished cue; however, in the simultaneous condition, transfer was strong both for the CC of another compound and for a trained and extinguished cue. The contrast in transfer between the serial and simultaneous conditions nicely demonstrated an effect of time on learning within a causal induction task. Given that we did not observe highly elemental responding in the simultaneous condition (documented by the relatively small effect of OS extinction), the basis for the difference in transfer behavior in the two conditions cannot be due to modulatory versus elemental responding (as seems to be the case with nonhuman animals; Holland, 1986). In the General Discussion section, we will consider two other accounts of the differential transfer that we observed: (1) less attention to a compound's elements when they are presented simultaneously rather than seri-

ally, and (2) presenting an event at a different time from the others may produce a more distinct role for that event.

Under serial presentation of the OS→CC compound, our observation that extinction of the OS failed to eliminate the OS→CC compound's perceived efficacy and that the occasion setting properties of an event transferred to a new CC parallel results documented in studies of occasion setting in animal learning preparations. Under simultaneous presentation of the OS→CC compound, extinction of the OS also failed to eliminate the OS→CC compound's perceived efficacy. Holland (1986, 1991a), however, observed a larger effect of OS extinction on the efficacy of a simultaneous OS→CC compound than the one observed here (in the simultaneous condition of Experiment 1, the OS→CC compound still produced effect predictions on 66.7% of these compound trials after OS extinction). Apparently, our participants learned that the compound had an efficacy that was more than the sum of its constituent elements, so that extinguishing one of the elements did not extinguish the compound's efficacy. This configural learning produced a smaller effect of extinction in the simultaneous condition than would be anticipated by elemental learning.

Experiment 1 thus revealed that the serial presentation of putative causes produces a distinctly different behavioral profile than that observed under simultaneous stimulus presentation. The most surprising result was the weaker learning of a compound's efficacy when the events were presented simultaneously. Although unexpected, there is some evidence of similar effects in nonhuman animals. For example, Nakajima (1992) has found that the simultaneous presentation of events in an occasion setting relation produces weaker overall learning than the serial presentation of events in the classical conditioning of pigeons (but only when there was no temporal delay between OS offset and CC onset).

In contrast, Holland (e.g., Holland, 1991a) has generally found similar learning rates under serial and simultaneous presentation in rats, and, when a difference has been observed, the learning advantage occurs following simultaneous presentation. We did not, however, find an overall advantage for serial over simultaneous presentation, but rather an interaction: Accuracies were very similar for compounds and their elements in the serial condition, but accuracy on compound trials was decidedly lower than accuracy on element trials in the simultaneous condition. We are unaware of any tests of relative accuracy involving nonhuman animals.

EXPERIMENT 2

In Experiment 1, we found differences in the causal relationship induced when an occasion setting compound's events were presented serially rather than simultaneously. In Experiments 2A and 2B, we sought to extend our paradigm to include situations in which the same OS served both to increase the likelihood of the effect following one CC (positive occasion setting) and to decrease the likelihood of the effect following a different CC (negative occasion setting). Our use of an *ambiguous occasion setting* task can determine whether observers are able to learn that the apparent ambiguity in the modulatory action of an OS can be resolved by attending to the CC with which it is paired. Additionally, Holland and Reeve (1991) have used this discrimination to rule out the possibility that the OS serves a general modulatory function that can be used to change the likelihood of an effect expectation after almost any event (see Rescorla, 1985). If the OS's modulatory power involved simply raising or lowering the activation threshold of the effect (as suggested by Rescorla, 1985), then people would learn either the positive occasion setting relationship or the negative occasion setting relationship, but not both; the OS could not both raise and lower that threshold, unless its power was specific to the CC with which it was paired.

In Experiment 2A, participants were required to learn an ambiguous occasion setting discrimination to determine whether they (1) could learn the task (thus ruling out simple modulation as an account of occasion setting) and (2) would evidence stronger OS→CC learning in the serial condition and stronger CC learning in the simultaneous condition (thus replicating the unexpected finding of Experiment 1). We did not use an extinction phase because direct OS–effect learning cannot support learning of the ambiguous occasion setting discrimination. We also did not test for transfer of occasion setting to new events. Due to the more complex CC histories in ambiguous occasion setting discriminations (transfer may depend on whether a CC was previously in a positive or negative occasion setting relation and whether the previous relation involved an ambiguous occasion setter), we are pursuing the effect of transfer in a future series of studies.

More Manipulations of Time

Participants in Experiment 2A were placed in one of three conditions: simultaneous, serial, or overlap. The simultaneous and serial conditions involved the same temporal relations as those used in Experiment 1. The overlap condition involved the same OS onset and CC onset times as that used in the serial condition, but the OS duration was 6,000 msec rather than 1,000 msec. This increase in the OS's duration was designed to mimic a situation in which the OS remained "active" throughout the entire trial, eventually overlapping with the CC and coterminating with it.

The overlap condition was included to determine whether the OS having a weaker memory trace than the CC at the time of effect occurrence is critical to anticipating a difference between learning under simultaneous and serial cue presentation. If performance in the overlap condition were similar to that in the serial condition, then we could conclude that a difference in the memory trace strength of the events at the time of effect occurrence was not the basis of any between-condition disparities observed, because the trace strengths of both events would be approximately equal at the time of effect occurrence.

If, however, performance in the overlap condition were similar to that in the simultaneous condition, then we could conclude that a difference in the memory trace strength of the events at the time of effect occurrence might be the basis of any between-condition disparities observed.

Flexibility in the Application of Causal Knowledge

Although cue order produced different patterns of behavior in the simultaneous and serial conditions of Experiment 1, we were further interested in the flexibility of "order" knowledge once it was acquired. Therefore, we added two new transfer tests in Experiment 2A in which the cue order within an occasion setting compound was reversed during testing. This change might produce a significant disruption in behavior because of the mismatch between the training order and the testing order (the encoding specificity principle; Thomson & Tulving, 1970); however, it might produce no clear disruption if participants were able to reconstruct the original order or if they no longer considered order relevant. Of course, participants in the simultaneous condition should be unaffected by this manipulation, because their cues have no temporal order.

Experiment 2A

Method

Participants. Sixty-six students enrolled in an introductory psychology course at the University of Iowa served as voluntary participants. They received course credit for their participation.

Procedure. The basic procedure was identical to that used in Experiment 1.

The probability of a chemical reaction following the cues is shown in Table 2. The experiment comprised two phases—the training phase and the testing phase—with no pause between the phases; the change between phases was not signaled in any way (although the change to the testing phase was obvious, given that none of the predictions provided feedback).

