2015

Reasonable Reasoner: The Influence of Intervention Strategy, System Parameters and their Representation on Causal Understanding

Beth Cristina Bullemer
Wright State University

REASONABLE REASONER: THE INFLUENCE OF INTERVENTION STRATEGY, SYSTEM PARAMETERS AND THEIR REPRESENTATION ON CAUSAL UNDERSTANDING

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

By

BETH CRISTINA BULLEMER
B.F.A., Carnegie Mellon University, 2004

2015
Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

AUGUST 28, 2015

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Beth Cristina Bullemer ENTITLED Reasonable reasoner: The influence of intervention strategy, system parameters and their representation on causal understanding BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science.

Valerie L. Shalin, Ph.D.
Thesis Director

Scott N.J. Watamaniuk, Ph.D.
Graduate Program Director

Debra Steele-Johnson, Ph.D.
Chair, Department of Psychology

Committee on
Final Examination

Valerie L. Shalin, Ph.D.

Dragana I. Claflin, Ph.D.

John M. Flach, Ph.D.

Robert E.W. Fyffe, Ph.D.
Vice President for Research and
Dean, Graduate School

ABSTRACT

Bullemer, Beth Cristina. M.S. Department of Psychology, Wright State University, 2015. Reasonable reasoner: The influence of intervention strategy, system parameters and their representation on causal understanding

The following study assessed how contingency and delay influence people's reasoning strategy and outcomes after interacting with a representation of a discrete and continuous system environment, in the context of controlling hypertension. The related causal reasoning and system dynamics research adopt different measurement paradigms and employ different system dynamics, making it difficult to resolve the empirical findings. Specifically, the causal reasoning literature has traditionally considered systems in which previous inputs do not influence future outcomes (e.g., a discrete system condition) while the system dynamics literature removes this constraint (e.g., a continuous system condition). Also, the system dynamics literature has focused on the ability to control pre-specified systems, whereas the causal reasoning literature has focused on the ability to discover and identify causal relationships. To examine reasoning under conditions comparable to hypertension management, I asked participants to consider causal scenarios involving causal variables (e.g., treatment options) with different amounts of contingency and delay in relation to a known outcome variable (i.e., level of blood pressure) with the representation of either a discrete or continuous system condition.

The findings address the relationship between causal attribution and system control, highlighting the effect of the system representation and dynamics on both reasoning behavior and outcomes, and challenging whether the efforts to build reasoning theory based on the combination of simplified paradigms paradoxically result in artificially complex problems and misleading theory. Participants' use of more observation-dependent intervention strategies with the discrete system condition indicates that they were aware of and responding to salient information. Additionally, differences in information accessibility explain why more extreme causal attributions were observed with the continuous system condition. Independent of system condition, specific intervention strategies (observation-independent and treatment-biased strategies) led to higher causal attributions, again reinforcing that system representation and underlying system dynamics directs reasoning outcomes.

Keywords: contingency, delay, strategy, system, causal attribution, system control

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

I. INTRODUCTION

People reason causally about simple and complex systems every day. This reasoning often originates with a problem to control (e.g., Why is my washing machine overflowing? Why does my car get better gas mileage in the summer? Why do I have a headache?). Consider the following scenario relating to hypertension, a problem dealt with by 29% of adults, 18 years and older, in the United States (e.g., Ostchega, Yoon, Hughes & Louis 2008; Yoon, Burt, Louis, & Carroll, 2012).

"Your blood pressure has been consistently elevated for the past year. Your doctor writes you a prescription for a common blood pressure lowering medication. Scared by this diagnosis, you obediently take your medication as prescribed and randomly refrain from indulging in a few of your unhealthy habits every now and then (e.g., bacon double cheeseburgers, reality TV on your sofa, and hourly doses of caffeine). Fortunately, your blood pressure measurements retract to a normal level after several weeks. Your problem appears to be resolved, meaning you know that something you have changed affected your condition. So, what do you do next? Do you continue to take your medication diligently and determine to forgo your past indulgences forever in order to maintain this normal blood pressure level? Alternatively, since you are fond of cheeseburgers, reality TV and caffeine, do you try to figure out if your medication alone affects your condition?"

