Information Search in an Autocorrelated Causal Learning Environment

Benjamin Margolin Rottman (rottman@pitt.edu)
Learning Research and Development Center, 3939 O'Hara St
Pittsburgh, PA 15260 USA

Abstract

When trying to determine which of two causes produces a more desirable outcome, if the outcome is autocorrelated (goes through higher and lower periods) it is critical to switch back and forth between the causes. If one first tries Cause 1, and then tries Cause 2, it is likely that an autocorrelated outcome would appear to change with the second cause even though it is merely undergoing normal change over time. Experiment 1 found that people tend to perseverate rather than alternate when testing the effectiveness of causes, and perseveration is associated with substantial errors in judgment. Experiment 2 found that forcing people to alternate improves judgment. This research suggests that a debiasing approach to teach people when to alternate may be warranted to improve causal learning.

Keywords: Information Search, Causal Inference, Autocorrelated Environment, Dynamic Environment

Introduction

As researchers we are all familiar with history as a threat to internal validity. For example, suppose that we are comparing two interventions. Designing an experiment in which all participants first experience Intervention 1 and then Intervention 2 is flawed because a historical event, maturational change, or order effect could confound the results and make it seem as if there is a real difference between 1 and 2 even if there is not.

Despite the dubiousness of such a learning strategy it seems common in every-day learning situations. For example, a person is prescribed a new blood pressure medicine, tries it for a week, notices an improvement, and concludes that the new medicine works better than the old medicine. This inference is flawed because any number of other changes over time such as changes in diet or stress could be responsible for the change in blood pressure. Or, consider a parent who starts to bribe his child to behave better and notices an improvement. The change could be due to the bribe or any number of other factors such as starting to play a sport or growing more mature.

One way to increase the validity of such a "single-subject" design is to alternate between the two conditions (e.g., 1, 2, 1, 2) (Barlow & Hayes, 1979). With more alternations it is less likely that the baseline trend would correlate with the two conditions, reducing the likelihood of being fooled into believing that there is a difference merely due to the baseline trend. The current manuscript examines what sort of "experiments" people tend to design [e.g., (1, 1, 1, 2, 2, 2) vs. (1, 2, 1, 2, 1, 2)], and whether the experimental design influences their conclusions.

Information Search in Decisions from Experience

In general, when a learner has the opportunity to choose a piece of information to sample it is called "active learning" or "information search." One common information search experimental task involves a learner repeatedly choosing between two or more options, x=1 or x=2, and after each choice the learner receives the outcome Y. By sampling the two choices the learner forms an expectation of the outcome of Y given the different choices of X, and can use that expectation to choose a value of X that produces a desired outcome of Y. Experiments of this sort can reveal the patterns that people use when selecting X, how the information search pattern influences what is learned, and how well the learner obtains the desired outcome.

Information search paradigms vary on many different dimensions; here I focus on the difference between "stable" and "dynamic" environments. In a stable environment the outcome of Y given a particular choice (e.g., x=1) is stable over time. For example, Y given x=1 could be determined by a normal distribution with mean=10 and SD=2 whereas Y given x=2 could be determined by a normal distribution with mean=12 and SD=2. Hills and Hertwig (2010) found that in a stable environment the sampling pattern that individual participants used influenced their beliefs about which choice produced higher payoffs. It appears that people who frequently switched back and forth between the two options were essentially comparing which choice produced a higher outcome on sequential choices. At the end they tended to choose the option that more frequently produced a higher outcome even though on average it produced a lower mean outcome. In contrast, people who switched less frequently tended to choose the option that on average produced the higher outcome. In sum, perseverating was associated with maximizing expected value.

Other experiments have investigated information search in <u>dynamic</u> environments. A dynamic environment is one in which the probability of reward does not remain stable over time. The main type of dynamic environment that has been studied is one in which sometimes x=1 produces a higher reward than x=2, and sometimes it produces a lower reward than x=2 (e.g., Biele, Erev, & Ert, 2009; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Yi, Steyvers, & Lee, 2009). The outcome is autocorrelated in the sense that if x=1 is the better choice at Time 5, it will likely be the better choice at Time 6, but participants do not know how long it will remain the better choice. Dynamic environments have been used primarily in conjunction with tasks that involve both exploration and exploitation; participants are instructed to maximize their reward over an entire set of choices. Because the rewards for the two options change, one should not perseverate too long on one option, because in the meanwhile it is possible that the other option has switched to giving a higher reward. Thus, the task requires a combination of exploring (alternating) and exploiting (perseverating).