The training phase consisted of seven randomized blocks comprising feedback trials. The participants observed 16 trials in each block that included 4 ambOS→posCC^{100}, 4 posCC^{0}, 4 ambOS→negCC^{0}, and 4 negCC^{100} trials, for a total of 112 training trials.

The testing phase consisted of three randomized blocks comprising test trials. The participants observed 7 trials in each block that included each of the cues or cue compounds experienced in training (posCC*, negCC,* ambOS→posCC*, and ambOS→negCC*), the never-before-presented ambOS by itself, and two reversal transfer tests (posCC→ambOS*, and negCC→ambOS*), for a total of 21 testing trials. Note that the reversal tests were irrelevant in the simultaneous condition because these test trials did not differ from those trials presented in training—there was no temporal order to reverse.

The participants were randomly assigned to one of three conditions: serial, simultaneous, or overlap. In the simultaneous condition, all cues occurring in a trial were 1,000 msec (plus the RT) in duration and started and ended together. In the serial condition, a cue that occurred first in a serial compound (ambOS) was presented for 1,000 msec, and its termination was followed by the second cue (posCC or negCC, with a duration of 1,000 msec plus the RT) after a delay of 4,000 msec. In the overlap condition, a cue that occurred first in a serial compound (ambOS) was presented for 6,000 msec (plus the RT), with the second cue's onset occurring 5,000 msec after the onset of the first cue; the second cue was 1,000 msec in duration (plus the RT), so that both cues coterminated. Single cues (posCC or negCC) in all three conditions were 1,000 msec in duration (plus the RT). All cues that were active at the onset of the "RESPOND" prompt (i.e., all of the cues in the simultaneous and overlap conditions and the CC cues in the serial condition) remained on the screen until a response was registered.

Assignment of cue identity to cue type was done through a 4 × 4 Latin square: 4 cues (Adelpine, Bucagon, Furval, and Glexus) × 4 roles (ambOS, posCC, negCC, and absent). Each of the four workstations used a consistent assignment of cue to role for a given condition. Condition was manipulated between subjects.

Table 2
The Percentage of Trials in Which a Chemical Reaction Followed the Addition of Chemical(s) in Experiments 2A and 2B

Chemical	Percentage (Number) of Trials With a Chemical Reaction
Positive Occasion Setting	
ambOS→posCC	100 (28)
posCC	0 (28)
Negative Occasion Setting	
ambOS→negCC	0 (28)
negCC	100 (28)

Note—The number in parentheses indicates the number of trials of that type within the phase. The participants in the serial and overlap conditions experienced the cue preceding the arrow before the cue following the arrow; the participants in the simultaneous condition experienced both of these cues simultaneously.

Results

Because there would have been differential attrition across conditions (the participants in the simultaneous condition had greater difficulty learning the required discrimination than did those in the other two conditions), we did not eliminate any participants in Experiment 2A for failing to reach the 50% performance criterion. Overall performance accuracy averaged 88.2% correct during the final block of training (Block 7) and 82.6% correct during testing (for those stimuli that received differential feedback during the training phase).

Training performance. Performance in all three conditions reflected the prevailing contingencies (Figure 3). The participants were clearly able to use a single OS as a positive OS for one event (posCC) and as a negative OS for a second event (negCC). Additionally, performance in the overlap condition was nearly identical to that in the serial condition. The participants in the simultaneous condition appear to have been more accurate on the negCC trials and less accurate on the ambOS→posCC trials and the ambOS (negCC trials than were the participants in the serial and overlap conditions.

To confirm these observations, we submitted accuracy scores across the entire training phase to a repeated measures, factorial ANOVA of condition (serial vs. simultaneous vs. overlap) and stimulus (posCC, negCC, ambOS→posCC, and ambOS→negCC), with block (1–7) as a blocking factor. A trial was given an accuracy score of 1.0 when the participant made the correct prediction (i.e., chose key "1" when the effect was scheduled to occur and key "3" when the effect was not scheduled to occur) and an accuracy score of 0.0 otherwise. The main effect for condition was not statistically significant ($F < 1$).

