Journal of Experimental Psychology: General
2010, Vol. 139, No. 4, 756–771

0096-3445/10/$12.00 DOI: 10.1037/a0020976

Temporal Predictability Facilitates Causal Learning

W. James Greville and Marc J. Buehner
Cardiff University

Temporal predictability refers to the regularity or consistency of the time interval separating events. When encountering repeated instances of causes and effects, we also experience multiple cause–effect temporal intervals. Where this interval is constant it becomes possible to predict when the effect will follow from the cause. In contrast, interval variability entails unpredictability. Three experiments investigated the extent to which temporal predictability contributes to the inductive processes of human causal learning. The authors demonstrated that (a) causal relations with fixed temporal intervals are consistently judged as stronger than those with variable temporal intervals, (b) that causal judgments decline as a function of temporal uncertainty, and (c) that this effect remains undiminished with increased learning time. The results therefore clearly indicate that temporal predictability facilitates causal discovery. The authors considered the implications of their findings for various theoretical perspectives, including associative learning theory, the attribution shift hypothesis, and causal structure models.

Keywords: causality, predictability, contiguity, time, learning

Causal learning is a crucial cognitive process that provides us with the ability to interact with our environment. Creating a representation of the causal structure of the world around us allows us not only to understand and predict the occurrence of events but also to intervene on the world and control our environment, directing our behavior in order to achieve goals and fulfill desires. The key difficulty in understanding how we learn causal structure was identified by the Scottish philosopher David Hume (1739/1888); causal relations cannot be detected directly by our sensory modalities. Instead, we must infer causal relations from the observable evidence that is available to us. Representations of causal relations must therefore be constructed in some way using information about the events that occur in the world around us. Hume proposed that there are crucial "cues to causality" that underpin causal learning and identified the most important determinants as (a) temporal order—causes must precede their effects; (b) contingency—regular co-occurrence of putative causes and effects; and (c) contiguity—the closeness in time and space of these events.

Hume's (1739/1888) first cue of temporal order is perhaps the most fundamental, and its importance is almost unanimously accepted across researchers; causes must occur prior to the effects they produce (although events may not always be perceived in this order, e.g., Waldmann & Holyoak, 1992). The vast majority of the literature on causal learning has thus instead focused on the second cue of contingency and how this information may be used to infer causality. Most researchers agree that the sensory input we receive with regard to the presence or absence of candidate causes and effects is computed in some way to assess the covariation between them, which is then used as a basis for a causal judgment. At the root of most covariation models is the 2×2 contingency matrix, which essentially describes, in the most simple format, the possible combinations of causes and effects. Exactly how this information is computed has been the subject of rigorous debate (see, e.g., Hammond & Paynter, 1983; Perales & Shanks, 2007), and numerous models with varying degrees of complexity have been proposed to account for this computation. A long-standing measure is the ΔP statistic (Jenkins & Ward, 1965), which makes a simple ratio calculation using the four cells of the contingency matrix: $\Delta P = A/(A + B) - C/(C + D) = P(e|c) - P(e|\neg c)$, as shown in Table 1. More recently developed models, for instance Cheng's (1997) power PC theory, have extended ΔP with the intention of accounting for some of the particular phenomena of causal inference that ΔP cannot represent. However, a major shortcoming of all covariation models is that they fail to represent temporal information.

The second of Hume's (1739/1888) tenets, contiguity, has received less empirical attention than contingency. As a result, it is less well understood, and its role in causal learning remains uncertain. According to Hume, temporal contiguity between cause and effect is essential to the process of causal induction. This supposition was corroborated in a systematic investigation by Shanks, Pearson, and Dickinson (1989), demonstrating the crucial role played by contiguity. Their task involved judging how effective pressing the space bar on a keyboard was in causing a triangle to flash on a computer screen. Participants were given a fixed amount of time to engage on the task and could gather evidence through repeatedly pressing the space bar and observing whether or not the outcome occurred. The apparatus was set up to deliver the outcome, with a 0.75 probability when the space bar was pressed. On each trial, if an outcome was scheduled, it would occur after a specific amount of time following the space bar. This interval varied between conditions from 0 s to 16 s. It was found that as the delay increased, participants' causal judgments decreased in systematic fashion. In fact, conditions involving delays of more than 2 s were no longer distinguished as causally effective

W. James Greville and Marc J. Buehner, Department of Psychology, Cardiff University, Cardiff, Wales, UK.

Correspondence concerning this article should be addressed to W. James Greville, Department of Psychology, Cardiff University, Cardiff, Wales, UK. E-mail: grevillewj2@cf.ac.uk

Table 1
Standard 2 × 2 Contingency Matrix

	EFFECT e	
CAUSE c	PRESENT	ABSENT
PRESENT	A e|c	B ¬e|c
ABSENT	C e|¬c	D ¬e|¬c

and were judged just as ineffective as noncontingent control conditions.

However, further research has since expanded on the role of time in causal learning. Buehner and May (2002, 2004) have demonstrated that prior expectation of time frames can alleviate even abolish the detrimental effect of delay, in accordance with Einhorn and Hogarth's (1986) knowledge mediation hypothesis. Greville, Cassar, Johansen, and Buehner (2010) have shown that delays of reinforcement no longer impair instrumental learning when the task environment highlights the underlying contingency structure. Indeed in some circumstances, delays may serve to facilitate causal attribution (Buehner & McGregor, 2006). Such work has suggested that the role of time in causal learning is not merely limited to the contiguity between events but may have a more complex and varied influence on the perception and interpretation of causal relations. It is with this in mind that we turn to address the potential role of temporal predictability. We begin by considering the following anecdote:

> Dave, Jon, and Tom are discussing their morning drives to work. Dave and Jon suffer a similar problem in which they encounter sets of traffic lights that take a very long time to change, even when no cars are coming through on the opposite side. Tom suggests that they try flashing their headlamps at the traffic lights to induce them to change, as he has heard a rumor that they are programmed to respond to the flashing lights of emergency service vehicles. Both take his advice. Dave notices that every time he flashes his headlamps, the traffic lights do in fact change after a consistent delay of around 10 seconds. Jon tries it at the set of lights on his route; sometimes the lights change very quickly, sometimes they take much longer, with little discernible pattern. Jon concludes the lights are operating on a fixed program and his headlamps are not influencing them. Dave, on the other hand, decides that his actions are effective and continues to flash his headlamps when held up at the traffic lights.

The above story is an example of how event timing influences the way in which we learn about causal relations. In the story, it is not the delay between candidate cause and effect on each instance that is important; rather, the overall variation in the timing of events across the set of instances is the determining factor. What eventually forms the basis for decision in this case is the consistency of the temporal interval across multiple events.

This consistency or regularity of the time interval separating events may be referred to as *temporal predictability*. In most situations, causes and effects are rarely encountered as unique, one-off incidences; rather, we experience repeated pairings of cause and effect over time. Over multiple cause–effect instances, we naturally experience multiple cause–effect intervals. When there is a degree of constancy in the duration of intervals, then one may, with some degree of accuracy, be able to *predict* when an effect will occur following administration of the cause. If the temporal interval is fixed and always takes the same value, then the relationship may be said to be maximally predictable. Conversely, if interevent intervals vary from case to case, then predicting future events becomes a much more difficult, if not impossible, task. The greater the variability of the intervals, the more unpredictable the relationship. Under the former scenario, one may develop particular expectations regarding the timing of events, whereas for the latter there is uncertainty as to when an outcome may occur. However, what influence this distinction may have, if any, in the detection or appraisal of causal relations is yet to be fully explored.

Empirical Research

To date, the feature of temporal predictability has received remarkably little empirical attention. One exception is a well-known early study on detecting response-outcome contingencies by Wasserman, Chatlosh, and Neunaber (1983). The authors studied causal learning in a free-operant paradigm, in which a response made during any given trial could increase or decrease the likelihood of a light to illuminate at the end of that trial. Their third experiment contrasted predictable conditions in which they used trial lengths fixed at a constant value of 3 s against unpredictable conditions in which trial lengths could take a value of 1 s, 3 s, or 5 s. Although fixed and variable conditions did not differ significantly, there was a general trend indicating that the variable conditions received uniformly, if marginally, lower ratings than their fixed counterparts. The implication of this study is therefore unclear, and a more in-depth examination of this feature is warranted. Indeed, Wasserman et al. stated:

> Our failure to find significant effects attributable to these factors in no way means that manipulation of the same variables over a broader range of values would also fail to yield reliable results; indeed, we still believe that such work would disclose discernible differences. Our research can thus be seen as a guide to others in their search for potent influences on the perception of response-outcome relations (p. 428).

With a dearth of previous experimental work in which this particular phenomenon has been examined, it is worth casting a more broad glance at findings from the learning literature. The nonsignificant trend in the study described above suggests that, if anything, causal relations with fixed temporal intervals may be seen as more robust than temporally variable relations. However, there is a wealth of evidence from the learning and memory literature that suggests that the reverse may be true.

A long-standing method for the exploration of how relations between responses and outcomes govern behavior is the use of schedules of reinforcement, which specify the input that is required for a reward to be delivered. It is common knowledge in the field of behavior analysis that animals have preferences among such schedules; naturally, it is not surprising that a schedule providing a faster rate of reinforcement, or requiring less input to receive a reward, will be preferred to a slower or more demanding schedule. But certain types of schedules are preferred over others even when the rate of reinforcement is the same. For instance, Cicerone (1976) used a free-operant procedure in which pigeons were presented with two, concurrently available, response keys. Variable-length delay intervals were superimposed on the reinforcers scheduled on one response key, whereas delay intervals of constant length were superimposed on the reinforcers assigned to the other. The results showed that pigeons preferred variable over constant delays of reinforcement and that this preference increased as the range of the interval lengths increased. Many other studies have also indicated that organisms prefer aperiodic over periodic schedules of reinforcement. Andrzejewski et al. (2005) suggest that such effects have arisen from foraging strategies and the utility of behavioral variability.

If animal behavior in conditioning paradigms is based on causal understanding (Blaisdell, Sawa, Leising, & Waldmann, 2006), this might suggest that the animal believes the potential for receiving a reward to be greater in variable than in fixed schedules and, consequently, that the action is more causally effective, under the former than the latter. Although it is clear that performance on such schedules of reinforcement is very different from contingency judgment in action-outcome learning studies with humans, the preference for variable reinforcement shown in nonhuman animals may be indicative of a general facilitatory effect of variability in learning preparations. As Reed (1993) pointed out, although a relationship linking a response to an outcome is not necessarily a reinforcement schedule, it is nevertheless possible that "human perception of the causal efficacy of responses may be influenced by such schedules of outcome presentation in some systematic manner" (p. 328). A consistent preference for variability may well be something that generalizes across learning domains.