Current Task

In the current study we investigated another type of dynamic environment. In this environment if the outcome is high at Time 5 it will likely also be high at Time 6 – the outcome is autocorrelated. However, unlike in the other task, one of the choices always produced a lower outcome than the other and the goal is to figure out which of the choices produces a higher vs. lower outcome. This type of situation is more in line with research on causal inference – the goal is to figure out the causal relationship between the choices and outcome, and the causal relationship is stable over time even though the outcome changes (c.f. Hagmayer, Meder, Osman, Mangold, & Lagnado, 2010).

In the current studies, participants were allowed to choose between two levels of a cause (two medicines) for 14 trials, and on each trial they saw the outcome (amount of pain) on the scale 0-100. At the end of the 14 trials participants had to judge which cause (Medicine 1 or 2) results in a lower effect (pain), and by how much. One of the medicines always produced a slightly lower outcome than the other.

The outcome was determined by a baseline trend that was autocorrelated over time, which is what makes the environment dynamic. The line in Figure 1 represents one example baseline trend of pain across 14 days. In Figure 1, Medicine 1 always reduced the pain 5 points relative to the baseline, and Medicine 2 did not reduce pain at all relative to the baseline. Participants did not know the baseline trend – they only observed the outcome of pain after choosing a medicine.

Figure 1 exemplifies how alternation is a much more useful strategy in the context of an autocorrelated effect. If one perseverates, then it is possible that the two levels of the cause will happen to coincide with high or low periods of the effect. For example, in Figure 1A, Medicine 1 happens to be tried when the baseline level of pain is fairly high, and Medicine 2 is tried when the baseline level of pain is fairly low. If one aggregates across these two periods one would likely conclude that Medicine 2 works much better (lower pain scores) then Medicine 1, perhaps by 20 points or so. Thus, perseveration can <u>sometimes</u> result in a strong inference in the wrong direction (inferring that Medicine 2 is better than Medicine 1 even though the opposite is true).

A quick thought experiment reveals that perseveration can also sometimes result in an inference in the right direction, but far too strong. For example, imagine Figure 1a but with Medicine 2 tried for the first 7 days and Medicine 1 tried for Days 8-14. It is also possible, depending on the underling baseline, that that the periods of the two levels of the cause will not line up with different levels of the baseline (e.g., if Medicine 1 is tried Days 1-7 and Medicine 2 is tried Days 8-14, and if the baseline is symmetric and peaks on Days 7 and 8). In this case the baseline would not be confounded with the medicine choices.

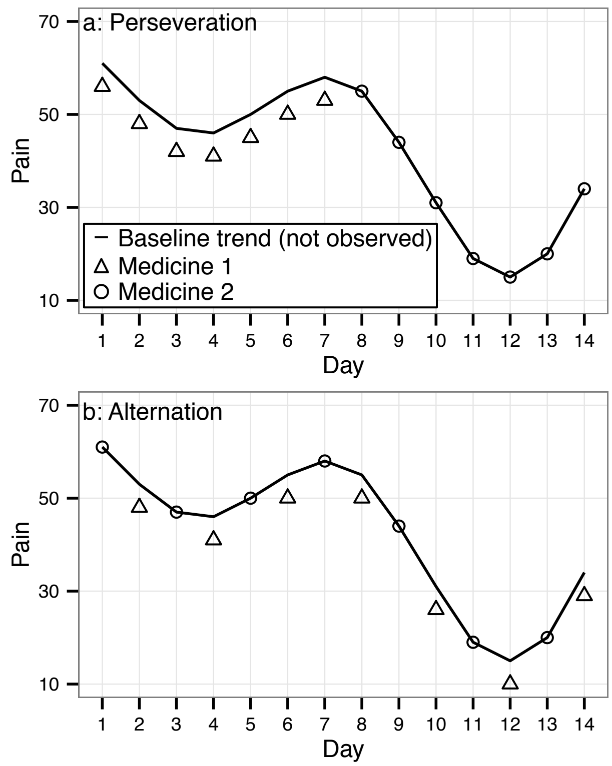

Figure 1: The Effect of Perseveration and Alternation in an Autocorrelated Environment