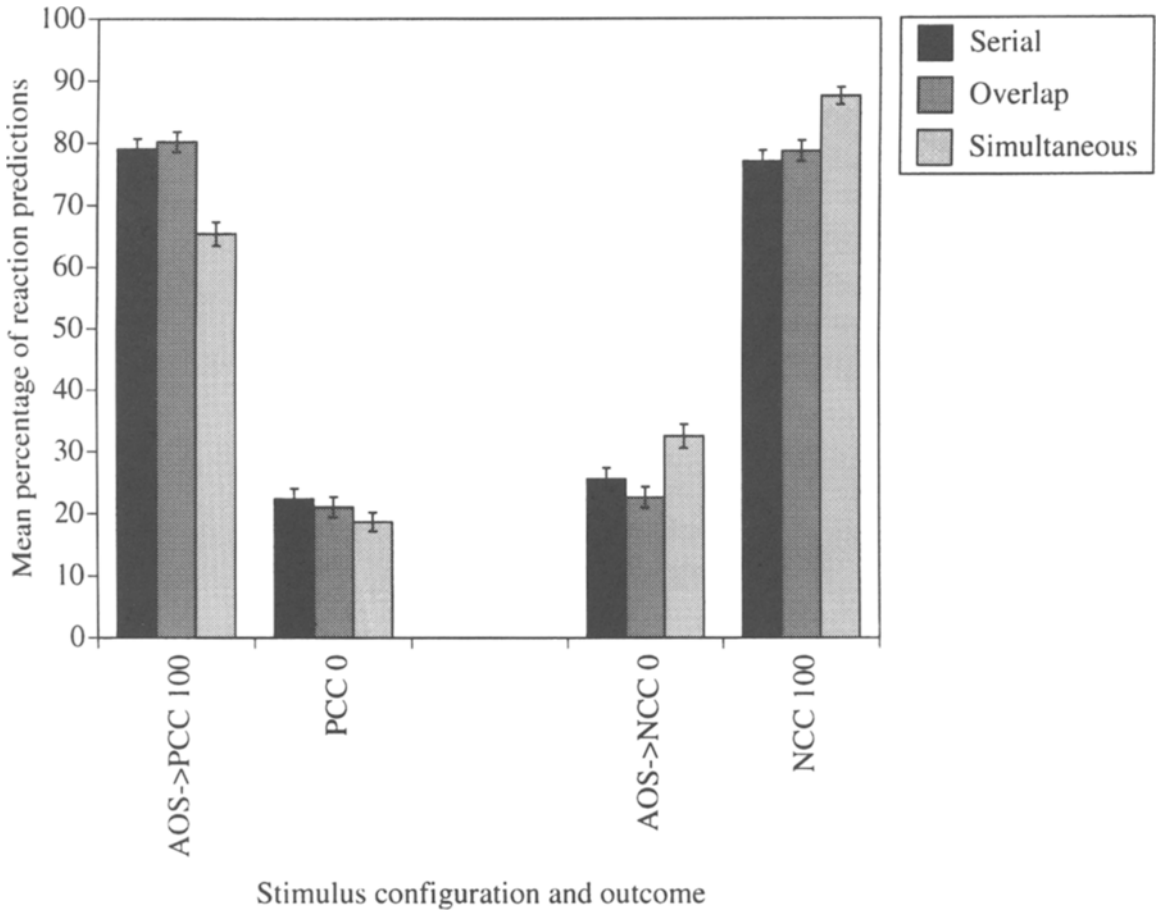

Figure 3. Mean percentage of reaction predictions for each stimulus or stimulus compound during the training phase of Experiment 2A.

There was a statistically significant main effect for stimulus [$F(3,189) = 7.01, MS_e = 0.374, p < .001$], and there was a statistically significant condition × stimulus interaction [$F(6,189) = 6.42, MS_e = 0.374, p < .0001$].

For the purposes of determining the presence of between-condition differences in learning, we performed planned orthogonal comparisons ($\alpha = .05$) of the occasion setting compounds and the elements in each of the three conditions. There were no statistically significant differences between accuracy on compound and element trials for the serial and overlap conditions ($M = 77.9\%$); however, in the simultaneous condition, accuracy on compound trials ($M = 66.4\%$) was significantly lower than accuracy on element trials ($M = 84.4\%$) [$t(63) = 7.31$, $p < .0001$]. These results are similar to those observed in Experiment 1: The participants in the simultaneous condition learned the efficacy of the elements faster than the efficacy of the compounds containing those elements, but there was no difference between learning of the efficacy of compounds and their elements in the serial and overlap conditions.

Transfer tests. A summary of testing performance is shown in Figure 4. This figure reveals that the similarity in the response profile between the serial and overlap conditions that was observed in the training phase persisted into the testing phase. The profile for the simultaneous condition remained distinctly different; responses to the compounds were consistently less accurate in this condition than in the other two conditions, and responses to the negCC element were more accurate in the simultaneous condition than in the other two conditions, whereas accuracy for the posCC element was similar across the three conditions.

Reversing the order of the cues between training and testing (relevant only for the serial and overlap conditions) had no apparent effect on predictive accuracy in either condition. In both conditions, the response profile for the ambOS→posCC compound was similar to that for the posCC→ambOS compound, and the response profile for the ambOS→negCC compound was similar to that for the negCC→ambOS compound (Figure 4).

To confirm our conclusions, the accuracy scores during the testing phase were submitted to a repeated measures, factorial ANOVA of condition (serial vs. simultaneous vs. overlap) and stimulus (posCC, negCC, ambOS→posCC, ambOS→negCC, and ambOS). The main effect for condition was not statistically significant [$F(2,63) = 1.68, MS_e = 0.454, p > .10$], and the main effect for stim-

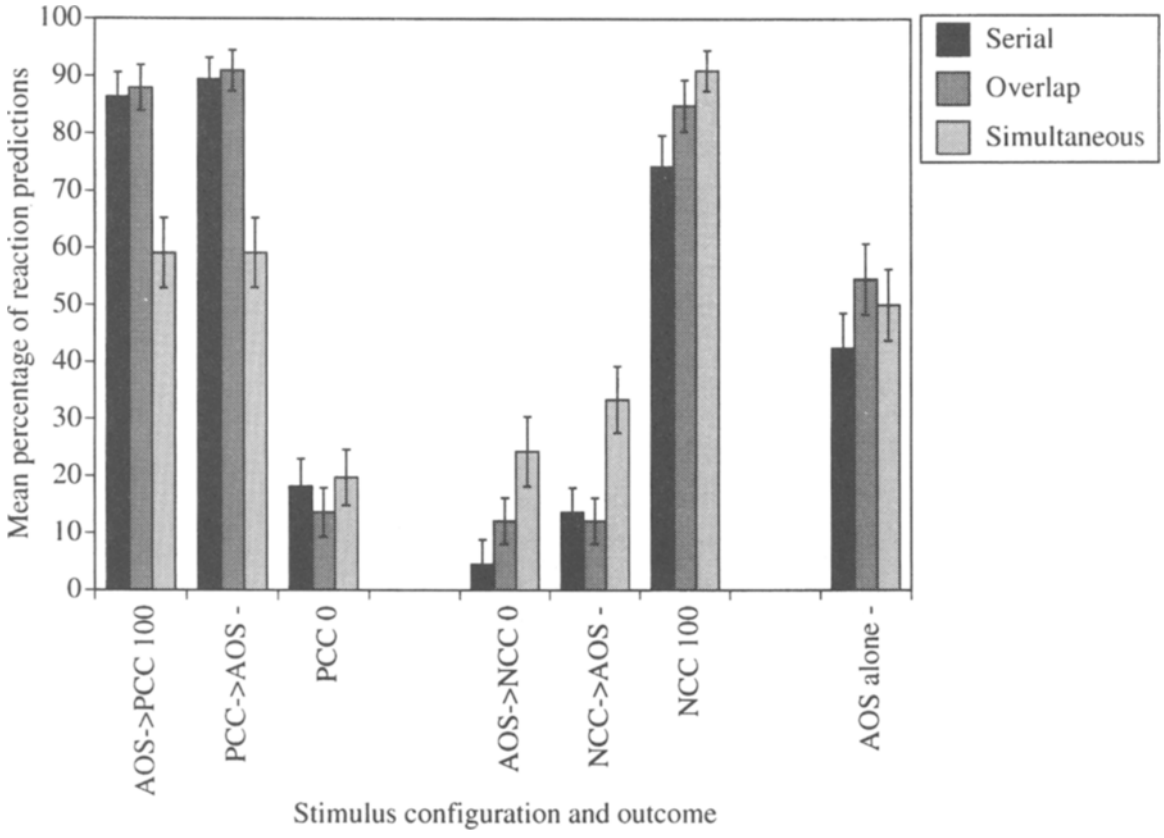

Figure 4. Mean percentage of reaction predictions for each stimulus or stimulus compound during the testing phase of Experiment 2A.

ulus was not statisically significant [$F(3,189) = 1.65$, $MS_e = 0.152$, $p > .10$]. The condition × stimulus interaction, however, was statistically significant [$F(6,189) = 4.726$, $MS_e = 0.152$, $p < .001$].