An Associative Analysis of Temporal Predictability

The dominant theory of animal behavioral processes is associative learning theory (Mackintosh, 1983; Rescorla & Wagner, 1972). Several authors (e.g., Dickinson, 2001; Dickinson, Shanks, & Evenden, 1984) have pointed out that there are many notable parallels between studies soliciting human judgments of causality and conditioning studies with animals, with a number of factors that influence both classical and instrumental conditioning also having been observed to influence human judgment of causal efficacy (Shanks & Dickinson, 1987). As a consequence, many researchers have attempted to reduce causal reasoning to an associative learning process.

According to associative learning theory, causal relations are represented by the strength of an association between putative causes and effects, which is determined by the increment (or decrement) of associative strength over repeated learning trials. Effects are considered to be reinforcers to a conditioned stimulus or response that is considered as the cause. Traditional associative theories tend to downplay the role of time in learning. Although contiguity has previously been identified as both necessary and sufficient for an association to be acquired (Damianopoulos, 1982; although see Rescorla, 1988), it is not considered that temporal information itself is encoded in this representation. Contiguity is therefore generally seen as merely adjunctive to the learning process. The impact of contiguity on causal learning is addressed by the supposition that the greater the temporal separation, the less associative strength that is acquired (Shanks, 1987; Shanks & Dickinson, 1987). The value of the reinforcer decays over time, so a delayed reinforcer contributes a smaller increment in associative strength compared with an immediate one.

Any anticipated effect of predictability would therefore depend on the rate at which associative strength changes with delay. Research on the phenomena of discounting (e.g., Myerson & Green, 1995) has shown that temporally delayed rewards have less subjective value than immediate rewards. If we suppose that causal induction is also liable to temporal discounting, then we may anticipate that delayed effects lose their capacity to increase the cause–effect association in an analogous manner. Indeed, prior research on reinforcement learning with animals has already suggested that response rates (taken as an indicator of associative strength) follow an exponential decay function with delay (e.g., Chung, 1965). To understand the implications this would have for predictability, consider comparing a fixed-delay with a variable-delay causal relation and further assume that event delays in the variable condition are equally distributed around either side of the fixed-relation midpoint. Due to the hyperbolic shape of the discounting function, the combined associative strength of one contiguous and one delayed event would be greater than that of two fixed-delay events, despite the average cause–effect delay being identical in the fixed and variable relations. In Figure 1, where ΔV is the change in associative strength, this could be expressed as follows: $\Delta V_x + \Delta V_z > 2\Delta V_y$. Consequently, we would expect that an equal amount of immediate and delayed effects would accrue more associative strength overall than one where all the events occur at some fixed intermediate point. Note that for this inequality to hold, the precise shape of the function is unimportant. Green and Myerson (1996), among others, proposed that a hyperbola-like function provides a better fit with experimental data than an exponential decay function; however, any negatively accelerated function would result in the same imbalance in accrued associative strength. Under this assumption, we would thus anticipate that temporally variable conditions would give a stronger impression of causality than predictable conditions, and thus attract higher causal ratings.

However, this prediction might be considered as somewhat counterintuitive. One might be more inclined to expect predict-

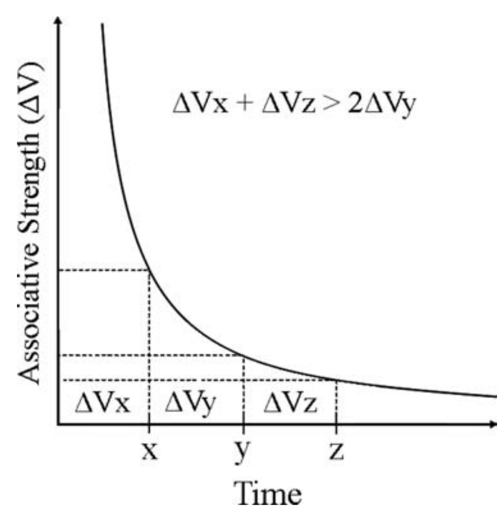

Figure 1. Potential differences in accrued associative strength between fixed-interval and variable-interval conditions according to a hyperbola-like discounting function of delayed events.

ability to be something that might provide confirmatory evidence for a causal relationship, as was the case in the anecdote above. Consistency of the temporal interval separating candidate cause and effect could be taken as symbolic of a genuine relationship between them, in much the same way as statistical co-occurrence. If causes are hypothesized to bring about their effects by means of a particular mechanism or sequence of events, it seems reasonable to suggest that (provided the mechanism remains unaltered) there should be a degree of regularity in the timeframe over which these events unfold. We therefore turn now to consider other theories of causal learning from which may be derived predictions in concordance with this intuition.

Attribution Shift Hypothesis

From a covariation perspective of causal learning, a potential explanation for the effect of predictability is the attribution shift hypothesis (Shanks & Dickinson, 1987). This has previously been proffered as an account for the detrimental effect of delay. Under this assumption, a delayed action–outcome pairing is perceived not as a cause–effect pairing, $c \rightarrow e$, but instead as one instance of an action with no outcome, $c \rightarrow \neg e$, and an outcome following no action, $\neg c \rightarrow e$, as illustrated in Figure 2. In terms of the 2×2 contingency matrix, this may be described as one instance of Cell B and one instance of Cell C rather than a single instance of Cell A.

However, this process is highly dependent on the size of the "temporal window" that is adopted for event parsing. If a reasoner assumes a more relaxed time frame over which events may unfold, this enables temporally distal effects to be correctly attributed to the candidate cause rather than disregarded as spurious. Previous work (e.g., Buehner, 2005) has suggested that prior knowledge about existing causal mechanisms can lead to the adjustment of this temporal window in this manner. In similar fashion, if the reasoner repeatedly encounters evidence that is contradictory to their initial time frame expectations, they may revise their assumptions and adopt a new, more lenient temporal window. Thus, if the cause and effect are temporally separated, but this interval is constant, this may be recognized over repeated instances and avoid the delayed effects being subjected to attribution shift. Temporal predictability, therefore, may enable a learner to bridge temporal gaps in causal induction through repeated exposure to the same temporal interval. In contrast, a variable interval might preclude recognition of the statistical regularity between cause and effect, which in turn would mean that actual cause–effect pairings will be parsed as instances of Cells B and C. The attribution shift hypothesis is therefore capable of forecasting an advantage for predictability through the reduction of erroneous attribution of delayed effects to random background processes.

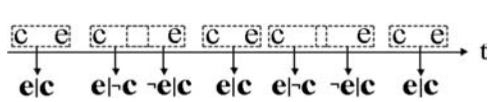

Figure 2. The effect of attribution shift in parsing an event stream with a fixed temporal window: $c \rightarrow e$ intervals that are longer than the temporal window simultaneously decrease impressions of $P(e|c)$ and increase impressions of $P(e|\neg c)$.

Causal Structure Models

One final perspective takes a broader, more holistic viewpoint on the causal learning process. Causal structure models (Griffiths & Tenenbaum, 2005, 2009) combine both bottom-up empirical processes, by which statistical inference from observable evidence forms the basis for causal induction, with top-down modulation in the form of preexisting ideas of causal structure based on prior knowledge, experience, and expectations. Under this framework, causal inference can take place when there is insufficient statistical information available for other models (such as ΔP) to calculate a metric of covariation. Structure models emphasize causal structure over causal strength, proposing that the primary decision in causal reasoning is whether there *exists* a causal relationship between two events, before assessing the *extent* of any such relationship. This involves essentially a binary decision between two hypotheses: H_0, in which there is no causal relation between cause c and effect e, and e instead occurs solely due to the influence of random background processes b; and H_1, where c has the generative power to produce e (and b still also produces e). Causal inference is then made by assessing the likelihood of obtaining an observed cause–effect sequence against these two hypotheses using Bayes' rule. Learning to impose structure on the world of sensation crucially depends on our ability to identify regularities, patterns, and consistencies in the environment that we can piece together to produce a coherent picture. Constancy of a temporal separation is a further regularity in the environment that an organism may be able to detect and use to construct an accurate representation of causality. From this Bayesian structure perspective, temporal predictability would serve to facilitate causal learning because temporal regularity between putative cause and effect is much more likely if there exists a causal relation than if no such relation exists (and the repeated regularity occurs by chance).

In summary, it is evident that temporal variability (or predictability) has the potential to be added as a fourth cue to causality (in addition to temporal priority, contiguity, and contingency). Given that empirical data about it is sparse and ambiguous, and that different theoretical perspectives allow contrasting predictions, the remainder of the present article is dedicated to an experimental analysis of its role on causal inference.

Experiment 1

Experiment 1 was modeled closely on the paradigm used by Shanks et al. (1989), which was also replicated by Reed (1993) and Buehner and May (2003). In each condition, participants were presented with a triangle on the screen and a button labeled *PRESS* just beneath it. Participants were instructed that their task was to investigate the extent to which their action (clicking on the button) could cause something to happen on a computer screen (the triangle lighting up).

Participants engaged in a free-operant procedure (FOP), meaning that they were free to choose whether and when to respond throughout the duration of the condition. Previous studies have found scheduling of response-outcome contingencies on an FOP to be a highly sensitive and unbiased method of investigating causal learning (Shanks et al., 1989; Wasserman et al., 1983). However, in many such studies, the learning experience is segmented into predefined "response bins" or learning trials (e.g., of 1-s duration).

If a response is made during this time bin, then it is reinforced at the end of the period. However, it is of course possible that the participant may respond again during the time between a reinforced response and the consequent outcome. This, and any further responses, would then go unreinforced. Consequently, such a procedure fails when participants respond at a faster rate than that corresponding to the predefined bin size, as only the first response within each bin will have the potential to produce an outcome. This was pointed out by Buehner and May (2003), who demonstrated that action-outcome delays in a standard FOP change $P(e|c)$ and $P(e|\neg c)$ so that the actual contingency experienced by the participant is lower on delayed than on immediate conditions. Furthermore, and of crucial importance for scrutinizing the influence of temporal predictability, using this underlying trial structure means that full control over the cause–effect interval cannot be maintained; whereas trial length can be held constant, a participant may respond at any point during this trial, hence the interval between action and outcome may still vary. Wasserman et al.'s third experiment should therefore more accurately be considered as a comparison of low variability against high variability, rather than predictability against variability.