Figure 1B demonstrates how alternation can produce more accurate inferences when there is an autocorrelated effect. Since the effect is autocorrelated, sequential observations will generally have fairly similar baseline levels. By comparing sequential days alternation reduces the influence of the baseline. Comparing sequential days is not always perfect. For example, Medicine 2 on Day 9 produces a lower pain level than Medicine 1 on Day 8 even though in general Medicine 1 reduces pain relative to Medicine 2. The reason for this effect is because there is a decreasing baseline trend during this period. However, even though Day 9 is lower than Day 8, Medicine 1 on Day 10 produces considerably lower pain than Medicine 2 on Day 9. In sum, in autocorrelated environments alternation produces "cleaner" data that would likely lead to more accurate inferences.

This paper addresses the following three questions:

1. Do people alternate?
2. Do people who alternate make better inferences?
3. Is alternation the cause of the better inferences?

Experiment 1

Methods

Participants 152 participants were recruited through MTurk. I intended to recruited 100, but due to a server error 52 were terminated early. Thus I ran another 52 to have a total of 100 participants who completed the entire study. All the data were analyzed from all participants. Participants were paid $1 and there was the possibility of a bonus for accuracy explained below.

Stimuli On each trial participants chose either Medicine 1 or Medicine 2. One of the medicines always reduced the amount of pain by 5 points relative to the baseline; the other medicine did not change the amount of pain from the baseline. Participants never directly observed the baseline trend – they only saw the pain outcome after having chosen one of the medicines. The baseline trend was a sum of three sine waves with different amplitudes and frequencies producing an unpredictable but highly autocorrelated sequence. Each baseline sequence of 14 days was chosen randomly along the length of function. Figure 2 shows 10 sample baseline trends. They include increasing, decreasing, peak in the middle, peaks at the ends, and other patterns.

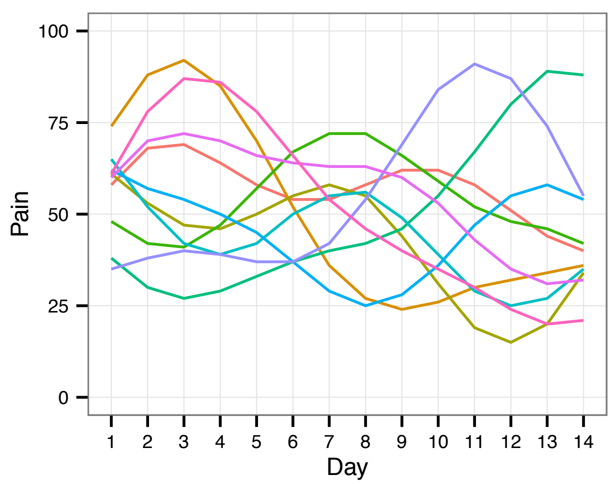

Figure 2. 10 Example Baseline Trends

Procedures Participants read the following instructions:

> "Please imagine that you are a doctor treating patients for chronic back pain. There are two medicines that you can use. These medicines are meant to be taken once a day in the morning and they work all day. For 50% of patients Medicine 1 works better, and for 50% of patients Medicine 2 works better. Thus, you are going to try to figure out which medicine works the best for each individual patient. Every morning you decide whether the patient should take Medicine 1 or Medicine 2. Then you will see how much pain the patient is in during the afternoon. You have 14 days to test the medicines. At the end of the 14 days, you will judge which medicine works better and by how much. You will receive a bonus according to how close your estimate comes to the true difference in the effectiveness of the two medicines for the patient. The table below (Table 1) shows how much bonus you can earn for each of the eight patients."

Table 1: Bonus Scale

Judgment within +/- points	2	4	6	8	>8
Reward (in cents)	20	15	10	5	0

On each day (trial), participants chose whether the patient would take Medicine 1 or Medicine 2, and after choosing they saw the level of the outcome, pain, presented as a number 0-100 (Figure 3A); the outcome was <u>not</u> presented as a graph. In reality, for a given patient, Medicine 1 or 2 was randomly chosen to work better; the medicine that worked better always reduced the pain by exactly 5 points from baseline. This amount was chosen to make the discrimination challenging but not impossible.