A significant amount of psychological research potentially describes what people may do in the above situation. The causal reasoning literature has generated theoretically driven findings related to how people learn associations (Gopnik, Sobel, Schulz, & Glymour, 2001), use heuristics (Mayrhofer & Waldmann, 2011), and are influenced by delay (Lagnado & Speekenbrink, 2010; Shanks, Pearson & Dickinson, 1989) and instability (Rottman & Keil, 2012). The associated research methods traditionally manipulate a small number of task variables or parameters, which are presented using simplified, discrete systems under the hope of developing theory by aggregating the findings of numerous simple studies. However, people often experience problems much more complex and intertwined than addressed within causal reasoning literature.

The system dynamics literature starts by considering how people interact with continuous systems with multiple underlying parameters. Realistically complex, applied decision making situations include how people diagnose an unknown disease (Kleinmuntz & Thomas, 1987), fight forest fires (e.g., Brehmer, 1989), and manage product inventories (Diehl & Sterman, 1995). However, numerous manipulations of the underlying system parameters obscure the relationship between reasoning behavior in these applied situations and the basic findings from the causal reasoning literature.

This study borrows from both research methods to better understand how people reason when encountering paradigms in daily life with a variable whose causal efficacy is unknown and whose effects may appear after an unknown period of delay. These are factors pertinent to hypertension management, which is a particularly challenging task requiring patients, physicians and caregivers to reason with the diagnosed patient's cardiovascular system, an unobservable, but complex environment. As the causal

reasoning literature has traditionally considered system conditions which previous events do not influence future outcomes (i.e., discrete systems) while the system dynamics literature removes this constraint (i.e., continuous systems), I considered both a discrete and continuous system condition within the same experimental hypertension management task.

Recommended treatment plans for hypertension are not equally effective across patients (e.g., Chobanian et al., 2003). To examine reasoning under comparable conditions, I asked participants to consider three patient scenarios each involving a single causal variable (i.e., a treatment option) and a known effect variable (i.e., blood pressure level). Each patient scenario differed relative to the frequency that the causal variable produced an effect or by the frequency that the effect occurred in the absence of the causal variable. Also, I manipulated the amount of delay applied to the patient scenarios, as differences in temporal interactions between treatments and blood pressure level add additional complexity to hypertension management.

Below I review the causal reasoning research that addresses the influence of contingency, delay, contextual cues, and intervention on reasoning with a single cause variable. Additionally, I highlight the methodological differences between the causal reasoning and system dynamics research and review two studies that have considered continuous systems within the causal reasoning framework. But first, let us review what I mean by a *discrete* system and a *continuous* system.

Definition of a Discrete and a Continuous System

I will use the definitions for a discrete and continuous system as they are presented within the causal reasoning literature. However, differences in terminology

and definitions exist between literatures; so let me clarify what I mean when I refer to a discrete and continuous system. In the system dynamics literature, a system's representation is the manner in which the underlying parameters and dynamics of a system are sampled and presented to the observer. In this context, discrete refers to a sampling procedure that presents isolated feedback or snap shots of the underlying system environment, whereas continuous refers to constant feedback. This is not the meaning of discrete and continuous that I use here. According to this systems dynamics definition, both the referenced literature and the system conditions that I apply in this study only consider discrete system representations. Instead, the specification and discussion of the system condition in this document refers to aspects of the underlying dynamics, and their implications on the visual representation of the system environment.

Figure 1 identifies four different possibilities for system dynamics relative to two dimensions (Flach, 2015): the independence of outcomes (independent or dependent) and the adaptability of outcomes (stationary or non-stationary). The dimension of independence categorizes the outcomes associated with system environments as either independent from one another (e.g., whether or not you flip a heads or tails is *unrelated* to previous coin tosses) or dependent on one another (e.g., your blood pressure level after exercising is *directly related* to your starting blood pressure level). Trial or time dependences between outcomes are a fundamental difference between causal reasoning and system control research, which I will discuss in more detail when comparing causal reasoning and system dynamics research. Much of the causal reasoning literature does not consider the influence of time-based parameters, like delay, and research that

considers these parameters either violates the independent framework (becoming a discrete-like system) or has issues relative to internal validity.

The dimension of adaptability identifies whether the outcomes are stationary (e.g. the outcome of a coin toss is *always* either heads or tails) or non-stationary (e.g., the effect of exercise on my blood pressure today *differs* from its effect 10 years ago). Both the discrete and continuous systems mentioned within this document only consider stationary outcomes (past and present research included), meaning the parameters that define the systems' dynamics are fixed or constant throughout the learning experience.