To avoid these problems, we did not use predefined learning trials or time bins in our experiments; instead, every response had the potential to generate an effect, regardless of when it was made. We used the same response-outcome contingency as Shanks et al. (1989): Every press of the button had a 75% chance of producing the outcome. If an outcome was scheduled, then the effect occurred following the programmed temporal delay, and we were able to fully control the temporal interval separating a cause from its generated effect. Thus, we were able to manipulate temporal variability and delay across conditions while keeping constant the *objective* contingencies. We are fully aware that this trial-free instrumental procedure carries inherent ambiguity with respect to matching individual responses to individual outcomes. For instance, a participant could perform several responses in quick succession and then observe a corresponding burst of effects after the relevant delay. It would be difficult to match individual responses to specific effects, and this would be amplified when the cause–effect interval is variable. Importantly, however, by allowing each response to produce the effect (without limitations imposed by trial structures), the overall objective contingency will remain unaffected by variations in delay and variability of delay. Whether the subjective impression of contingency (and in this case also causality) remains unaltered by our manipulations is of course a different question altogether, and is in fact at the heart of the research reported here.

We used two fixed delays—2 s and 4 s, and three different types of temporal predictability. The first was a fixed, predetermined delay that remained constant throughout a given condition, and thus constituted maximal predictability. However, most natural causal relations rarely involve precise and perfectly predictable cause–effect delays. Epidemiologists, for instance, have long postulated that disease outbreak follows infection after an incubation period described by a log-normal distribution (Evans, 1993) centered around a mean expected wait time. Consequently, our second type of temporal predictability used cause–effect intervals sampled from a normal distribution, centered around a midpoint corresponding to one of the fixed intervals (see the Method section below for more details). Finally, as a maximally uncertain control, we used a uniform random distribution, in which the delay could take any value within a predefined range, with an equal probability of taking any particular value. Importantly, these manipulations are distinct from Experiment 3 of Wasserman et al. (1983); rather than restricting intervals to a small set of fixed values, we instead allowed intervals to vary freely across a continuum.

Most real-world causal relations are assessed against a background of alternative causes. For instance, although an illness may be the cause of a headache, a headache could also potentially arise as a result of stress, tiredness, or dehydration. Identifying the crucial relation from other spurious connections is a fundamental part of the induction process. In order to preserve ecological validity in this respect, we also introduced three different levels of background effects in our paradigm. This was done by scheduling the effect to occur a predefined number of times, independently of the participant's action, at random points in time during the condition.

Method

Participants. Thirty undergraduate students with a median and modal age of 19 years were recruited via an online participation panel hosted at Cardiff University, Cardiff, Wales, United Kingdom. They received either £4 payment (about $6 U.S.) or partial course credit for participation.

Design. Three factors were manipulated in this experiment—temporal distribution, background effects, and delay. Temporal distribution had the levels fixed, normal, and random; background effects had the levels zero, low, and high; delay had the levels 2 s and 4 s. Factorial combination of these levels resulted in a $3 \times 3 \times 2$ within-subjects design, producing 18 different conditions each of 90-s duration.

The probability of an outcome following an action, $P(e|c)$, was .75 throughout all conditions. Note that this probability was not defined relative to a particular unit of time; instead, each button press had a 75% chance of causing the triangle to flash. If an event was generated, then the effect occurred after the appropriate temporal interval had elapsed.

The three types of temporal distribution provided a manipulation of predictability by controlling the variation of the temporal intervals in each condition. The interval for any given action–outcome pairing was determined according to the particular combination of delay and temporal distribution. In the fixed conditions, the temporal interval was always the same, held at a constant value within the condition (i.e., 2 s or 4 s). These values then served as "midpoints" for the comparable normal and random conditions. For the random conditions, the temporal interval for any given cause–effect pair was given by generating a random value within the specified range. So, for example, in the "Random-2" condition, the interval could take any value between 0 s and 4 s, with any value equally as likely to occur as another. For the normal conditions, the delay was specified according to a normal probability distribution with a range of 4 s, centered around the midpoint. So, for example, in the "Normal 4" condition, interval lengths were drawn from a normal distribution centered around 4 s, with minima and maxima of 2 s and 6 s. Accordingly, values closer to the midpoint of 4 s were more likely than values closer to the extreme boundaries of 2 s and 6 s. Thus, the delay variance for normal conditions should be smaller with respect to the random conditions.

In addition to this, three levels of noncontingent "background" effects were used, where the outcome occurred independently of the response. As a baseline, a zero rate of background effects was first applied—the effect did not occur in the absence of the cause and $P(e|\neg c) = 0$. In addition, a medium rate, equivalent to 1 effect every 10 s and a high rate equivalent to 1 every 5 s, was created. With a total condition time of 90 s, this gave nine and 18 background effects in total for the medium and high levels, respectively, which were distributed randomly throughout the condition.

Two questions were used as dependent measures to gauge participants' impressions of causal strength. One was based on a covariational understanding of causality couched within a counterfactual question: "Imagine you had pressed the button 100 times in this condition. How many of these 100 presses would have caused the triangle to light up?" The other was slightly more ambiguous and was aimed to appeal to the degree of perceived control beyond pure covariation: "Overall, to what extent do you feel pressing the button controlled the triangle lighting up in this condition?"

Participants provided a rating between 0 and 100 for both questions.

Apparatus, materials, and procedure. The experiment was programmed in Python 2.4 and conducted on Apple Macintosh computers situated in individual testing booths. Participants used the mouse to click on the *PRESS* button and used the keyboard to type in their responses at the end of each condition. After being welcomed by the experimenter and giving consent to participate, participants read on-screen instructions that outlined the nature of the task.

In each condition, a triangle was presented in the center of the screen, along with a button that participants were able to press, by clicking on it with the mouse. If a response triggered an outcome, then the triangle lit up for 250 ms. Participants engaged in 18 different FOPs, as described above, presented in a random order, with each condition lasting 90 s. At the end of each, the screen cleared and participants were asked to respond to the two questions described previously. Participants then typed in their answers into the appropriate text box and clicked on the *SUBMIT* button to proceed to the next condition. In total, the experiment lasted around 35 min.

Results and Discussion

Causal judgments. We adopted a significance level of .05 for all analyses in this and the following experiments, except where otherwise noted. Two different questions were posed at the end of each condition, intending to try and capture fully all aspects of the participants' causal impressions. The "contingency" question is a well-established measure that has been used in many previous studies (Buehner, Cheng, & Clifford, 2003). The "control" question, meanwhile, was more ambiguous, which may propel participants to take temporal information into account in providing their rating, and thus may provide a more useful measure for capturing any influence of predictability. Accordingly, it seems appropriate to focus initially on this latter measure. Figure 3 shows mean ratings provided by participants for the control question, for all 18 conditions. As expected, ratings were considerably higher in the

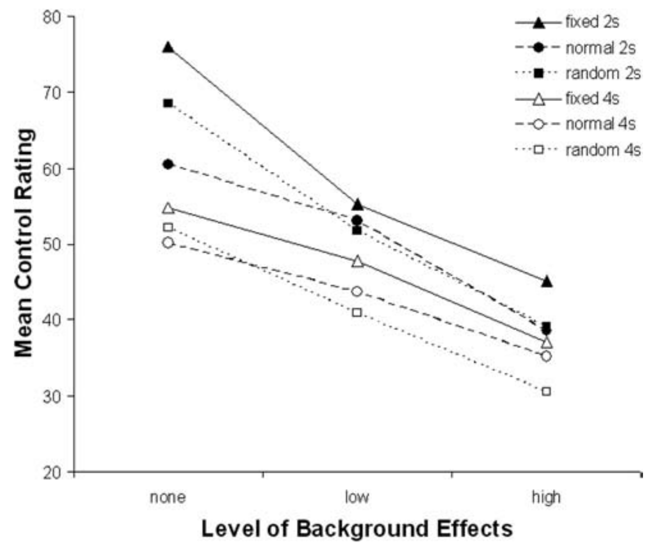

Figure 3. Mean control ratings for all conditions in Experiment 1 as a function of background effects. Filled and unfilled symbols refer to mean delays of 2 s and 4 s, respectively. Delay variability is noted by different symbol and line styles.

shorter delay compared with the longer delay conditions. Also in accordance with previous findings, ratings declined as the rate of background effects increased. The effect of temporal predictability, which is the factor of principal interest, is less immediately apparent. It can, however, be seen that the fixed conditions consistently received higher causal ratings than their normal and randomly distributed counterparts, while there appeared to be little difference between the two distributed conditions.

A $3 \times 3 \times 2$ within-subjects repeated measures analyses of variance (ANOVAs) corroborated these impressions, finding significant main effects of temporal distribution, $F(2, 58) = 3.37$, $MSE = 626$, $\eta_p^2 = 0.104$; delay, $F(1, 29) = 22.85$, $MSE = 714$, $\eta_p^2 = 0.441$; and background effects, $F(2, 58) = 27.97$, $MSE = 799$, $\eta_p^2 = 0.491$. None of the possible interactions were significant. Further investigation of our main manipulation of interest—temporal variability—using orthogonal Helmert contrasts confirmed that fixed conditions ($M = 52.71$, $SD = 12.19$) received significantly higher ratings than normal ($M = 46.86$, $SD = 11.57$) or random ($M = 47.19$, $SD = 9.85$) conditions, $F(1, 29) = 4.98$, $MSE = 1271$, $\eta_p^2 = 0.147$, which in turn did not differ significantly from each other, $F(1, 29) = 0.007$, $MSE = 810$.

Participants' ratings for the contingency question followed the exact same pattern as for the control question, with significant main effects for temporal distribution, $F(2, 58) = 3.78$, $MSE = 573$, $\eta_p^2 = .115$; delay, $F(1, 29) = 23.10$, $MSE = 661$, $\eta_p^2 = 0.443$; and background effects, $F(2, 58) = 11.46$, $MSE = 593$, $\eta_p^2 = 0.283$. Participants therefore apparently made little distinction between the two dependent measures, with both eliciting similar responses. Indeed, close inspection reveals they were treated as identical by a considerable proportion of participants, with scores matched in over one third of the total cases. We therefore decided to deploy only a single dependent measure in subsequent experiments.