For the purpose of determining the presence of between-condition differences in asymptotic performance (Figure 4), we again performed planned orthogonal comparisons ($\alpha = .05$) of the occasion setting compounds and the elements in each of the three conditions. There were no statistically significant differences between accuracy on compound and element trials for the overlap condition (*M*s = 87.9% and 85.6%, respectively); however, in the serial condition, accuracy on compound trials ($M = 90.9\%$) was significantly higher than accuracy on element trials ($M = 78.0\%$) [$t(126) = 2.68$, $p < .01$], whereas in the simultaneous condition, accuracy on compound trials ($M = 67.4\%$) was significantly lower than accuracy on element trials ($M = 85.6\%$) [$t(126) = 3.79$, $p < .001$]. The participants' stronger learning of the efficacy of the elements in the simultaneous condition persisted through the testing phase. Interestingly, we also observed the first difference between the serial and overlap conditions: The participants in the serial condition evidenced a larger disparity in prediction accuracy between compounds and elements than did the participants in the overlap condition.

Cue order. Figure 4 reveals no apparent effects of presenting the elements of the occasion setting compounds in the order opposite to that used during testing. The difference between the trained order and the opposite order in the serial and overlap conditions was little different from that observed in the simultaneous condition (where there should be no effect because the "orders" used in training and testing are identical—they are both simultaneous compounds).

To confirm our observations, we first submitted the reaction predictions during the testing phase to a repeated measures, factorial ANOVA of condition (serial vs. simultaneous vs. overlap) and stimulus (posCC, negCC, ambOS→posCC, ambOS→negCC, ambOS, posCC→ambOS, and negCC→ambOS). The main effect for condition was not statistically significant [$F(2,63) = 1.24$, $MS_e = 0.602$, $p > .10$]. There was a statistically significant main effect for stimulus [$F(6,378) = 69.33$, $MS_e = 1.157$, $p < .0001$], and there was a statistically significant condition × stimulus interaction [$F(12,378) = 3.47$, $MS_e = 1.157$, $p < .0001$].

Given the significant interaction, we performed planned orthogonal comparisons ($\alpha = .05$) of the trained (e.g., ambOS→ posCC) and reversed order (e.g., posCC→ambOS) compounds in each of the three conditions. None

of these differences were statistically significant (all $ps > .25$). Thus, although serial ordering of the elements during training did have an effect on learning, the acquired knowledge appears to be quite flexible in its later application. Did this flexibility come at any cognitive cost?

To answer this question, we performed one final analysis of asymptotic performance by submitting testing phase RTs to a repeated measures, factorial ANOVA of condition (serial vs. simultaneous vs. overlap) and stimulus (posCC, negCC, ambOS→posCC, ambOS→ negCC, ambOS, posCC→ambOS, and negCC→ambOS). The main effect for condition was not statistically significant [$F(2,63) = 1.01, MS_e = 0.791, p > .10$]. There was a statistically significant main effect for stimulus [$F(6,378) = 28.33, MS_e = 0.110, p < .0001$], and there was a statistically significant condition × stimulus interaction [$F(12,378) = 9.65, MS_e = 0.110, p < .0001$].

This analysis was followed by planned orthogonal comparisons ($\alpha = .05$) of the trained (e.g., ambOS→ posCC) and reversed order (e.g., posCC→ambOS) compounds in each of the three conditions. We again found that any differences due to order in the overlap and simultaneous conditions failed to reach statistical significance ($ps > .25$). However, we did find significantly longer RTs when the order was reversed in the serial condition; the RT for the trained order was significantly shorter than that for the reversed order for both the ambOS→posCC compound ($M = 511$ and 733 msec, respectively) and the ambOS→ negCC compound ($M = 456$ and 661 msec, respectively).

Discussion

The participants were clearly able to learn an ambiguous occasion setting relationship whether those cues were presented simultaneously or successively. This finding makes Rescorla's (1985) original proposal untenable; the OS is not functioning either to increase or to decrease the likelihood of a response for any target event. Combined with the specificity of transfer noted in Experiment 1, our evidence strongly suggests that the action of an OS is at least partially dependent on the identity of its target CC. These findings parallel those from analogous animal experiments (Holland, 1991b).

In tests of the relative accuracy of predictions for compounds and elements in the three conditions, we found that the participants' responses were more accurate with a compound when the cues were presented serially and more accurate with the compound's elements when the cues were presented simultaneously (Figures 2 and 3). Thus, the stronger learning of an element's efficacy under simultaneous cue presentation and the stronger learning of a compound's efficacy under serial presentation appear to be robust findings.

We found that performance in the overlap condition was very similar to that observed in the serial condition, leading us to conclude that the relative memory trace strength of the events at the onset of a prediction is not the basis of the observed differences in accuracy for compounds and elements. Our results indicate that any appeal to the relative strengths of the two cues must be modified to include a possible difference in the *associative strength* for each event rather than a differential in memory trace strength.

In real-time theories of occasion setting in the conditioning of nonhuman animals (Brandon & Wagner, 1998; Schmajuk, Lamoureux, & Holland, 1998), learning occurs throughout the entire trial. Thus, as the duration of a predictive event is lengthened, during most of the event's duration, the outcome is absent, thus producing a weak (direct) association between the long-duration event and the outcome. In the model proposed by Schmajuk et al. (1998), occasion setting can produce two associations between an OS and the outcome: a direct OS–outcome association and an indirect (OS→CC)–outcome association. When the conditions for the development of a direct OS–outcome association are favorable (e.g., in the simultaneous condition), then this association may overshadow the indirect association. In contrast, when the conditions for the development of a direct OS–outcome association are unfavorable (e.g., in the serial and overlap conditions), then the indirect association will overshadow the direct association. The reasons that the conditions are unfavorable in the serial and overlap conditions are quite different. In the serial condition, the memory trace strength of the OS is weak at the time of the onset of the effect. In the overlap condition, however, the memory trace strength is still strong at effect onset, but the effect association is weakened because, during most of the duration of the OS, the effect was absent.