So, the primary distinction between discrete and continuous systems as defined here, is the independence of outcomes. Figure 1 highlights that quadrants A and C distinguish the discrete and continuous systems mentioned within this document. Also, it is important to mention that neither of these systems truly reflects the human cardiovascular system (quadrant D), which is a potential limitation of this study.

	Stationary	**Non-stationary**
Independent	A) **Traditional Discrete Systems**/ Bernoulli Process	B) Multiple Bernoulli Processes
Dependent	C) **Continuous System/ Discrete-like (and the Discrete system in this study)** Fixed parameters (relative to either time or trial)	D) **Cardiovascular System/**Variable Parameter (relative to either time or trial)

Figure 1. Possible system dynamics of underlying environmental parameters. Adapted from Flach (2015).

Causal Reasoning

The causal reasoning literature focuses on how people establish the source(s) of effects or determine the likelihood that an effect will occur. Until recently, the typical research paradigm involved the presentation of a series of observations of the underlying parameters of the environment presented in a discrete manner. With such discrete system representations, each observation is independent of the next, so the interpretation of these observations (an assessment of causality) is the same regardless of whether the observations presented as a collection or a sequence. The scope of the reasoning task limits search for plausible causal candidates (Buehner & Cheng, 2005, p. 145). Researchers generally restrict the problem space to the relationship between a single causal variable and a known effect. Furthermore, researchers seldom combine more than two dimensions that may affect causal attributions. Relative to the multi-dimensional setting of hypertension management, this bottom-up approach results in a disjointed set of research finding regarding: contingency, delay, contextual cues, and direct engagement through intervention on the system.

Contingency. The causal reasoning literature considers differences relative to the frequency of an effect (*e*) given the presence (*P(e|c)*) and absence (*P(e|~c)*) of a causal variable *(c)*, and the *contrast* between these two values: ΔP[1] = *P(e|c) - P(e|~c)*. Although contingency is a common manipulation, researchers have not systematically explored consistent sets of contingency manipulations in isolation, with other manipulations, or with a consistent dependent measure between studies.

[1] $\Delta P = 0$ when *P(e|c) = P(e|~c),* meaning the probability of an effect is the same regardless of whether the suspected causal variable is present.

Buehner, Cheng, and Clifford's (2003) series of experiments is one of the few attempts to examine how a range of contingency values influences causal attribution (see also Wasserman, Elek, Chatlosh, & Baker, 1993). Buehner et al. (2003, experiment 1) examined 15 experimental conditions that assessed combinations of five contingency values (1.00, .75, .50, .25, .00) applied to the presence and absence of a suspected causal variable. Using a within-subject design, participants passively observed precisely eight applications and eight absences of the suspected causal variable presented randomly for each experimental condition. The intervention (causal variable (*c*), no causal variable (~*c*)) and the outcome (effect (*e*), no effect(~*e*)) appeared on the same trial, with the 16 observations presented serially rather than simultaneously.

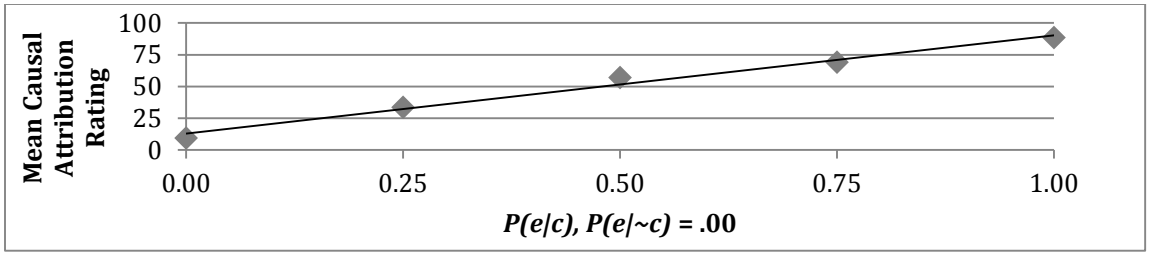


Figure 2. Positive slope associated with Buehner et al. (2003, experiment 1) findings when *P(e|c)* varies and *P(e|~c)* is held constant at 0.