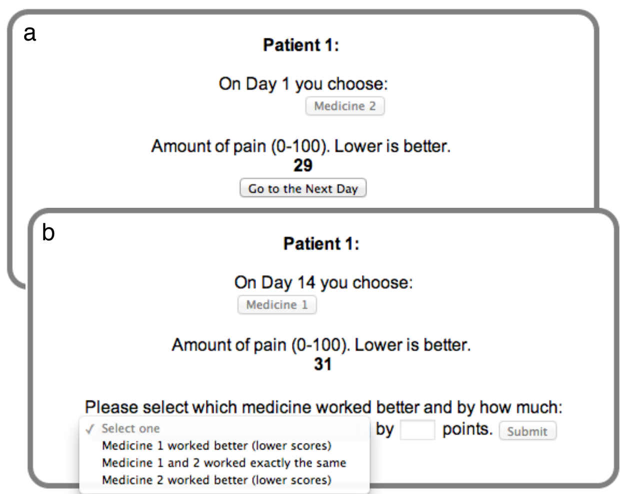

Figure 3. Screenshots of Learning (A) and Judgment (B)

After making the 14 choices and seeing 14 outcomes for a given patient, participants were asked to "select which medicine worked better and by how much" (Figure 3B). If they answered that Medicine 1 or Medicine 2 worked better they were prompted to answer by how many points better. If they said that Medicine 1 and 2 worked exactly the same the program automatically entered zero points difference. Participants worked with 8 scenarios each of which represented a different patient. For each of the 8 scenarios the baseline trend was chosen randomly along the length of the baseline trend function; no two participants ever saw the exact same baseline trend.

Results

Do people tend to alternate or perseverate? Figure 4 shows a histogram of the number of alternations or switches between medicines across all scenarios and across all participants. Since there are 14 days there could be up to 13 alternations. As is obvious from the histogram, the most common strategy was to only switch once (e.g., try Medicine 1 for 7 days and then try Medicine 2 for 7 days).

In Figure 4 there is also a second peak representing scenarios in which participants alternated between the two medicines on every single trial (number of alternations = 13). Out of the total 1001 scenarios across 152 participants, there were 118 scenarios in which participants alternated exactly 13 times. 10 participants accounted for 68 of these 118 instances. Another 117 participants never alternated 13 times within a single scenario. In sum, most of these instances of high number of alternations can be attributed to a relatively small percent of participants.

Additionally, out of a total of 1001 scenarios across all 152 participants, there were also 29 scenarios in which participants did not alternate at all (e.g., just tried Medicine 1 for all 14 trials). 22 of these 29 instances were committed by just 6 participants. Since is not possible to know which medicine works better if only one medicine was tried, these instances are not plotted in Figure 4 and are omitted from future analyses.

In sum, in the current task most participants perseverated, but there is a minority who frequently alternated.

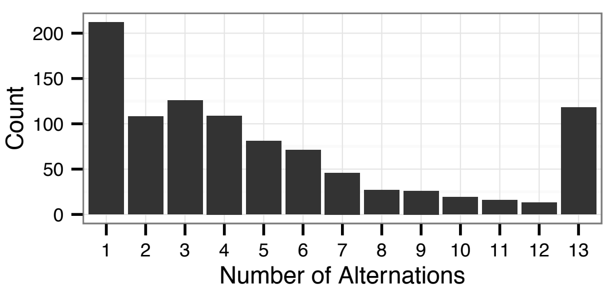

Figure 4. Histogram of number of alternations per scenario.

Do people who alternate make better inferences? Accuracy of inference was assessed in two ways. The first accuracy measure was binary - whether the participant inferred the correct direction (e.g., Medicine 1 works better than Medicine 2). This meant that inferences in which participants inferred that Medicine 1 and Medicine 2 worked exactly the same were ignored.

Figure 5 plots the percent of correct direction inferences by the number of alternations; more alternations was associated with a higher likelihood of inferring the correct direction. The error bars in Figure 5 are 95% confidence intervals, but they do not account for repeated measures. The error bars for 8-12 alternations are particularly wide because there are few numbers of scenarios in which participants alternated 8-12 times (see Figure 4).

A logistic regression tested whether more alternations was associated with a higher likelihood of inferring the correct direction. The regression had random effects for the intercept and random effects for the slope of number of alternations to account for repeated measures. The slope was significantly positive: 95% CI=[0.12, 0.23].