Finally, although the cue order experienced during learning produced significant differences between the simultaneous and serial presentation of cues in Experiments 1 and 2A, cue order appeared to be largely irrelevant in the later application of the acquired causal knowledge. The participants' predictions for trials in which the OS and the CC were presented in the opposite order did not differ from predictions for trials in which the original order was retained. Interestingly, there was a cost for this flexibility, but only in the serial condition: The participants' RTs were over 200 msec slower for the reversed compounds. Note that the OS and CC are presented together during a trial's final 1,000 msec for the overlap and simultaneous conditions; the reversal manipulation has no effect on which cues were present at the time of a response. In contrast, only the CC is present during the trial's final 1,000 msec in the serial condition. This critical difference among the conditions may have produced the slower responding under reversal that was observed in the serial condition (where the OS would be present during the trial's final 1,000 msec rather than the CC), but not in the simultaneous and overlap conditions (where both the OS and the CC would be present).

Experiment 2B

In addition to the important difference between the simultaneous and serial presentation of the cues in Exper-

iment 2A, the simultaneous condition also differed from the serial and overlap conditions in terms of the average length of a trial. Although the participant had control over how much time passed between trials, the trial itself unfolded on a predetermined schedule. For the simultaneous condition, the cues were presented for 1,000 msec before the response prompt was displayed, whereas for the serial and overlap conditions, the cues spanned 6,000 msec before the response prompt was displayed. It is therefore possible that the presentation of shorter trials produced stronger learning of the elements and that the presentation of longer trials produced stronger learning of the compounds.

Therefore, in Experiment 2B, we used three versions of the simultaneous condition, to determine whether making trials longer would produce stronger learning of the compound and weaker learning of the elements. We lengthened the trial either by increasing the duration of the cues to 6,000 msec (the long-stimulus condition) or by inserting an additional passage of time (5,000 msec) before the presentation of the 1,000-msec cues (the long-ITI condition, so called because this manipulation effectively increased the intertrial interval [ITI] by 5,000 msec). We anticipated that these manipulations might produce an overall improvement in learning (e.g., increasing the ITI is known to improve occasion setting in conditioning preparations; Holland, 1995, 1999; Holland & Morell, 1996), but we were uncertain as to their possible effect on the element versus compound difference observed in Experiments 1 and 2A.

Method

Participants. Forty-two students enrolled in an introductory psychology course at the University of Iowa served as voluntary participants. They received course credit for their participation.

Procedure. The basic procedure was identical to that used in Experiments 1 and 2A.

The probability of a chemical reaction following the cues was identical to that in Experiment 2A and is shown in Table 2. The delivery and blocking of trials were also identical to those in Experiment 2A.

The participants were randomly assigned to one of three conditions: simultaneous, long ITI, or long stimulus. In the simultaneous condition, all cues occurring in a trial were 1,000 msec in duration (plus the RT) and started and ended together. In the long-ITI condition, all cues occurring in a trial were 1,000 msec in duration (plus the RT) and started and ended together; however, a trial began with a 5,000-msec delay before cue presentation, thus making the effective ITI much longer. In the long-stimulus condition, all cues occurring in a trial were 6,000 msec in duration (plus the RT) and started and ended together.

In Experiment 2B, we used the same counterbalancing scheme used in Experiment 2A.

Results and Discussion

In order to parallel the analysis of Experiment 2A, we did not eliminate any participants in Experiment 2B for

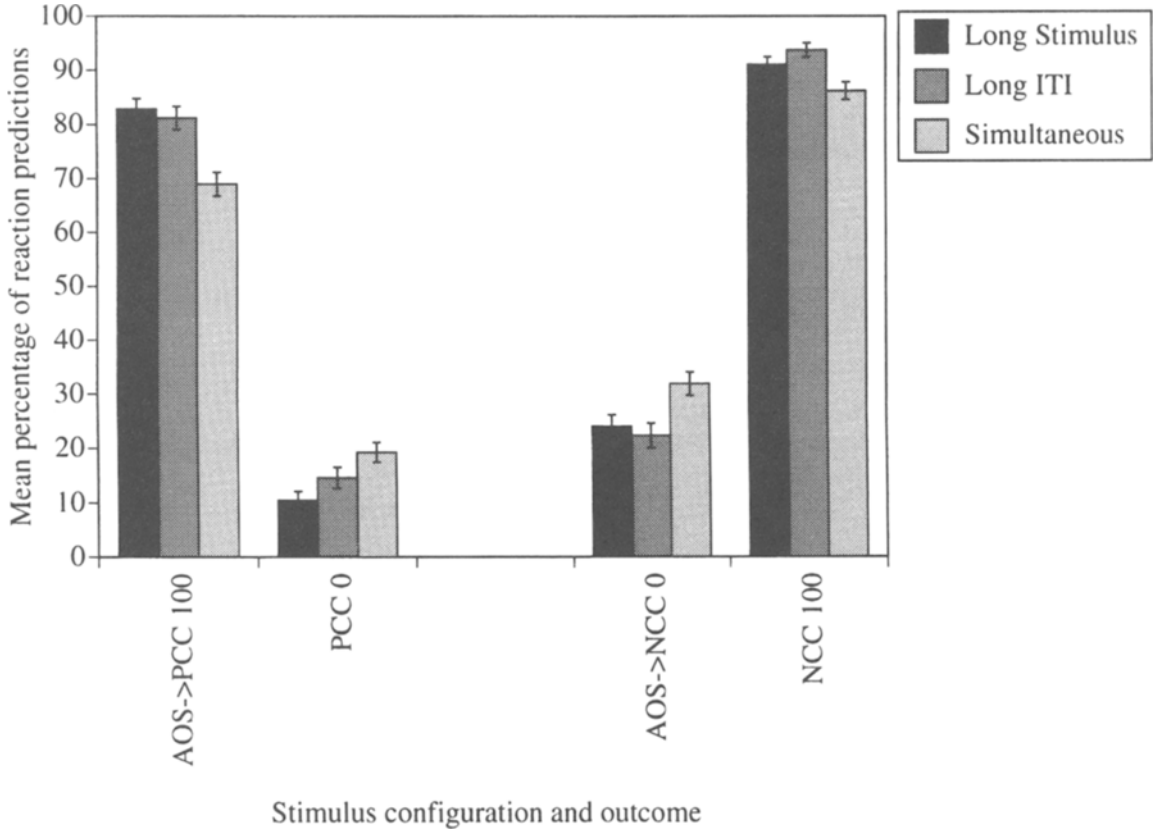

Figure 5. Mean percentage of reaction predictions for each stimulus or stimulus compound during the training phase of Experiment 2B.

failing to reach the 50% performance criterion. Overall performance accuracy averaged 94.2% during the final block of training (Block 7) and 89.3% during testing (for those stimuli that received differential feedback during the training phase).