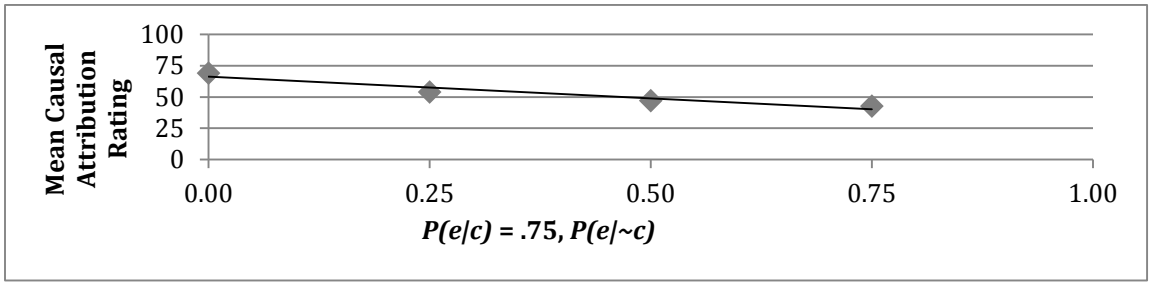


Figure 3. Negative slope associated with Buehner et al. (2003, experiment 1) findings when *P(e|~c)* varies and *P(e|c)* is held constant at .75.

Figures 2 and 3 depict the mean causal attribution values, the primary dependent measure, relative to a subset of the contingency manipulations considered. With both figures, causal attributions weakened as the ΔP approaches zero. When the probability of

an effect in the absence of the causal variable (*P(e|~c)*) was held constant at zero, causal attributions increased in combination with increases in the probability of an effect in the presence of the causal variable (*P(e|c)*) resulting in the positive slope depicted in Figure 2 (n = 260, SE = 1.39, β = 77.40). Conversely, increasing the value associated with the presence of an effect in the absence of the causal variable (*P(e|~c)*) reduces causal attributions when the probability of an effect given the causal variable (*P(e|c)*) is held constant at .75, forming the negative slope shown in Figure 3 (n = 208, SE = 1.73, β = -34.46).

These outcomes correspond to previous findings (e.g., Wasserman et al., 1993) and Buehner et al.'s predictions with the exception of when the probability of an effect was the same given the presence and absence of the causal variable (ΔP = zero). When the effect is occurring frequently, but ΔP = zero (as with the *P(e|c)* = .75, *P(e|~c)* = .75 condition), people fail to realize that the effect is present more often than not across all trials and ascribe causality to the suspected causal variable rather than concluding the variable is non-contingent. This finding is indicative of the base-rate neglect/fallacy (i.e., focusing on, or placing greater weight on, specific information (like the overall probability of an effect (*P(e)*) rather than considering all the information presented (e.g., the probability of a causal variable (*P(c)*) or the probability of the effect given the causal variable (*P(e|c)*). Buehner et al. (see also Buehner and Cheng, 2005) posit that participants' higher than expected attributions of non-contingent variables reflect a) an excessive demand on working memory or b) ambiguity in the semantics of the dependent measure. However, 16 trials may not be enough time to detect a difference between the probability of an effect given the cause (*P(e|c)*) and the probability of an effect without

the cause ($P(e|\sim c)$) across all contingency values. (See Appendix A for a detailed analysis that considers whether differences can be detected between two proportions.)

Delay. People rely on the cause preceding the effect as a cue to identify causal relationships (Lagnado & Sloman, 2006). A common paradigm for the study of human causal reasoning with delay is the experimental design used by Shanks et al. (1989). Participants freely administer a suspected causal variable (e.g., pressing a space bar) and observe whether an effect occurs (e.g., triangle appearing on a screen). After performing this task for a controlled amount of time or number of applications specified by the experimenter, participants determine the degree to which the suspected causal variable causes the effect.

Initial findings indicated that increasing the amount of time between the application of the cause and the onset of the effect disrupts causal attribution (e.g., Shanks et al., 1989). However, the implementation of the delay manipulation, rather than delay itself may be responsible for these findings. Specifically, the seemingly continuous learning period (e.g., 120 seconds) was divided into separate time segments (e.g., 1 second) to maintain a discrete rather than continuous learning environment in which there is no carry-over between trials. This time interval was extended for the delay condition (e.g., 2 seconds) relative to the control condition (e.g., 1 second). Yet, only the first action within each time interval was registered by the computer and produced an outcome with both the delay and non-delay conditions. So, only one corresponding outcome was produced regardless of whether the causal variable was applied two or ten times. Longer delay intervals provide more opportunity for participants to apply the potential causal variable, which may have reduced the observed contingency because the probability of

the cause (*P(c)*) increased relative to the probability of the effect (*P(e)*). Thus, the observed decrease in causal attribution may have been a function of a mathematical decrease in contingency rather than an increase in delay.