Instrumental behavior and outcome patterns. Table 2 shows the behavioral data from the first experiment, for each of the

Table 2
Behavioral Data for Experiment 1

Variable	Level of backgrounds effects	Fixed 2 s	Fixed 4 s	Temporal distribution Normal 2 s	Temporal distribution Normal 4 s	Random 2 s	Random 4 s
Mean response rate (/min)	Zero	21.53	21.78	22.44	20.13	21.31	17.93
	Low	19.02	18.07	16.64	21.49	18.44	19.22
	High	17.13	18.69	16.58	20.18	21.91	18.98
Mean outcome rate (/min)	Zero	16.42	16.36	16.38	15.20	16.40	13.56
	Low	13.82	13.60	12.82	16.11	13.96	14.60
	High	12.51	13.62	12.47	14.91	16.33	14.80
Actual $P(e\|c)$	Zero	0.772	0.751	0.733	0.768	0.768	0.754
	Low	0.730	0.749	0.770	0.758	0.761	0.753
	High	0.743	0.734	0.754	0.733	0.756	0.794
Mean actual delay (ms)	Zero	2000 (0)	4000 (0)	1992 (167)	4000 (203)	2071 (249)	4016 (237)
	Low	2000 (0)	4000 (0)	2027 (202)	4036 (188)	2017 (325)	3949 (290)
	High	2000 (0)	4000 (0)	2072 (172)	4046 (136)	2049 (273)	3948 (327)
Mean control rating	Zero	76.10 (26.51)	54.87 (21.60)	60.57 (17.40)	50.10 (23.39)	68.53 (22.71)	52.27 (21.79)
	Low	55.20 (21.87)	47.80 (28.41)	53.07 (22.32)	43.67 (27.90)	51.77 (18.19)	40.87 (27.85)
	High	45.17 (21.55)	37.10 (24.27)	38.60 (22.93)	35.17 (24.51)	39.23 (21.80)	30.47 (20.31)
Mean contingency rating	Zero	75.87 (26.54)	50.43 (23.83)	58.37 (19.67)	44.07 (25.49)	65.40 (24.06)	52.20 (19.75)
	Low	57.93 (20.37)	53.13 (28.00)	58.93 (18.31)	53.57 (24.13)	55.47 (20.53)	45.17 (26.46)
	High	56.00 (26.69)	45.57 (30.49)	46.80 (23.62)	42.80 (27.36)	45.53 (25.66)	37.67 (22.70)

Note. Standard deviations appear in parentheses.

18 conditions. This includes response rate (i.e., mean presses per minute) within each condition and the corresponding rate of effects (outcome density). The experienced $P(e|c)$ is also shown, calculated as the proportion of responses that generated an effect (ignoring background effects), for each participant in each condition. The mean interval between cause and effect was likewise computed, and is shown with the standard deviation, as an indication of temporal interval variance, in parentheses. In addition, the mean ratings provided for the contingency and control questions are also provided, again with standard deviations in parentheses.

Although the number of responses produced is fairly consistent across conditions, it appears that conditions without background effects produced the highest response rates in general, whereas the Random-4 conditions (random distribution, 4-s delay) received lower response rates. If, for some reason, different conditions are producing different response rates in participants, then the effect of our manipulation may not be directly on causal rating but instead mediated through changes in response (and subsequent outcome) density. It was thus necessary to verify whether our independent variables influenced ratings indirectly by exerting an effect on behavior. In addition, some fluctuations in the actual delay and $P(e|c)$ from the programmed values are also expected; although these were assumed to eventually cancel out throughout the course of each condition (and certainly across participants), it is possible that differences between conditions could remain and be driving any observed differences in causal ratings.

To address these concerns, we carried out 2×3 within-subjects repeated measures ANOVAs on the data derived from participants' instrumental behavior. Due to a small number of participants responding at a very high rate, the distribution of data for response and outcome rate is positively skewed; hence, we normalized response rates by taking the square root. No significant effects of temporal distribution, $F(2, 52) = 1.95$, $MSE = 1.09$; delay, $F(1, 26) = 0.002$, $MSE = 1.83$; or background effects, $F(2, 52) = 2.67$, $MSE = 1.05$, were found on response rate. Mean delay naturally differed between different delay conditions, but it was not significantly affected by either temporal distribution or background effects (both ps > 0.4). Actual $P(e|c)$ was also unaffected by all three independent variables (all ps > 0.2). Participants' causal judgments were therefore not impacted by uncontrolled differences in instrumental behavior or deviations from programmed values.

Our results thus replicate the well-established findings that (a) in the absence of delay expectations, cause–effect delays are detrimental to learning and (b) decreasing the contingency by introducing effects caused by alternative background causes likewise reduce causal ratings. We are thus confident that our paradigm is reliable. Of central interest, however, was the influence of temporal predictability. Although our analyses confirmed that holding the cause–effect interval constant facilitated causal attribution, in line with predictions derived from causal structure models, we did not find any distinction between the normal and random conditions. Arguably, normally distributed delays could have been expected to elicit higher ratings than their uniformly distributed random counterparts, due to the smaller variability of delay in the former compared with the latter. A possible explanation for the failure to find a significant difference here is that the normal and random conditions were much more similar to each other than either was to the fixed conditions. Although the fixed conditions had no variability of delay, for the two distributed conditions, there was a maximum range of 4 s within which the effect could occur following a reinforced response, the only difference between these two being the likelihood of the effect occurring at a particular point within this range. Rather than increasing or decreasing the temporal range within which an effect could occur, we changed the probability distribution according to which any given temporal interval was determined. Although the variance of the delay was greater for random than for normal conditions (see Table 2), the maximum range of interval variability was the same for each. It therefore seems pertinent to investigate the effect of modifying temporal predictability by instead varying the size of the interval range. Will an increase in interval variability, and concomitant unpredictability, lead to a corresponding decline in causal evaluations? We sought to address this question in Experiment 2.

Experiment 2

In Experiment 2, we implemented variations in the degree of predictability by modifying the range over which intervals could vary, rather than the type of distribution from which they were drawn. We also improved the paradigm as follows: First, we increased the time participants could learn about each causal relation from 90 s to 120 s, comparable to earlier studies (Buehner & May, 2002, 2003; Shanks et al., 1989). In Experiment 1, we used a shorter exposure time primarily to prevent participant fatigue when working though 18 conditions. Because this experiment involved fewer conditions, we felt it safe to increase exposure time. Second, we presented just a single question of causal effectiveness as the dependent measure. We found no systematic differences in Experiment 1 between the two questions, so the focus on one question is economical both in terms of participant time and analysis.

Finally, we removed all background effects from this study, because Experiment 1 showed that, although producing the expected main effect, they do not interact with either delay or predictability. However, previous studies on operant causal learning always used paradigms, including conditions in which $P(e|\neg c)$ has some nonzero value (Shanks et al., 1989; Wasserman et al., 1983) or noncontingent-yoked conditions in which outcomes were predetermined and responding was ineffective (Reed, 1993; Shanks & Dickinson, 1991). Both manipulations guarantee that participants will experience situations in which the outcome occurs independently of their actions, creating an element of uncertainty as to whether an outcome that occurs is due to their action or to alternate causes. We felt that we needed to include this element of uncertainty in order to ensure that our task is not trivial.

Thus, we introduced noncontingent conditions using a yoking technique. Specifically, we played back outcome sequences that were generated by a previous participant's performance during a pilot experiment. In these conditions, the action of pressing the button had no causal efficacy itself, and the effects that occurred were therefore noncontingent on the current participant's behavior. Yoking to outcome patterns from a pilot experiment, rather than to participants' own behavior in the present experiment, was preferred for two reasons. First, yoking to one's own behavior places considerable restriction on the ordering of conditions, because a yoked condition cannot be presented until a participant has worked through the corresponding master condition. Second, it is very possible that participants might notice that the same outcome stream they previously generated is being played back to them, particularly if they are responding in a structured way (such as using response bursts or specific patterns of responding), and this would therefore make the task trivial.

Method

Participants. Sixty undergraduate students from Cardiff University, with a median and modal age of 20 years, were recruited via an online participation panel. Either £4 payment (about $6 U.S.) or partial course credit was awarded for participation.

Design. Three independent variables were manipulated: condition (either master or yoked), mean programmed delay, and range of temporal interval values as a measure of predictability. In similar fashion to the random conditions in Experiment 1, the value of a temporal interval on any given cause–effect pairing could take any value within the defined range, with uniform probability across the range. The wider the allowable range of temporal interval values, the greater the variation in the value that a temporal interval could take on any one particular cause–effect instance, and thus the greater the variability of temporal intervals throughout the experimental condition.

Delay had two levels, 3 s and 6 s. Range had three values: 0 s, which meant that there was no variation in the temporal intervals and the value was fixed, so the cause–effect interval was constant throughout the condition; 3 s, which meant the temporal interval on any one cause–effect instance could take any value within a range of 3 s across the central midpoint value of delay, or in other words 1.5 s either side of the midpoint; and 6s, which meant temporal interval could take any value within 3 s either side of the programmed mean delay. Thus, for instance, in the 3-s range 3-s delay condition, cause–effect intervals could take on any value between 1.5 s and 4.5 s. Combining all levels of delay and range provided six master conditions. Condition was either yoked or mastered. The yoked conditions were generated from previous participants' contributions to the six conditions in a pilot experiment (which were identical to the six master conditions). Each participant in Experiment 2 was paired randomly (with replacement) with a pilot participant; the outcome pattern generated by this pilot participant was then played back in the yoked conditions, and responses made during these yoked conditions were ineffective in influencing the outcome pattern.

Factorial combination of delay, range and condition in a 2 × 2 × 3 fully within-subjects design produced 12 different conditions. The first condition presented was always a master condition, and which of the six master conditions was selected as the first was counterbalanced across participants. The remaining conditions were then presented in random order.

Apparatus, materials, and procedure. The experiment took place in a large computer lab. Participants were tested in small groups, seated in a quiet area of the lab to work on the task. Each participant used a PC running Windows XP and Python version 2.4.1, with a 19-in. LCD widescreen display. The paradigm was a straightforward adaptation from the previous study, with the visual appearance in terms of size and shape of stimuli and the speed of stimulus presentation consistent with Experiment 1. The basic experience for participants was thus virtually identical to that from Experiment 1, except that condition time was extended to 120 s, and we opted to use only a single dependent measure: "On a scale of 0–100, how effective was pressing the button at causing the triangle to light up?" As in the previous experiment, participants used the mouse to click on the button and the keyboard to type in responses. The experiment took approximately 15 min to complete.

Results and Discussion

Causal ratings. Figure 4 shows mean causal ratings for Experiment 2. First, there is a very clear distinction between ratings

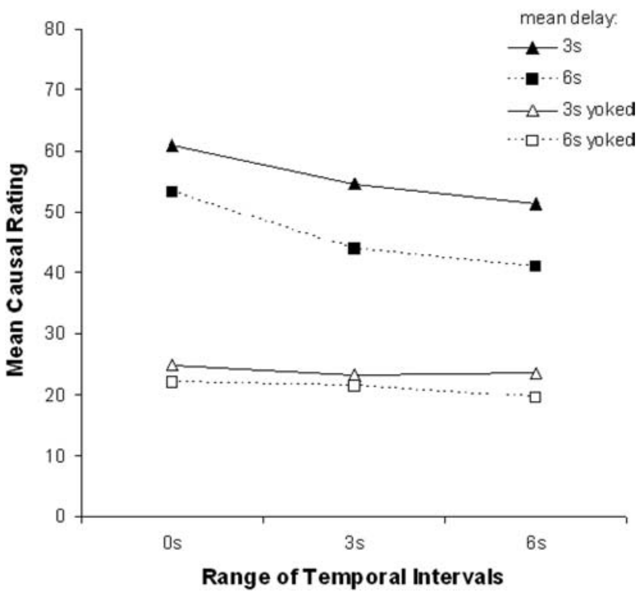

Figure 4. Mean causal ratings from Experiment 2 as a function of interval range. Filled and unfilled symbols refer to master and yoked conditions, respectively. Mean delays are noted by different symbol and line styles.

for the master and the yoked conditions. This indicates that participants had little difficulty in correctly distinguishing the contingent and noncontingent causal relations within the experimental set. The yoked conditions themselves all appear to have elicited very similar, low causal ratings, as expected, because there is no connection between response and outcome. The fact that ratings are above zero is likely attributable to the occasional random coincidence of participants' responses with the preprogrammed outcomes, or a reluctance to endorse ratings at the extreme end of the scale.