The second measure of accuracy was log absolute error. Perseveration is expected to produce high error both in the right direction and in the wrong direction (e.g., Figure 1A and the corresponding experiment if Medicine 2 was tried first). To capture both types of error this measure uses the absolute deviation from the correct answer. For example, if Medicine 1 reduced pain by 5 points relative to Medicine 2 and a participant inferred that Medicine 1 increased pain by 20 points, the absolute error was 25 points. Additionally, this measure uses the log of the absolute error because the absolute error was skewed with some very high errors.

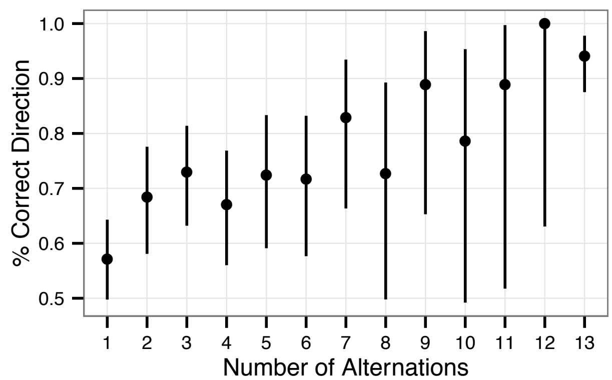

Figure 5. Percent of Inferences that were in the Correct Direction by Number of Alternations.

Figure 6 plots the error by number of alternations within a given scenario, with jitter on both axes to reduce overplotting. There is a clear trend such that the amount of error decreases with more alternation. A linear regression with by-subject random effects for the intercept and slope of number of alternations tested whether there was less error with increasing number of alternations. This decreasing slope was significant, 95% CI=[-0.11, -0.06]. To ensure that this effect was not driven by the scenarios in which participants alternated 13 times, the same regression was performed on the scenarios with 1-7 alternations with the same results: 95% CI=[-0.14, -0.06].

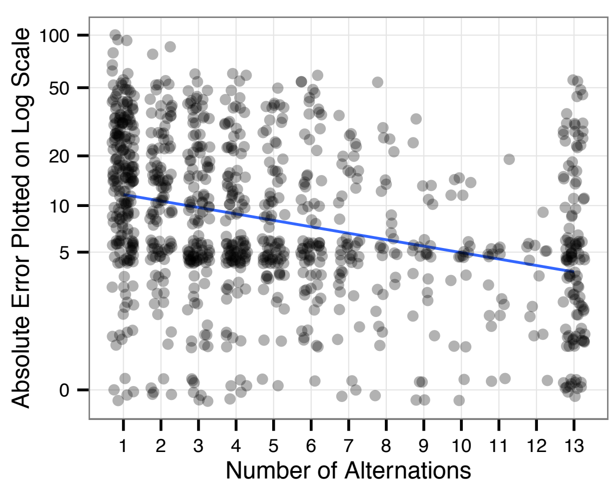

Figure 6. Absolute Error by Number of Alternations.

Experiment 2

Experiment 1 found that alternation was associated with more accurate inferences of the direction and size of the causal effect. However, it is possible that people who tend to alternate happen to be better at this information search task, but that alternating itself does not cause inferences to be more accurate. In Experiment 2 this was tested by forcing participants to either alternate or perseverate.

Another explanation for Experiment 1 is that alternating does indeed improve causal inference, but the improvement is due to a cognitive factor (e.g., alternation produces better memory), not that alternating produces cleaner data in an autocorrelated environment. To test this possibility, Experiment 2 compared the autocorrelated environment from Experiment 1 with a “random” (stable) environment in which the baseline trend varied randomly from day to day. If a cognitive benefit of alternating over perseverating is the only reason for the difference from Experiment 1, then the same difference in accuracy would appear in the random condition in Experiment 2. Alternatively, if the reason that alternation produces better inferences in the autocorrelation condition is because it results in cleaner data, then there would no longer be a benefit of alternation in the random condition. When the trial order is random both perseveration and alternation have exactly the same probability that the baseline would happen to be confounded with the medicine choice (e.g., Figure 7). In the random environment it should be fairly hard to tell which medicine produces lower pain because there is considerable variation in the pain scores and only a 5 point difference between the two medicines.