Training performance. Performance in all three conditions reflected the prevailing contingencies (Figure 5). Additionally, accuracy in the simultaneous condition was poorer than that in the long-ITI and long-stimulus conditions, although this difference did not appear to interact with cue identity.

To confirm these observations, we submitted accuracy scores across the entire training phase to a repeated measures, factorial ANOVA of condition (simultaneous vs. long ITI vs. long stimulus) and stimulus (posCC, negCC, ambOS→posCC, and ambOS→negCC), with block (1–7) as a blocking factor. The main effect for condition was statistically significant [$F(2,39) = 3.25$, $MS_e = 1.298$, $p < .05$]. There was also a statistically significant main effect for stimulus [$F(3,117) = 20.41$, $MS_e = 0.309$, $p < .0001$], but the condition × stimulus interaction did not approach significance ($F < 1$). This analysis was followed by a series of planned comparisons ($\alpha = .05$). The training accuracy in the simultaneous condition ($M = 76.0\%$) was lower than that in the long-ITI ($M = 84.5\%$) and long-stimulus ($M = 84.9\%$) conditions, with the latter two not differing from each other. Note that the accuracy in the simultaneous condition in Experiment 2B was nearly identical to that observed in the identical condition in Experiment 2A ($M = 75.4\%$). A comparison of the accuracy on compound trials ($M = 75.8\%$) and element trials ($M = 87.8\%$) revealed that the efficacy of the compounds was lower than the efficacy of one of its elements, replicating the stronger element learning observed in the simultaneous condition of Experiment 2A. The size of this difference did not vary across conditions, as revealed by the nonsignificant interaction.

Transfer tests. Accuracy scores during the transfer phase were very similar for all four of the stimuli (not shown in any figure). A repeated measures, factorial ANOVA of condition (simultaneous vs. long ITI vs. long stimulus) and stimulus (posCC, negCC, ambOS→posCC, ambOS→negCC, and ambOS) revealed no significant effects for condition [$F(2,39) = 2.531$, $p < .10$], stimulus [$F(3,117) = 1.635$, $p > .10$], or condition × stimulus ($F < 1$). Ceiling effects on accuracy scores and the relatively small number of trials in the testing phase may have been responsible for the absence of any significant effects. The differences we observed in testing, however, were consistent with those we observed in training: Accuracy on compound trials ($M = 87.3\%$) was lower than accuracy on element trials ($M = 91.3\%$), and accuracy in the simultaneous condition ($M = 83.3\%$), was lower than that in the long-ITI ($M = 93.8\%$) and long-stimulus ($M = 92.3\%$) conditions.

GENERAL DISCUSSION

The participants in our study observed one event, the conditional cause (CC), followed by the effect only when it was accompanied by another event, the occasion setter (OS). We obtained a number of notable behavioral results. First, the participants rapidly learned to predict the effect only when it was preceded by both the OS and the CC; the likelihood of an effect prediction was significantly higher when the pair was given than when either the OS or the CC was given alone (Experiment 1). Second, asymptotic predictions of the causal efficacy of the events closely accorded with the true contingencies (Experiments 1 and 2). Third, the percentage of reaction predictions following the OS→CC compound remained high even when the direct OS–effect relation was extinguished in a later training phase (Experiment 1). Fourth, the modulatory properties of an OS presented in a serial relationship with a CC were especially strong for events that were CCs either in the same or in a different occasion setting relationship with the same outcome; this transfer was clearly different from that observed for an OS presented simultaneously with the CC (Experiment 1). Fifth, reversing the order of serially presented cues during testing had no effect on prediction accuracy. Reversing cue order, however, produced slower predictions when the original training involved an OS that terminated before offset of the CC; reversing cue order did not produce slower predictions when the original training involved an OS that coterminated with the CC (Experiment 2A). Finally, we consistently observed that learning the efficacy of an OS→CC compound relative to that of its CC element was improved by presenting the events serially rather than simultaneously (Experiments 1 and 2). These results provide strong evidence for the effect of time in a serial causal induction task.

As long as learning involves a speeded, attention-limited presentation of events that engender few preexperimental expectations, we anticipate that our findings will generalize to analogous domains. When people have unlimited time, however, they may engage more rule-based strategies that overshadow the associative mechanisms we have documented (Sloman, 1996). Additionally, the presence of preexperimental expectations would be expected to considerably attenuate the patterns of behavior we have documented. We specifically chose the domain of chemistry to ensure that the participants would readily accept that the candidate causes could interact with one another, that the effects of these causes might be delayed, and that one chemical could mimic the effects of another. In contrast, if we had used a chemical as an OS in one occasion setting relation and examined transfer to a CC trained with a very different OS (e.g., a buttonpress), then the results may have turned out quite differently, because prior experience does not lead one to anticipate that a chemical and a buttonpress might have

analogous effects. Young children, however, might readily reveal transfer patterns under conditions in which adults would not, because children lack the experience that produces skepticism about certain causal relationships (e.g., Ausubel & Schiff, 1954; Berzonsky, 1971).

The serial ordering of the causal antecedents opens the field of causal induction to further explorations of the importance of time to causal judgments. None of the currently popular models of causal induction (e.g., Cheng, 1997; Rescorla & Wagner, 1972; Van Hamme & Wasserman, 1994) explicitly represent temporal relations; they all assume that causal candidates and their effects are sufficiently contiguous to support a possible causal relationship. Events either are or are not candidate causes—there is no representation of degree of contiguity. These models do not make differential predictions when the events in an occasion setting relation are presented serially rather than simultaneously. This lack of a theoretical framework for the representation of time is made more relevant when the importance of time is documented empirically, as we have done here.

Table 3
Percent Correct for Compound and Element Trials in Experiments 1, 2A, and 2B

Condition	Type of Trial: Compound	Type of Trial: Element
Experiment 1: Initial Phase		
Serial	77.8	72.0
Simultaneous	67.5	74.7
Experiment 1: Extinction Phase		
Serial	90.3	94.9
Simultaneous	82.4	94.4
Experiment 2A		
Serial	90.9	78.0
Simultaneous	67.4	85.6
Overlap	87.9	85.6
Experiment 2B		
Simultaneous	68.5	83.5
Long ITI	79.5	89.6
Long Stimulus	79.4	90.3

Alternative Accounts

The participants consistently showed more general transfer to novel OS compounds and weaker learning of the compound's efficacy when the compound's elements were presented simultaneously, but they showed more specific transfer to novel OS compounds and stronger learning of the compound's efficacy when its elements were presented serially.[2] There are three possible reasons that we will consider for these observations.