Of primary interest, however, are the master conditions, in which delay and delay variability actually affected the timing of outcome following responses, and consequently our analyses focused on these only. As can be seen in Figure 4, range exerted a linear influence on causal ratings, $F(1, 59) = 10.97$, $MSE = 651$, $\eta_p^2 = 0.157$, confirming that judgments of causal effectiveness declined as a function of interval range and thus temporal uncertainty. The effect of delay is also immediately apparent, with short-delay conditions receiving uniformly higher ratings than the long delay, $F(1, 59) = 14.07$, $MSE = 590$, $\eta_p^2 = 0.193$, in line with Experiment 1 and prior research. There was no significant interaction between range and delay, $F(2, 118) = 0.19$, $MSE = 444$.

Instrumental behavior and outcome patterns. Table 3 shows the behavioral data for the six master conditions in Experiment 2. We used 2 × 3 within-subjects ANOVAs and found that actual $P(e|c)$ remained unaffected by both range and delay (both ps > 0.5), and mean delay was unaffected by range, $F(2, 118) = 0.32$, $MSE = 7.02$. We are thus confident that the programmed manipulations delivered the appropriate event streams to participants. Response rates (normalized by taking the square root) were not significantly influenced by range, $F(2, 118) = 0.46$, $MSE = 1.92$; however, there was a significant effect on response rate of delay, $F(1, 59) = 5.20$, $MSE = 1.61$, $\eta_p^2 = 0.088$. An inspection of Table 3 suggests that response rate was slightly lower in the long-delay conditions, in line with previous reports (cf. Buehner & May, 2003).

Experiment 2 has therefore provided a clear illustration that temporally predictable cause–effect relations are perceived as more causal compared with variable and unpredictable relations. Furthermore, increasing temporal variability within unpredictable relations results in a corresponding linear decrease in causal judgments. This is the first time, as far as we are aware, that this has been demonstrated in a free-operant response-outcome learning task. It would appear, therefore, that our results are more in line with attribution shift or causal structure models and problematic for associative perspectives on causal learning.

However, our results need not altogether be incompatible with findings from reinforcement learning; there remains an alternative explanation that must be explored. Drawing on the wider literature on learning and memory, it has been widely reported that the progression of learning is highly dependent on the type of training or practice undergone. In particular with regard to motor learning and skill acquisition, researchers have compared constant practice, where participants practice using a consistent set of materials and skills, with variable practice, where performance takes place in a variety of different conditions. Constant practice generally produces better performance in the short term, whereas variable

Table 3
Behavioral Data for Experiment 2

	Delay					
	3 s			6 s		
	Range of temporal intervals					
Variable	0 s	3 s	6 s	0 s	3 s	6 s
Mean response rate (/min)	21.83	19.06	19.51	18.01	19.08	18.28
Mean outcome rate (/min)	16.31	14.04	14.86	13.62	14.32	13.71
Actual $P(e\|c)$	0.750	0.745	0.784	0.759	0.756	0.750
Mean actual delay (ms)	3000 (0)	2968 (213)	2993 (359)	6000 (0)	5960 (208)	5969 (566)
Mean causal rating	60.92 (29.72)	54.63 (26.02)	51.28 (30.04)	53.20 (32.66)	43.80 (28.51)	41.02 (28.05)

Note. Standard deviations appear in parentheses.

practice leads to better retention in the long run (Gluck, Mercado, & Myers, 2008). Thus, although learning under consistent conditions may initially result in more rapid acquisition, over time, variable conditions result in the formation of stronger associations. According to Schmidt (1975), variations in practice of a motor skill result in superior learning, which is demonstrated by a better ability to transfer the skill to different contexts. Wulf and Schmidt (1997), for example, found that performance on a continuous pursuit tracking task in transfer tests with novel scaling was generally enhanced by variable compared with constant practice. Until fairly recently, though, there has been little interest in whether this finding generalizes to higher level cognitive tasks. However, Goode, Geraci, and Roediger (2008) investigated the effects of constant versus variable practice on performance with the verbal priming task of anagram solution. The results from this study showed that although initially a greater proportion of anagrams were correctly solved following constant rather than variable practice, by the third practice session this trend had reversed.

Thus, there is converging evidence from a range of learning paradigms and contexts for a facilitatory effect of variability, provided enough learning time is provided. We fully recognize that causal or contingency learning is very different from motor skill acquisition. Nonetheless, we took inspiration from this literature to explore the possibility of an analogous role of temporal variability with respect to causal learning. Specifically, we asked whether learning reaches asymptote faster with consistent temporal intervals, compared with variable ones, and whether perhaps any apparent advantage conferred by temporal predictability is simply due to learning having failed to reach asymptote for the variable conditions in the time we provided. This short-term advantage for predictability may then disappear over long enough learning trials, and even be reversed in the long run.

However, a computational perspective might instead suggest that, if anything, temporal predictability may have more of an impact as learning progresses: Increasing learning time is likely to enhance any potential temporal contribution to a mental computation of causality, because more temporal information becomes available over extended learning periods. Moreover, temporal predictability is only capable of exerting an influence when an observer experiences multiple intervals. The more cause–effect intervals a reasoner experiences during a learning period, the greater the total amount of variation that may be experienced, and the more apparent a distinction between a predictable, fixed relation and a variable, unpredictable relation may become. We endeavored to examine these two opposing hypotheses in Experiment 3.

Experiment 3

The previous experiments have clearly demonstrated a facilitatory effect of temporal predictability in causal learning. However, a possible consideration in the interpretation of these results is that the rate of acquisition may differ with temporally predictable conditions compared with temporally variable conditions. Variable–interval causal relations may take longer to discover but may then lead to formation of a stronger associative bond, and thus prove more resilient to extinction. If enough learning time is provided, then it might be expected that judgments of causal strength for temporally variable causal relations should match or even exceed those for temporally predictable conditions.

Accordingly, we set out to investigate the influence of condition time on temporal predictability in a free-operant causal learning experiment. If, as might be suggested by associative accounts, the effect of predictability observed thus far is merely a failure of learning to reach asymptote, then increasing condition time should bring causal ratings for variable conditions in line with predictable conditions.

To investigate this question, we added conditions that lasted double the length of time as those in previous experiments (i.e., 4 min) and contrasted these with corresponding 2-min conditions. If the "failure to reach asymptote" argument holds, we should expect some reduction in the difference between predictable and variable temporal relations for the 4-min conditions with respect to the 2-min conditions. We might even see the variable conditions judged as stronger if in fact variability leads to the formation of stronger associations provided enough learning time is allowed, as might be suggested from the literature on variability of practice. Experiment 3 thus served as a sterner test of the influence of temporal predictability.

Method

Participants. Thirty-three undergraduate psychology students based at Cardiff University, with a median and modal age of 19 years, were recruited via an online participation panel and received partial course credit for completing the experiment.

Design. This experiment introduced exposure time (to each condition) as an additional factor. Two levels were used; 2 min, to be consistent with methods obtained thus far and attempts to replicate the findings, and 4 min, which by doubling the sampling opportunity were considered ample time for participants to fully investigate, discover, and make a judgment on any causal relationship that might exist. Delay and range remained as factors, although to simplify and condense the experiment, the "intermediate" level of temporal interval range was removed (i.e., 3 s). Thus, there were two levels of delay, 3 s and 6 s, and two levels of range –0 s (fixed, predictable) and 6 s (variable, unpredictable). These factors combined to produce eight different conditions, all of which were experienced by each participant, thus providing a $2 \times 2 \times 2$ fully within-subjects design. The order of which condition was experienced first was counterbalanced across participants, with the remaining conditions then appearing in random order. Participants provided causal ratings from 0 to 100 at the end of each condition as the dependent measure.

In order to add a degree of difficulty to the task and avoid making the contingency too apparent, a steady rate of noncontingent background effects was applied to each condition. This was equivalent to one every 10 s, and each effect could occur at any point within a given 10-s segment (i.e., the first background effect could occur somewhere between 0 s and 10 s, the next between 10 s and 20 s, and so on). Yoked conditions could instead have been implemented as in Experiment 2, but given that there were eight master conditions, it seemed that matching each of these with a noncontingent condition would be somewhat uneconomical, and a more streamlined experiment would be less tedious for participants.

Apparatus, materials, and procedure. The experiment was conducted in a small computer lab, using identical apparatus as for Experiment 2, and was once again developed and run using the

Python programming language. Participants were tested in small groups, seated at individual workstations that were screened off from each other. The paradigm and procedure were identical to those of the previous experiments, using the same visual stimuli and layout, with only the key differences described above, and corresponding modifications to the instructions informing participants that they would experience conditions of different durations.

Results and Discussion

Causal ratings. Figure 5 summarizes the results from Experiment 3. As can be clearly seen, we once again have a noticeable influence of range, with a decline in ratings evident with all but one of the temporally variable conditions compared with the corresponding temporally predictable conditions (with the same combination of delay and condition time), and an overall significant main effect of range, $F(1, 32) = 6.13$, $MSE = 571$, $\eta_p^2 = 0.161$. Delay also again has an immediately apparent influence, with the 3-s conditions receiving significantly higher ratings than 6-s conditions, $F(1, 32) = 5.15$, $MSE = 823$, $\eta_p^2 = 0.139$. Of central interest in this experiment, we see that there is no significant influence of the duration of the experimental conditions, $F(1, 32) = 0.80$, $MSE = 694$, and crucially no significant Range $\times$ Duration interaction, $F(1, 32) = 2.26$, $MSE = 588$, confirming that the advantage for predictability over variability is maintained for the longer (4-min) conditions. None of the other possible interactions were significant.