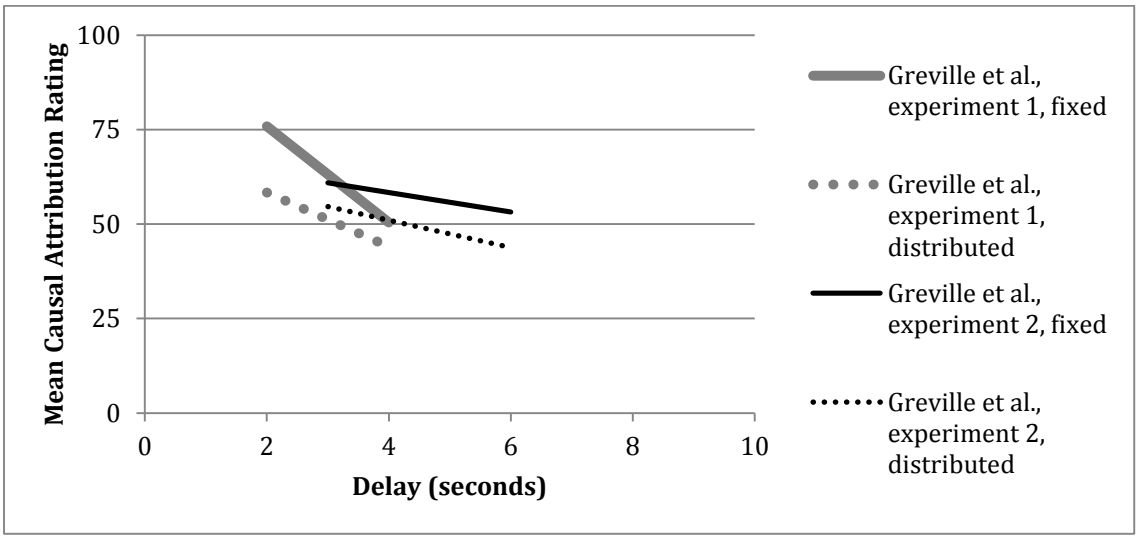


Figure 4. Negative slopes associated with Greville et al. (2010) findings.

To rule out this confounding explanation, Greville and Buehner (2010) maintained Shanks et al.'s (1989) basic spacebar/triangle experimental design, with an implementation of delay that is more analogous to, if not categorically a continuous system. Greville and Buehner did not parse their learning period into separate time segments, allowing for carry-over or overlap between participants' application of the suspected causal variable and outcomes. Figure 4 depicts Greville and Buehner's (2010, experiment 1 and 2) findings. With all four of Greville and Buehner's conditions, the negative slopes illustrate the same pattern of decline in causal attributions with increases in the interval of delay as observed by Shanks et al. (1989). Also, Greville and Buehner's findings suggest that a consistent temporal relationship between cause and effect increases causal attributions as compared to normally distributed (variable) delay intervals. Similarly, Lagnado and Speekenbrink (2010) found that causal attributions

declined when exposed to variable delay intervals, but that these declines in attribution did not affect participants' ability to correctly differentiate the active casual variable from a set of three options.

Other causal reasoning research has suggested that contiguity affects people's reasoning capabilities in conjunction with other factors (e.g., context, extraneous variables, ability to interact with the system), that moderate the relationship between delay and causal attribution. For instance, Buehner and May (2002) demonstrated that context could offset the impact of delay by providing participants with temporal expectations. This finding is compelling given that the common paradigm for contiguity within the causal reasoning literature (i.e., Shanks et al., 1989) essentially asks people to reason about nothing (the relationship between a space bar and the appearance of a triangle).

Contextual Cues. The use of substantive contextual scenarios has varied substantially within the causal reasoning literature. Some studies have asked participants to reason about abstract relationships, like which shapes make a box glow (Frosch, McCormack, Lagnado & Burns, 2012) and if a space bar makes a triangle appear on a computer screen (Shanks et al., 1989). Other studies have incorporated scenarios that are more realistic such as asking participants to identify which bacteria causes stomach cramping (Lagnado & Speekenbrink, 2010) or which variables prevent a rocket from launching (Lagnado & Sloman, 2002).