Methods

Participants 100 participants were recruited. 12 returned the hit with partial completion, resulting in a total of 112 participants. All data were analyzed from all participants.

Design The design of the study was a 2 Amount of Alternation (alternate vs. perseverate; between-subjects) × 2 Baseline Trend (autocorrelated vs. random; within subjects). Participants were randomly assigned either to alternate (switch back and forth between the two medicines resulting in 13 alternations across 14 days) or perseverate (try Medicine 1 for 7 days and then try Medicine 2 for 7 days).

The baseline trends were created in the following way. Every other participant was assigned to the alternate or perseverate condition. A pair of participants (e.g., Participant 1 and 2) received exactly the same baseline trends. Every pair was assigned four autocorrelated baseline trends like in Experiment 1. Then a parallel set of randomized baseline trends was created by randomizing the order within each of the four autocorrelated trends. Each pair of participants received a unique set of baseline trends.

Participants worked with all eight scenarios in blocks of autocorrelated vs. random. Half the participants received the autocorrelated block first and half received the random block first. Aside from these differences Experiment 2 was the same as Experiment 1.

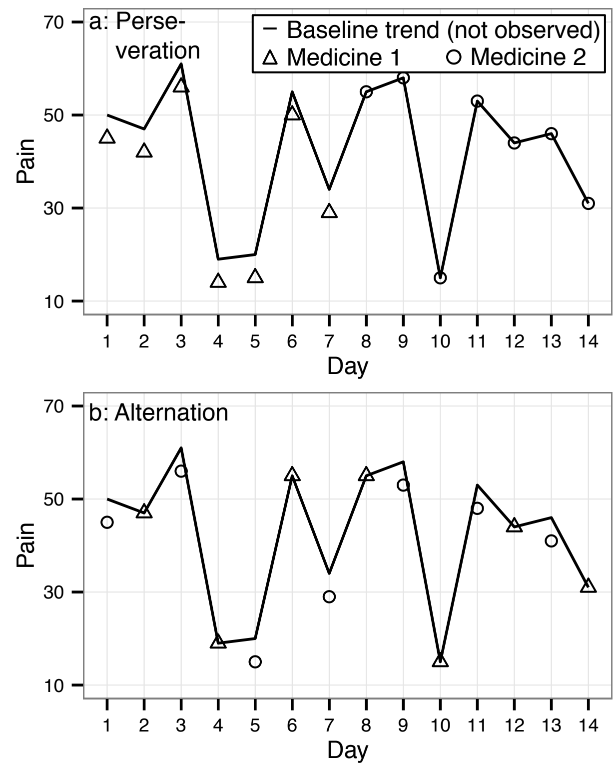

Figure 7: The Effect of Perseveration and Alternation in a Random Environment

Results

Just as in Experiment 1, the data were analyzed two ways, in terms of percent of inferring the correct direction and mean absolute error (see Table 2 for summary statistics).

Table 2: Results of Experiment 2

Condition	Alternate	Perseverate
	Percent Correct Direction	
Autocorrelated	91%	57%
Random	65%	68%
	Mean of Absolute Error	
Autocorrelated	13	22
Random	20	17

Percent of Inferences in the Correct Direction A logistic regression with random effects on the intercept and the slope of Amount of Alternation revealed that participants who alternated were much more likely to infer the correct causal direction than those who perseverated, 95% CI=[0.16, 0.30]. The slope of alternation is in the same unit as in Experiment 1; alternation was coded as 13 and perseveration as 1. In fact, the percent correct for the alternate and perseverate conditions (91% and 57%) were very close to the percent correct of participants who alternated 13 times vs. 1 time in Experiment 1 (Figure 5).

The same regression showed no effect of alternating vs. perseverating in the random condition, 95% CI =[-.05, .03]. Furthermore, there was a significant interaction between the amount of alternation and the type of baseline trend 95% CI of interaction term = [0.17, 0.32].

Log Absolute Error All of the same effects were also seen when analyzing the log of absolute error. Parallel regressions found that 1) there was a significant effect of alternation in the autocorrelated condition, CI=[-0.08, -0.03], but 2) there was not a significant effect of alternation in the random condition, CI=[-0.01, 0.04], and 3) there was a significant interaction between the amount of alternation and the type of baseline trend, CI=[-0.09, -0.05].

In sum, Experiment 2 found that alternating causes an improvement in accuracy and that the improvement only occurs in an autocorrelated environment.