Instrumental behavior and outcome patterns. Table 4 shows the behavioral data from Experiment 3. As can be seen, response rates were fairly consistent across levels of range and delay, though naturally there were more responses in total in the 4-min conditions than in the 2-min conditions. We used within-subjects ANOVAs and found that response rate (square-rooted), mean experienced delay, and actual $P(e|c)$ were not significantly

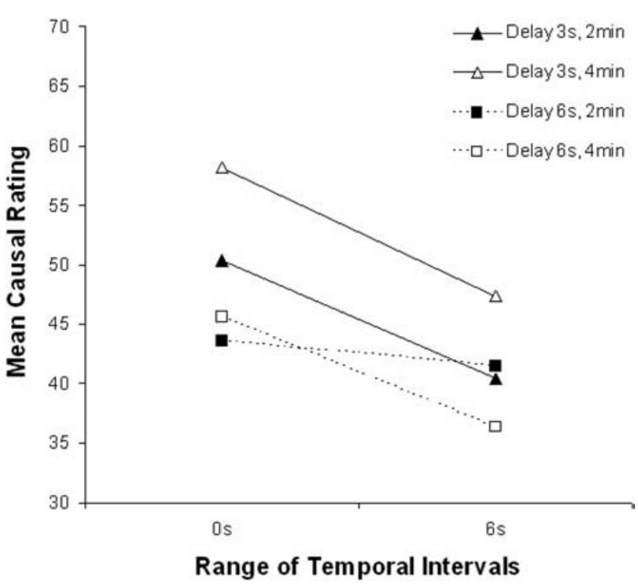

Figure 5. Mean causal ratings from Experiment 3 as a function of interval range. Filled and unfilled symbols refer to 2- and 4-min training, respectively. Mean delays are noted by different symbol and line styles.

affected by range (all $ps > 0.1$); mean delay and $P(e|c)$ were unaffected by condition duration (all $ps > 0.2$); and response rate and $P(e|c)$ were unaffected by delay (all $ps > 0.2$); therefore, the effects of our manipulations are not mediated through these potential confounds.

In summary, then, this experiment once again confirms that temporally predictable causal relations are judged as more causal than temporally variables ones. This effect of temporal predictability remains undiminished even as condition time increases, with condition time itself appearing to have little influence; the extent of information sampling does not moderate or mediate any effects associated with predictability. We can therefore be confident that the effect of predictability observed thus far (and demonstrated once again in this experiment) cannot be attributed to a mere failure of learning to reach asymptote. Temporal regularity remains as a cue to causality regardless of duration of learning.

General Discussion

In the present article, we attempted to broaden the perception of the role of time in causal learning. In particular, we have highlighted that temporal predictability can act as an empirical cue in the induction of causal relations from a real-time response-outcome schedule. More precisely, our results demonstrate that fixed, predictable temporal intervals attract higher causal ratings than variable ones and that causal ratings decrease as a function of temporal uncertainty. We now turn to consider the implications of these empirical findings with respect to the theories of causal learning discussed earlier.

An associative perspective on causal learning is partly motivated by the multitude of apparent similarities between conditioning in animals and causal learning in humans (Shanks & Dickinson, 1987). Endorsements of an associative perspective have considered phenomena such as outcome-density bias, sensitivity to cue competition, and super-learning as deep structural similarities between human causal learning and animal conditioning (Shanks, Holyoak, & Medin, 1996). We asked whether a similar commonality arises between the influence of temporal predictability on human causal learning and manipulations of temporal variability in studies of animal learning. The results from our experiments, however, have shown that humans tended to form the opposite view and instead drew the conclusion that causes producing their effects over a stable, reliable timeframe were more effective than those for which the effect occurred with variable latencies. What is the reason for this distinction?

One important conceptual difference between studies of animal conditioning and human causal learning, which might account for the divergent results with respect to temporal variability, is that the real appetitive or aversive stimuli (e.g., food rewards or shocks) are nearly always used in the former, whereas they are generally never used in the latter (e.g., triangles flashing)—or if they do, then only in hypothetical, imagined examples (e.g., food allergy scenarios, stock market "games"). Consequently, conditioning studies involve the concept of *utility*: A food reward is pleasant, and a foot shock is painful. Human causal learning studies, in contrast, do not call upon utility: Participants most likely do not care very much whether the triangle flashes or whether imaginary Mister X experiences an allergic reaction. Theories of delay discounting thus might not apply to human causal learning because

Table 4
Behavioral Data for Experiment 3

	Condition time							
	2 min				4 min			
	Delay							
	3 s		6 s		3 s		6 s	
	Range of temporal intervals							
Variable	0 s	6 s	0 s	6 s	0 s	6 s	0 s	6 s
Mean response rate (/min)	25.65	24.45	26.55	25.03	28.78	26.19	28.56	24.60
Mean outcome rate (/min)	19.18	18.29	20.00	19.26	21.68	19.53	21.57	18.61
Actual $P(e\|c)$	0.741	0.740	0.745	0.764	0.762	0.746	0.754	0.761
Mean actual delay (ms)	3000 (0)	3031 (343)	6000 (0)	5881 (343)	3000 (0)	3126 (244)	6000 (0)	5924 (240)
Mean causal rating	50.36 (28.44)	40.36 (28.03)	43.61 (32.92)	41.52 (28.17)	58.12 (26.70)	47.36 (23.54)	45.58 (25.66)	36.36 (25.91)

Note. Standard deviations appear in parentheses.

they address how (positive and negative) subjective utility decreases as a function of time-to-event. If the event has no intrinsic utility (as is arguably the case in human causal learning studies), then there is nothing to discount. In contrast, rewards and punishments are very clearly liable to discounting, both in human and nonhuman animals. The advantage of variable over fixed intervals in studies of animal learning thus may well be grounded in the shape of the discounting function and commensurate differences in subjective utility of the obtained outcomes. But because studies of human causal learning do not involve utility, discounting does not apply. The influence of temporal predictability on human causal learning thus must have another theoretical basis.

A traditional associative perspective therefore encounters difficulty when trying to deal with results such as these, and more well-documented findings of knowledge mediation such as those of Buehner and May (2002, 2003, 2004). However, more recent formulations of associative accounts have challenged the limited conception of time that these earlier models were hamstrung by. The *temporal coding hypothesis* (e.g., Miller & Barnet, 1993) argues that animals do indeed learn the temporal intervals in a conditioning preparation, and this information is important in determining not only whether an association is acquired, and the strength of this association, but also whether it is subsequently expressed in behavior. This idea has steadily accumulated support, because it has proved capable of addressing findings concerning variation in the CS–US timing that previous associative models (e.g., Pearce & Hall, 1980; Rescorla & Wagner, 1972) could not account for (e.g., Blaisdell, Denniston, & Miller, 1998; Savastano, Yin, Barnet, & Miller, 1998). By acknowledging that animals encode temporal information as part of the association, this view could potentially address findings for which the role of time appears to go beyond mere contiguity. For instance, Allan, Tangen, Wood, and Shah (2003) argued that the temporal coding hypothesis can be adapted to accommodate the results of Buehner and May (2004), and their own findings, that delayed causal relations receive higher causal evaluations than contiguous relations under certain circumstances (the basis of this argument is that knowledge mediation serves as an initial training phase in which the observer “learns” the delay). A similar extrapolation of this theory might apply here; if an organism learns the temporal interval between events and carries this forward, subsequent variation of the intervals might negatively impact CS–US association (as does a disruption of continuity between training phases, e.g., in latent inhibition or Hall & Pearce, 1979, negative transfer). Indeed, Denniston, Blaisdell, and Miller (1998) have already demonstrated an adverse effect of temporal incongruence in inhibitory conditioning.

The temporal coding hypothesis can not only account for the superiority of temporal regularity, but it paradoxically also appears capable of addressing the preference for variability observed in studies in which reinforcement schedules are used. The notion that contiguity is a key determinant of associative strength remains a fundamental tenet of the temporal coding hypothesis, as outlined by Blaisdell et al. (1998): “Contiguity is sufficient for the formation of an association. The degree of spatial and temporal proximity between two events (stimuli or responses) determines the extent to which they are associated” (p. 72). Thus, the association will depend on how associative strength changes as a function of delay, and the shape of this function may be highly dependent on the context. As mentioned previously, because utility is crucial for animal reinforcement learning, it may well be that the associative strength of delayed events does, in such cases, decline in a manner consistent with delay discounting (see Figure 1).

The difficulty, then, seems to lie in determining the specific predictions of the temporal coding hypothesis; what are the circumstances that govern whether a facilitatory or an inhibitory effect of variability on learning is anticipated from this perspective? As far as we are aware, the temporal coding hypothesis does not per se put differential weights on the extent versus the constancy of the reinforcement delay. Consequently, it could potentially be adapted to fit any set of results via a post hoc reconceptualization of the learning task (see, e.g., Allan et al., 2003). What is therefore needed is some extension or restriction of this theory that would enable it to specify a priori the expected progression of learning given a particular input or data set. Clearly, the temporal coding hypothesis represents an important step in the development of associative learning theory; the fundamental principle that temporal information is encoded in an association enables the multifaceted influences of time in learning to be accommodated. It remains unclear, however, whether the anticipation of a definitive influence of temporal predictability in a given situation can be derived from this theory. We therefore now turn to consider other theoretical approaches that make more concrete predictions regarding predictability.

Attribution Aide or Cognitive Component?

Having struggled to reconcile our findings with associative learning theory, we move now to consider our results in light of the covariation perspective. We described in the introduction how the attribution shift hypothesis could extend a covariational perspective to account for the effect of predictability by reducing erroneous attribution of delayed effects to random background processes. With a temporally predictable cause, repeated experience of a constant interval may lead the reasoner to adjust their temporal window such that delayed events are attributed to the candidate cause rather than disregarded. However, there remains the compelling question of whether time merely serves to facilitate or inhibit the detection and interpretation of events or whether temporal information itself is actually computed to form an integral part of the mental representation of causality. According to this account, temporal information is not considered to form part of a mental representation of causality, but merely determines the attribution of events to the cells of a contingency table. However, if this were the case, and predictability improves causal judgments simply by enabling the reasoner to correctly detect cause–effect pairings, then the *degree* of separation between cause and effect should not matter. If repeated experience of the same interval enables detection of delayed events, then there should not be a simultaneous effect of delay. Under these assumptions, then, although an effect of predictability could be accounted for, effects of predictability and delay are mutually exclusive and could not occur in tandem as demonstrated by our results. Besides, Greville and Buehner (2007) have already demonstrated that contiguity and covariation act in concert to influence causal judgment, even in situations in which the extent of contingency is unambiguous.

Additionally, the covariation account and attribution shift hypothesis encounter difficulty with the results from Experiment 3. If participants are given more time to explore the causal relation in question, they most likely will (and in our case indeed did) experience more action-outcome pairings. The more exposure participants have to a particular contingency, the more likely it is that they will be able to recognize it correctly. Although it is clear that temporal cues such as contiguity or predictability may assist in the recognition of cause–effect pairings in the short term at least (and conversely, temporal delay or unpredictability may impede the attribution of effect to the cause), given enough exposure, participants should be able to detect contingencies independently of temporal information. If participants do in fact come to notice the contingency, and this is the determinant of their causal representation, then temporal information should cease to be important. However, as Experiment 3 revealed, judgments of causality did not move significantly closer to ΔP as learning time was increased, and the effects of predictability and delay persisted. The implication is that cues such as contiguity and predictability are in and of themselves components of a computation of causal strength, rather than just an aide to event parsing for the calculation of covariance, as a purely statistical or contingency-based approach to learning would suggest.