Divided attention. When the two events in a compound are presented simultaneously, observers must divide their attention between the events; but, when the two events are presented serially, observers can first attend to one event and then attend to the other. Thus, during simultaneous stimulus presentation, this division of attention would produce less attention to the elements of a compound. During transfer, this lack of attention to the compound's elements would make it easier to overlook the identity of the event that is paired with the OS, thus producing greater transfer. Under serial presentation, however, attention can be wholly focused on each event as it is experienced, thus making it harder for an observer to overlook changes in the event that is paired with the OS.

Differences in the attentional requirements of the two tasks might also explain the disparity in element and compound learning under serial and simultaneous cue presentation. Observers who saw the events presented simultaneously may have exhibited poorer compound learning because it was more difficult to attend to both elements of the compound during the trial. In contrast, observers who saw the events presented serially would have been able to fully attend to both of the elements of the compound. Although the presentation of two cues (rather than one) may seem unlikely to tax attentional resources, this possibility cannot be ruled out.

There are two pieces of evidence indicating that attention may be more divided during simultaneous stimulus presentation than during serial stimulus presentation. First, we have consistently observed in this study and other studies that RTs on simultaneous compound trials are 70–100 msec longer than RTs on element trials. For example, in the simultaneous condition of Experiment 2A, RTs on correct compound trials were 78 msec longer (M = 458 msec) than RTs on correct element trials (M = 380 msec). Thus, the participants not only showed lower accuracy on compound trials but also took longer to make those decisions. In contrast, the RTs on serial compound trials are over 200 msec shorter than RTs on element trials. For example, in the serial condition of Experiment 2A, RTs on correct compound trials were 263 msec shorter (M = 433 msec) than RTs on correct element trials (M = 696 msec). Second, the difference between accuracies under serial and simultaneous presentation is most notable for compounds (as summarized in Table 3); accuracy on element trials is relatively constant across conditions, whereas accuracy on compound trials is always lower in conditions in which the cues were presented simultaneously.

Although attention may be more divided during simultaneous presentation than during serial presentation, increasing the duration of the stimuli should help offset these deleterious effects. As revealed in Experiment 2B, both increasing stimulus duration and increasing the ITI increased overall performance accuracy (although not necessarily for the same reasons). Neither manipulation, however, had an effect on the qualitatively superior learning of the elements relative to the compound under simultaneous presentation. If the superior compound learning under serial presentation is the result of increased attention to the compound's cues, then this benefit must arise only when attention is focused on each cue succes-

sively rather than vacillating between the two (as most likely occurred when stimulus duration was increased).

Distinctiveness. A second possible reason for the effect of time on learning is that a compound's elements may be more distinctive when they are presented serially than when they are presented simultaneously. In effect, the temporal signature of an OS is distinctly different from that of a CC during serial presentation, but not during simultaneous presentation; in more cognitive terms, the functional role of each event would be clearer to an observer under serial presentation than under simultaneous presentation.

If, under serial presentation, the functional role of the OS is distinctly different from that of other events, then its action may be judged to be quite specific (thus explaining the specificity in transfer), and this distinctiveness may make it easier to learn to conditionalize the CC's efficacy (thus explaining the stronger learning of the compound's efficacy). Under simultaneous presentation, the functional role of the OS may be ambiguous and easily confused with the role of other simultaneously presented events; the OS may thus have no special functional role and its action judged to be quite general (thus explaining the generality in transfer), and the OS's lack of distinctiveness may make it more difficult to learn that it functions to conditionalize the CC's efficacy.

Holland (e.g., Holland, 1989a, 1989c) has observed that occasion setting is more likely in nonhuman animals when the two cues are distinctly different from one another (e.g., they involve different sensory modalities, have different intensities, or are presented at different times); unfortunately, the mechanism that produces better occasion setting when the memory codes are more distinct is not yet known.

Competition between elemental and configural learning. Prominent theories of occasion setting in nonhuman animals (Brandon & Wagner, 1998; Schmajuk et al., 1998) predict the greater specificity of transfer under serial stimulus presentation than under simultaneous stimulus presentation that we observed. These accounts, however, rely on the existence of strong elemental learning and very little configural learning when the events are presented simultaneously. This prediction is in direct opposition to the highly configural learning that we observed in the simultaneous conditions of Experiments 1 and 2.

These models' failures to account for our data without appealing to a process that clearly was not in operation in our task make them problematic foundations for future accounts of occasion setting in human causal induction. Both the differential attention account and the distinctive code account, however, show promise. Unfortunately, the lack of any strong formalizations of the underlying mechanisms posited in these accounts currently make it difficult to favor one over the other. We anticipate that further empirical evaluation and theoretical development will clarify the possible roles of attention or distinctiveness in causal induction.

Conclusion

Allowing causes to occur at different times relative to an effect is an important step in our investigation of the processes underlying everyday causal induction. Doing so, however, reveals the limitations of current models of these cognitive processes. In the few studies in which events preceded the effect but were temporally ordered with respect to one another (Reed, 1992; Shanks, 1989), current theories were also found wanting. As we accumulate evidence in which temporal relations among causes and their effects are manipulated, the explanatory demands of this evidence should foster the development of a new generation of causal induction theories.

REFERENCES

ALLOY, L. B., & TABACHNIK, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review*, **91**, 112-149.

AUSUBEL, D. P., & SCHIFF, H. M. (1954). The effect of incidental and experimentally induced experience in the learning of relevant and irrelevant causal relationships by children. *Journal of Genetic Psychology*, **84**, 109-123.

BERZONSKY, M. D. (1971). The role of familiarity in children's explanations of physical causality. *Child Development*, **42**, 705-715.

BRANDON, S. E., & WAGNER, A. R. (1998). Occasion setting: Influences of conditioned emotional responses and configural cues. In N. A. Schmajuk & P. C. Holland (Eds.), *Occasion setting: Associative learning and cognition in animals* (pp. 343-382). Washington, DC: American Psychological Association.