General Discussion

The current research investigated how people learn about the relative efficacy of two causes when the effect was autocorrelated - underwent higher and lower periods. Autocorrelated environments are extremely common; one's blood pressure, mood, an investment, or almost any other real-world variable undergos high and low periods. When an effect is autocorrelated, repeatedly trying one cause and then repeatedly trying another (perseverating) is risky because the causes may be confounded with other changes over time. Alternating reduces the likelihood of confounding.

However, in the current experiments most participants did not appear to be aware of the benefits of alternating. In Experiment 1 most participants perseverated. Higher amounts of alternating were associated with more accurate inferences. In Experiment 2, forcing people to alternate produced substantially better inferences than forcing people to perseverate; the benefit of alternating is not limited to people who self-generate the alternating search strategy.

One could argue that it is not always feasible to alternate. For example, some drugs (e.g., antidepressants) can take weeks to start to work and can stay in the body for long periods of time, making it infeasible to repeatedly alternate between the two drugs. It can also be unethical to switch from a treatment that appears to be working. However, there are many real-world situations in which alternating is possible such as for medicines with limited-duration effects, and alternating is a common strategy in applied behavior analysis (Barlow & Hayes, 1979).

There are a variety of open questions about when people are more likely to alternate and how the choice to alternate vs. perseverate in the real world could affect causal inference. First, alternating matters in autocorrelated environments but not in stable environments. Do people start to alternate upon becoming aware that the environment is autocorrelated? However, note that the hot hand fallacy implies that people tend to believe that the environment is autocorrelated even when it is not (Wagenaar, 1970).

Second, autocorrelated environments in which the effect goes through high and low periods may potentially be interpreted as the cause having increasing or decreasing effectiveness over time (i.e., tolerance and sensitization, Rottman & Ahn, 2009). Are people able to disentangle these two causal mechanisms? Third, given a longer period of information search, would people who initially perseverate ever realize that their interventions could be confounded by the baseline and start alternating more frequently? Fourth, do people have explicit beliefs about whether alternation or perseveration is more useful? Fifth, how do people actually make the judgment of which cause works better given the data they observe, and is the process the same given alternated vs. perseverated data? Finally, are there cognitive costs to alternating (Arrington & Logan, 2004)?

In conclusion, information search in autocorrelated environments is common. The current findings raise the possibility that in some situations accurate causal inference may be facilitated by nudging people to alternate more than they would on their own. Teaching people to recognize when to alternate may be a useful debiasing strategy both for lay people and possibly also for professionals who perform single-subject type experiments such as doctors who practice personalized medicine.

References

Arrington, C. M., & Logan, G. D. (2004). The cost of a voluntary task switch. *Psychological Science*, *15*(9), 610–5. doi:10.1111/j.0956-7976.2004.00728.x

Barlow, D., & Hayes, S. (1979). Alternating treatments design: One strategy for comparing the effects of two treatments in a single subject. *Journal of Applied Behavior Analysis*, *12*(2), 199–210.

Biele, G., Erev, I., & Ert, E. (2009). Learning, risk attitude and hot stoves in restless bandit problems. *Journal of Mathematical Psychology*, *53*(3), 155–167. doi:10.1016/j.jmp.2008.05.006

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–9. doi:10.1038/nature04766

Hagmayer, Y., Meder, B., Osman, M., Mangold, S., & Lagnado, D. (2010). Spontaneous causal learning while controlling a dynamic system. *The Open Psychology Journal*, *3*, 145–162.

Hills, T. T., & Hertwig, R. (2010). Information search in decisions from experience. Do our patterns of sampling foreshadow our decisions? *Psychological Science*, *21*(12), 1787–92. doi:10.1177/0956797610387443

Rottman, B. M., & Ahn, W. (2009). Causal learning about tolerance and sensitization. *Psychonomic Bulletin & Review*, *16*(6), 1043–9. doi:10.3758/PBR.16.6.1043

Wagenaar, W. A. (1970). Appreciation of conditional probabilities in binary sequences. *Acta Psychologica*, *34*, 348–356.

Yi, S. K. M., Steyvers, M., & Lee, M. (2009). Modeling Human Performance in Restless Bandits with Particle Filters. *The Journal of Problem Solving*, *2*(2), 81–102.