Thus, the evidence from this study is incompatible with a covariation perspective even when its assumptions are relaxed as per the attribution shift hypothesis. We would suggest that it is still possible that the process of attribution shift does in fact take place during event parsing but that the constraints of the covariation account on this process are invalid. According to a strict covariation account, having determined whether or not event pairings are causal or spurious, temporal information then plays no further role in the learning process. However, if, instead, temporal information is still represented in the mental computation, then what we have is essentially a trade-off between contingency and contiguity. For instance, suppose that predictability does indeed result in a shift of the temporal window. In a delayed but predictable relation, it is likely that attribution shift will not occur; because all the effects happen after the same interval, they should be attributed to the cause. However, because they are all delayed, the overall impression of contiguity will be weak. For a delayed but variable relation, however, although later events may be disregarded as spurious, there will also be earlier events that occur with closer contiguity than events in the fixed interval relation, which should be attributed to the cause. Subjective contingency therefore is decreased relative to the fixed condition; however, because the remaining $c \rightarrow e$ pairings that are counted will all have equal or shorter intervals than the fixed delay, the overall impression of contiguity is stronger for the variable condition. Thus, whether variable or predictable causal relations are perceived as stronger would crucially depend on the trade-off between contingency and contiguity (see Buehner & McGregor, 2009).

A Bayesian Perspective

As discussed previously, Bayesian models of causal learning assess the likelihood of the obtained data under two opposing hypotheses: one for which there is a genuine mechanistic link between candidate cause and effect and one for which no such links exists, and the effect is the result of alternative unseen causes. Regularity is more likely under the former hypothesis than the latter, so it is taken as evidence for the existence of a causal relation. Though Griffiths and Tenenbaum's (2005) causal support model was originally developed as a computational account of assessing causal structure from contingency information, a logical extension of this perspective could easily be applied to temporal information. Applying the same principles, the prediction of the structure account with regard to the phenomenon addressed in this article is clear: Temporal regularity should facilitate learning. Indeed, in a more recent framework, Griffiths and Tenenbaum (2009) extended the structure account and highlighted the importance of patterns of spatial or temporal coincidences, with a set of regularly spaced events being much more probable under an identified potential mechanism than a spontaneous activation of an unseen alternative cause.

From such a perspective, predictability may further facilitate causal learning through the process of Bayesian updating (see, e.g, Lagnado & Sloman, 2002; Lagnado, Waldmann, Hagmayer, & Sloman, 2007). For instance, a reasoner may, in the first few instances of experiencing a delayed causal relation, decide that the effect was not actually generated by the cause. However, if the temporal interval is fixed, then after several exposures, the reasoner may revise and update their causal beliefs about the relation in question and adopt a new expectation of the timeframe. If they then continue to experience effects that occur at the time they now expect, then this will reinforce the impression of a causal relation. Additionally, events that had previously been classed as noncausal may also be reevaluated as causal, further contributing to the overall impression of causal strength. However, one problem with a simple formulation of the Bayesian account is that it, too, like the attribution shift hypothesis, would seem incapable of simultaneously accounting for a joint influence of delay and temporal predictability. Presumably, if a temporal interval is highly predictable, and therefore provides good support for a causal structure model, the extent of delay should not matter. One way to address this would be for future models to include priors of delay assumptions that reflect the consistent bias to prefer contiguous over delayed relations.

Classical Conflict, Concerns, and Caveats

Our results appear to contradict recent work from Young and Nguyen (2009). Their experiment could in essence be conceived as a classical conditioning analogue of the study reported here, with participants observing events rather than taking instrumental action. Their participants navigated a 3-D virtual world in which they could encounter groups of three "orcs" firing crossbows onto a distal target. The task required them to decide which of the orcs was responsible for causing explosions and to select that orc for destruction. The delay between orcs' firing and the explosions and the variability of this delay were varied, along with presence of auditory fillers during the delay. Contrary to our findings, constancy of delay did not appear to provide an advantage, and indeed variability often lead to a higher percentage of correct target selection. However, there are a multitude of differences between the studies in the present article and the paradigm devised by Young and Nguyen. Although a full discussion of these is not appropriate here, we briefly highlight a few of the key distinctions. First, different dependent measures were used. Rather than providing a judgment of causal strength, participants instead were given a forced-choice discrimination task, having to select the correct target from multiple causal candidates, meaning that competing agents could often come between the true cause and the outcome. This would be particularly true for a constant, high-delay causal candidate, which Young and Nguyen anticipated. In running Monte Carlo simulations prior to the experiment, the authors discovered that "highly variable long delays produced a larger number of experiences of the true cause being more contiguous to the effect whereas consistent long delays produced more experiences of one of the foils being more contiguous" (p. 300). This certainly appears to have been reflected in the results, with correct identification of the target for fixed, long-delay causal candidates falling as low as under 20%. This problem may well have been exacerbated by the fact that sampling times were self-truncated; no restriction was made regarding minimum observation time, and indeed Young and Nguyen reported that players "were not motivated to obtain large observation samples" (p. 309). Therefore, if an observer experienced a contiguous candidate early on and is particularly "trigger-happy," they may select this as the target. The likelihood that early selection of a target based on a small amount of contiguous pairings results in correct target identification naturally is much higher in variable than in fixed-delay conditions. Curiously, however, the effect was not very consistent across experiments and seemed to be restricted to male participants. This, combined with research suggesting that impulsivity is more likely in males than in females (e.g., D'Zurilla, Maydeu-Olivares, & Kant, 1998), would produce the distinction seen in Young and Nguyen's results. Different results may well have been obtained had learners been given sampling opportunities of a predetermined duration (as they were in our studies). To sum up, for the purposes of the present article, our results and those presented by Young and Nguyen are likely a reflection of the considerable differences in methodology rather than a contradiction with regard to the influence of temporal predictability. Accordingly, we turn now to consider other issues that raise more immediate concerns for the present study.

One important methodological aspect of the experiments presented in this article is the assumption that the psychological mean of the temporal intervals is equivalent to the arithmetic mean. To adequately compare variable and fixed delays, it was necessary to ensure that the mean of the intervals in the variable condition was (approximately) equal to that of the predictable condition, because a discrepancy would imply that the differences in predictability were confounded with different actual experienced delays. Indeed, in all such types of experiment, there is bound to be some fluctuation of the mean-experienced delay from the nominal programmed delay set by the experimenters (though an analysis of these data for our experiments showed a good degree of isomorphism between the two). However, it is not necessarily a given that the mean of these experienced intervals is functionally equivalent to the psychological mean. If subjective perceived duration of a temporal interval differs from the veridical duration, then the perceived mean duration will likewise differ from the recorded mean. This need only be cause for concern for our studies if subjective duration is some nonlinear function of actual duration. Wearden (1991) has shown that subjective time increases linearly as a function of real time in interval reproduction (e.g., Humphreys & Buehner, 2009), but even if time perception instead followed a logarithmic function in line with the Weber-Fechner law, this would still not cause problems for the interpretation of our results. In this case, longer intervals would be increasingly underestimated relative to shorter intervals, and the (subjective) net delay would thus be smaller when considering a short and long delay compared with two instances of a constant delay formed by the arithmetic mean of the short and long interval. Therefore, this discrepancy would only work against our hypothesis and make it less likely for predictable relations to draw higher ratings than variable ones. Because we still find an advantage for predictable conditions, this is not really a concern. In fact, in light of this consideration, the obtained findings are all the more noteworthy.

In summary, the effects of temporal predictability we have demonstrated, combined with the pervasive (and already established) effects of delay, suggest that an alternative conception of the contribution of time in causal induction may provide a better model for the learning process. We propose that, in line with the structural account, temporal information should be regarded in a similar manner to statistical information, which is to say that regularities in this input are used by reasoners to infer causal relations. Therefore, just as statistical regularity facilitates causal discovery, so does temporal regularity. The rationale behind this argument is that reasoners evaluate the likelihood of obtaining the observed data that is available to them within two hypothetical universes. In one universe, there is a genuine mechanistic link between candidate cause and effect, and in the other, there is not (and the effect happens solely because of random background conditions). Under the latter hypothesis, any form of cause–effect regularity is unlikely. If there is consistently a reliable timeframe of event occurrence such that cause and effect are routinely separated by the same temporal interval, then this provides growing evidence of a causal relation.

Thus, when an agent evaluates the causal effectiveness of their intervention on the environment, the effects of time may be seen as fourfold. First, as has been pointed out many times previously in the literature, causal relations with short delays are much easier to learn than those with long delays. If there is a temporal separation between cause and effect, then establishing a causal link between them requires far greater cognitive effort; the events must be held in memory for longer periods, and other events that occur in the intervening period must be ignored. Second, there is also the cognitive or pragmatic component of delay. In the case of a generative cause, if two different events produce an outcome but one does so more rapidly than the other, then that event may be judged as the stronger cause, particularly if considerations of utility figure in the evaluation of the relation. For instance, if a person has a splitting headache, then the sooner a medication can provide relief, the better. Third, any temporal interval between cause and effect may be compared with an existing hypothesis about the causal mechanism and the expected time frame of event occurrence. Evidence that conforms to this will strengthen the causal relation, whereas that which deviates from expectation will weaken the impression. Fourth, evidence of a regular (delayed) temporal interval between cause and effect might either facilitate the discovery of the statistical regularity between cause and effect or may result in the reasoner modifying prior assumptions about a different interval, or both. Because temporal regularity is highly unlikely to occur by random chance, the likelihood of H_1 (that there is a causal relation) compared with H_0 (that there is not) increases.

We accept that it is not a given that the results obtained in the present article will necessarily generalize to other types of learning situations, and further research may consider alternative preparations. Young and Nguyen (2009) have already suggested observational paradigms and multiple-cue contingency learning as potential avenues, and these warrant further exploration. An additional possibility is a scenario in which the operational relationship between cause and effect is already clearly defined, with no ambiguity regarding which response generates which outcome. Such a scenario would provide further clarity as to whether temporal variability weakens impressions of causality by degrading the subjective perception of contingency or purely due to the uncertainty regarding effect timing. Ultimately, however, the implication that we hope to impart from this article, beyond our empirical findings, is that causal induction is a flexible and dynamic process that occurs within time, and time therefore must be an integral component of the learning process. Models of causal learning therefore crucially need to represent temporal information as well as frequencies or rates of causes and effects. Among present popular perspectives on learning, two divergent approaches would seem to offer key insights to this issue. The temporal coding hypothesis offers the flexibility to incorporate differential effects of time dependent on the learning situation by positing that organisms learn temporal relationships along with associations and suggests that the nature of behavior depends on the integration of temporal relationships between multiple cues (Blaisdell et al., 1998). Meanwhile, the causal structure approach presents the twofold argument that causality is based on a mechanistic connection between cause and effect, and such mechanisms reveal themselves through environmental regularities. Combining the relative merits of these two perspectives may be a useful approach to the direction of future research.