CHENG, P. W. (1993). Separating causal laws from causal facts: Pressing the limits of statistical relevance. In D. L Medin (Ed.), *The psychology of learning and motivation* (Vol. 30, pp. 215-264). San Diego: Academic Press.

CHENG, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, **104**, 367-405.

COHEN, J., MACWHINNEY, B., FLATT, M., & PROVOST, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers*, **25**, 257-271.

DICKINSON, A., & BURKE, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *Quarterly Journal of Experimental Psychology*, **49B**, 60-80.

GILOVICH, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life*. New York: Simon & Schuster.

HOLLAND, P. C. (1983). Occasion-setting in Pavlovian feature positive discriminations. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Discrimination processes* (Vol. 4, pp. 183-206). New York: Ballinger.

HOLLAND, P. C. (1986). Temporal determinants of occasion setting in feature-positive discriminations. *Animal Learning & Behavior*, **14**, 111-120.

HOLLAND, P. C. (1989a). Acquisition and transfer of conditional discrimination performance. *Journal of Experimental Psychology: Animal Behavior Processes*, **15**, 154-165.

HOLLAND, P. C. (1989b). Feature extinction enhances transfer of occasion setting. *Animal Learning & Behavior*, **17**, 269-279.

HOLLAND, P. C. (1989c). Occasion setting with simultaneous com-

pounds in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, **15**, 183-193.

HOLLAND, P. C. (1991a). Acquisition and transfer of occasion setting in operant feature positive and feature negative discriminations. *Learning & Motivation*, **22**, 366-387.

HOLLAND, P. C. (1991b). Transfer of control in ambiguous discriminations. *Journal of Experimental Psychology: Animal Behavior Processes*, **17**, 231-248.

HOLLAND, P. C. (1995). The effects of intertrial and feature–target intervals on operant serial feature-positive discrimination learning. *Animal Learning & Behavior*, **23**, 411-428.

HOLLAND, P. C. (1999). Intertrial interval effects in Pavlovian serial feature positive discriminations. *Animal Learning & Behavior*, **27**, 127-139.

HOLLAND, P. C., & MORELL, J. R. (1996). The effects of intertrial and feature-target intervals on operant serial feature negative discrimination learning. *Learning & Motivation*, **27**, 21-42.

HOLLAND, P. C., & REEVE, C. E. (1991). Acquisition and transfer of control by an ambiguous cue. *Animal Learning & Behavior*, **19**, 113-124.

HUME, D. (1955). *An inquiry concerning human understanding*. Indianapolis: Bobbs-Merrill. (Original work published 1748)

NAKAJIMA, S. (1992). The effect of temporal relationship of stimulus compound on ambiguous discrimination in the pigeon's autoshaping. *Behavioural Processes*, **27**, 65-73.

NISBETT, R. E., & ROSS, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

REED, P. (1992). Effect of a signalled delay between an action and outcome on human judgement of causality. *Quarterly Journal of Experimental Psychology*, **44B**, 81-100.

RESCORLA, R. A. (1985). Conditioned inhibition and facilitation. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 299-326). Hillsdale, NJ: Erlbaum.

RESCORLA, R. A., & WAGNER, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

SCHMAJUK, N. A., LAMOUREUX, J. A., & HOLLAND, P. C. (1998). Occasion setting: A neural network approach. *Psychological Review*, **105**, 3-32.

SHANKS, D. R. (1985). Forward and backward blocking in human contingency judgement. *Quarterly Journal of Experimental Psychology*, **37B**, 1-21.

SHANKS, D. R. (1989). Selectional processes in causality judgment. *Memory & Cognition*, **17**, 27-34.

SHANKS, D. R., & DARBY, R. J. (1998). Feature- and rule-based generalization in human associative learning. *Journal of Experimental Psychology: Animal Behavior Processes*, **24**, 405-415.

SHANKS, D. R., DARBY, R. J., & CHARLES, D. (1998). Resistance to interference in human associative learning: Evidence of configural processing. *Journal of Experimental Psychology: Animal Behavior Processes*, **24**, 136-150.

SHANKS, D. R., PEARSON, S. M., & DICKINSON, A. (1989). Temporal contiguity and the judgement of causality by human subjects. *Quarterly Journal of Experimental Psychology*, **41B**, 139-159.

SLOMAN, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, **119**, 3-22.

SWARTZENTRUBER, D. (1998). Perspectives on modulation: Modulator- and target-focused views. In N. A. Schmajuk & P. C. Holland (Eds.), *Occasion setting: Associative learning and cognition in animals* (pp. 167-197). Washington, DC: American Psychological Association.

THOMSON, D. M., & TULVING, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology*, **86**, 255-262.

VAN HAMME, L. J., & WASSERMAN, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning & Motivation*, **25**, 127-151.

WALDMANN, M. R., & HOLYOAK, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General*, **121**, 222-236.

WASSERMAN, E. A., & BERGLAN, L. R. (1998). Backward blocking and recovery from overshadowing in human causal judgment: The role of within-compound associations. *Quarterly Journal of Experimental Psychology*, **51B**, 121-138.

WASSERMAN, E. A., & NEUNABER, D. J. (1986). College students' responding to and rating of contingency relations: The role of temporal contiguity. *Journal of the Experimental Analysis of Behavior*, **46**, 15-35.

WILLIAMS, D. A., SAGNESS, K. E., & MCPHEE, J. E. (1994). Configural and elemental strategies in predictive learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 694-709.

YOUNG, M. E. (1995). On the origin of personal causal theories. *Psychonomic Bulletin & Review*, **2**, 83-104.

YOUNG, M. E., & WASSERMAN, E. A. (1998). [The effect of delay interval on occasion setting]. Unpublished raw data.

YOUNG, M. E., WASSERMAN, E. A., & JOHNSON, J. L., & JONES, F. L. (2000). Positive and negative patterning in human causal learning. *Quarterly Journal of Experimental Psychology*, **53B**, 121-138.

NOTES

1. In a pilot study using a relatively short interval (1,000 msec) between the termination of the OS and the onset of the CC, we found effects in the predicted direction, but these effects were too small to be statistically significant with a reasonable sample size. So, in Experiment 1, we used a much longer interval (5,000 msec) in order to foster stronger occasion setting in our participants (Holland, 1986, has documented that longer ISIs facilitate stronger occasion setting).

2. For simplicity, we will continue to refer to the two events as an OS and a CC whether they were presented serially or simultaneously.

(Manuscript received February 16, 1999;
revision accepted for publication December 10, 1999.)