References

Allan, L. G., Tangen, J. M., Wood, R., & Shah, T. (2003). Temporal contiguity and contingency judgments: A Pavlovian analogue. *Integrative Psychological and Behavior Science, 38,* 214–229.

Andrzejewski, M. E., Cardinal, C. D., Field, D. P., Flannery, B. A., Johnson, M., Bailey, K., & Hineline, P. N. (2005). Pigeons' choices between fixed-interval and random-interval schedules: Utility of variability? *Journal of the Experimental Analysis of Behavior, 83,* 129–145. doi:10.1901/jeab.2005.30-04

Blaisdell, A. P., Denniston, J. C., & Miller, R. R. (1998). Temporal encoding as a determinant of overshadowing. *Journal of Experimental Psychology: Animal Behavior Processes, 24,* 72–83. doi:10.1037/0097-7403.24.1.72

Blaisdell, A. P., Sawa, K., Leising, K. J., & Waldmann, M. R. (2006, February 17). Causal reasoning in rats. *Science, 311,* 1020–1022. doi:10.1126/science.1121872

Buehner, M. J. (2005). Contiguity and covariation in human causal inference. *Learning and Behavior, 33,* 230–238.

Buehner, M. J., Cheng, P. W., & Clifford, D. (2003). From covariation to causation: A test of the assumption of causal power. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 1119–1140. doi:10.1037/0278-7393.29.6.1119

Buehner, M. J., & May, J. (2002). Knowledge mediates the timeframe of covariation assessment in human causal induction. *Thinking and Reasoning, 8,* 269–295. doi:10.1080/13546780244000060

Buehner, M. J., & May, J. (2003). Rethinking temporal contiguity and the judgement of causality: Effects of prior knowledge, experience, and reinforcement procedure. *The Quarterly Journal of Experimental Psychology, 56*A(5), 865–890.

Buehner, M. J., & May, J. (2004). Abolishing the effect of reinforcement delay on human causal learning. *The Quarterly Journal of Experimental Psychology, 57*B(2), 179–191.

Buehner, M. J., & McGregor, S. (2006). Temporal delays can facilitate causal attribution: Towards a general timeframe bias in causal induction. *Thinking and Reasoning, 12,* 353–378. doi:10.1080/13546780500368965

Buehner, M. J., & McGregor, S. J. (2009). Contingency and contiguity trade-offs in causal induction. *International Journal of Comparative Psychology, 22,* 19–42.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104,* 367–405. doi:10.1037/0033-295X.104.2.367

Chung, S.-H. (1965). Effects of delayed reinforcement in a concurrent situation. *Journal of the Experimental Analysis of Behavior, 8,* 439–444.

Cicerone, R. A. (1976). Preference for mixed versus constant delay of reinforcement. *Journal of the Experimental Analysis of Behavior, 25,* 257–261. doi:10.1901/jeab.1976.25-257

Damianopoulos, E. N. (1982). Necessary and sufficient factors in classical conditioning. *Integrative Physiological and Behavior Science, 17,* 215–229.

Denniston, J. C., Blaisdell, A. P., & Miller, R. R. (1998). Temporal coding affects transfer of serial and simultaneous inhibitors. *Animal Learning & Behavior, 26,* 336–350.

Dickinson, A. (2001). Causal learning: An associative analysis. *Quarterly Journal of Experimental Psychology, 54*B, 3–25.

Dickinson, A., Shanks, D., & Evenden, J. (1984). Judgement of act-outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 36*(1-A), 29–50.

D'Zurilla, T. J., Maydeu-Olivares, A., & Kant, G. L. (1998). Age and gender differences in social problem-solving ability. *Personality and Individual Differences, 25,* 241–252. doi:10.1016/S0191-8869(98)00029-4

Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin, 99,* 3–19. doi:10.1037/0033-2909.99.1.3

Evans, A. S. (1993). *Causation and disease: A chronological journey*. New York, NY: Plenum Medical Book Company.

Gluck, M. A., Mercado, E., & Myers, C. E. (2008). *Learning and memory: From brain to behaviour*. New York, NY: Worth Publishers.

Goode, M. K., Geraci, L., & Roediger, H. L., III. (2008). Superiority of variable to repeated practice in transfer on anagram solution. *Psychonomic Bulletin & Review, 15,* 662–666. doi:10.3758/PBR.15.3.662

Green, L., & Myerson, J. (1996). Exponential versus hyperbolic discounting of delayed outcomes: Risk and waiting time. *American Zoologist, 36,* 496–505.

Greville, W. J., & Buehner, M. J. (2007). The influence of temporal distributions on causal induction from tabular data. *Memory & Cognition, 35,* 444–453.

Greville, W. J., Cassar, A., Johansen, M. K., & Buehner, M. J. (2010, August). *Structure awareness in action-outcome learning eradicates the detrimental effect of reinforcement delays.* Poster presented at the 2010 meeting of the Cognitive Science Society.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology, 51,* 334–384. doi:10.1016/j.cogpsych.2005.05.004

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review, 116,* 661–716. doi:10.1037/a0017201

Hall, G., & Pearce, J. M. (1979). Latent inhibition of a CS during CS-US pairings. *Journal of Experimental Psychology: Animal Behavior Processes, 5,* 31–42. doi:10.1037/0097-7403.5.1.31

Hammond, L. J., & Paynter, W. E. (1983). Probabilistic contingency theories of animal conditioning: A critical analysis. *Learning and Motivation, 14,* 527–550.

Hume, D. (1888). *Hume's treatise of human nature* (L. A. Selby-Bigge, Ed.). Oxford, UK: Clarendon Press. (Original work published 1739)

Humphreys, G. R., & Buehner, M. J. (2009). Magnitude estimation reveals temporal binding at super-second intervals. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1542–1549. doi:10.1037/a0014492

Jenkins, H., & Ward, W. (1965). Judgment of contingencies between responses and outcomes. *Psychological Monographs, 7,* 1–17.

Lagnado, D. A., & Sloman, S. A. (2002). Learning causal structure. *Proceedings of the twenty-fourth Annual Conference of the Cognitive Science Society* (pp. 560–565). Mahwah, NJ: Erlbaum.

Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2007). Beyond covariation: Cues to causal structure. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 154–172). New York, NY: Oxford University Press.

Mackintosh, N. J. (1983). *Conditioning and associative learning.* Oxford, UK: Clarendon Press.

Miller, R. R., & Barnet, R. C. (1993). The role of time in elementary associations. *Current Directions in Psychological Science, 2,* 106–111. doi:10.1111/1467-8721.ep10772577

Myerson, J., & Green, L. (1995). Discounting of delayed rewards: Models of individual choice. *Journal of the Experimental Analysis of Behavior, 64,* 263–276. doi:10.1901/jeab.1995.64-263

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87,* 532–552. doi:10.1037/0033-295X.87.6.532

Perales, J. C., & Shanks, D. R. (2007). Models of covariation-based causal judgment: A review and synthesis. *Psychonomic Bulletiin & Review*, *14,* 577–596.

Reed, P. (1993). Influence of the schedule of outcome presentation on causality judgements. *The Quarterly Journal of Experimental Psychology, 46*A(2), 327–345.

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist, 43,* 151–160. doi:10.1037/0003-066X.43.3.151

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 theory and research* (pp. 64–99). New York, NY: Appleton-Century-Crofts.

Savastano, H. I., Yin, H., Barnet, R. C., & Miller, R. R. (1998). Temporal coding in Pavlovian conditioning: Hall-Pearce negative transfer. *The Quarterly Journal of Experimental Psychology, 51*B(2), 139–153.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82,* 225–260. doi:10.1037/h0076770

Shanks, D. R. (1987). Acquisition functions in contingency judgment. *Learning and Motivation, 18,* 147–166.

Shanks, D. R., & Dickinson, A. (1987). Associative accounts of causality judgment. In G. H. Bower (Ed.), *Psychology of learning and motivation: Advances in research and theory* (Vol. 21, pp. 229–261). San Diego, CA: Academic Press.

Shanks, D. R., & Dickinson, A. (1991). Instrumental judgment and performance under variations in action outcome contingency and contiguity. *Memory & Cognition, 19,* 353–360.

Shanks, D. R., Holyoak, K. J., & Medin, D. L. (Eds.). (1996). *The psychology of learning and motivation*: *Causal learning* (Vol. 34). San Diego, CA: Academic Press.

Shanks, D. R., Pearson, S. M., & Dickinson, A. (1989). Temporal contiguity and the judgment of causality by human subjects. *The Quarterly Journal of Experimental Psychology, 41*B(2), 139–159.

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. doi:10.1037/0096-3445.121.2.222

Wasserman, E. A., Chatlosh, D. L., & Neunaber, D. J. (1983). Perception of causal relations in humans. *Learning and Motivation, 14,* 406–432. doi:10.1016/0023-9690(83)90025-5

Wearden, J. H. (1991). Do humans possess an internal clock with scalar timing properties? *Learning and Motivation, 22,* 59–83. doi:10.1016/0023-9690(91)90017-3

Wulf, G., & Schmidt, R. A. (1997). Variability of practice and implicit motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 987–1006. doi:10.1037/0278-7393.23.4.987

Young, M. E., & Nguyen, N. (2009). The problem of delayed causation in a video game: Constant, varied, and filled delays. *Learning and Motivation, 40,* 298–312. doi:10.1016/j.lmot.2009.02.002

Received March 10, 2010
Revision received July 9, 2010
Accepted July 12, 2010 ■

Correction to Greville and Buehner (2010)

In the article "Temporal Predictability Facilitates Causal Learning," W. James Greville and Marc J. Buehner (*Journal of Experimental Psychology: General*, 2010, Vol. 139, No. 4, pp. 756–771), Figure 2 (p. 759) contained an error. The terms e|¬c and ¬e|c were mislabelled as ¬e|c and e|¬c. The corrected figure appears below.

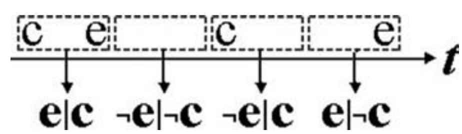

Figure 2. The effect of attribution shift in parsing an event stream with a fixed temporal window: c→ e intervals that are longer than the temporal window simultaneously decrease impressions of $P(e|c)$ and increase impressions of $P(e|\neg c)$.

DOI: 10.1037/a0024735