Differences in realism can influence people's reasoning behavior and interpretations of causality. For instance, Johnson-Laird, Legrenzi, and Legrenzi (1972) famously demonstrated the influence of context on participants' reasoning tendencies.

Using the Wason card task, they illustrated that participants' selections differed when exposed to symbolic representations (i.e., letters and numbers) and realistic representations (i.e., letters and stamps). Similarly, Buehner and May (2002) illustrated that contextual assumptions pertaining to delay could influence reasoning by presenting participants with scenarios in which delay between cause and effect was expected (i.e., launching a grenade) and unexpected (i.e., turning on a light bulb). Although Buehner and May found participants' causal attributions were higher experimental conditions with no delay regardless of the contextual scenario, their contextual manipulation that insinuated some delay (grenade) did mediate the decline in causal attributions with conditions containing delay.

Einhorn and Hogarth (1986) suggested that the influence of cues present in realistic representations could depend on whether or not they increase or decrease uncertainty. So, contextual cues presented in experimental paradigms, or applied contexts may also impair successful identification of causal relationships. Specifically, false assumptions pertaining to a specific context may prevent the recognition of causal relationships. For instance, if I assume that all medication works like Advil on a headache (i.e., feel permanent relief after 30 minutes), I may falsely conclude that other medications are ineffective if they do not exhibit these characteristics.

Though contextual cues can misguide participants, removing them altogether, or simplifying reasoning paradigms may inadvertently remove essential information that assists people in recognizing causal relationships. This is a criticism of research that considers decomposed or discrete system environments independent of the amount of contextual framework added to the paradigm. Arguably, the potential for cue elimination

is why system dynamics researchers refrain from altering the contextual framework of their task environments and is why it is essential to assess reasoning capability with this complexity added back in.

Intervention. One approach to the study of causal reasoning controls the learning environment in order to ensure a consistent participant experience, typically by presenting an equivalent number of observations related to the presence and absence of the causal variable (i.e., $P(c) = P(\sim c)$). This allows researchers to attribute differences in data to variables other than differences in the frequency of causal variables under the guidance of strategy. The alternative to a fixed approach (e.g., $P(c) = P(\sim c)$) allows participants to interact freely with a system for a set amount of time or trials (i.e., free-operant learning). Causal reasoning with delay research has extensively employed free-operant learning environments (e.g., Buehner & May, 2003; Shanks et al., 1989; Wasserman & Neunabber, 1986), and free-operant learning environments are equally as prevalent in the causal reasoning with contingency research (e.g., Jenkins & Ward, 1965; Wasserman et al., 1993) as controlled learning environments (e.g., Buehner et al., 2003; Perales & Shanks, 2003).

Steyvers, Tenenbaum, Wagenmakers, and Blum (2003, experiment 2) demonstrated that people were markedly more successful in identifying causal structures when allowed to interact with the system representation. Lagnado and Sloman's research (2002, 2006) also indicates that people are more successful when learning through active intervention, as opposed to passive observation. However, they propose that confounded temporal cues (Lagnado and Sloman, 2006) or the relative frequencies of causal variables (e.g., $P(c)$ relative to $P(\sim c)$) (Lagnado and Sloman, 2002) enable correct causal structure

identification rather than the act of interacting with the system representation. I examined a series of treatment intervention strategies to evaluate Lagnado and Sloman's (2002) latter explanation. Specifically, I considered what intervention strategies minimize standard error (Appendix B) and an observation-dependent strategy relative to two different observation-independent strategies (Appendix C). These theoretical analyses highlight that intervention strategy shapes the amount and type of information available to the observer, which supports Lagnado and Sloman's (2002) proposition, as well as provides a viable explanation as to why Hagmayer, Meder, Osman, Mangold, and Lagnado (2010) found differences in reasoning outcomes as a function of intervention strategy. Additionally, these analyses emphasize that limiting research to the study of one treatment intervention strategy (i.e., the unbiased intervention strategy used in controlled learning environments) only realizes a simple effect, which may not generalize to (or even bear on) the complete set of treatment intervention strategies. Or, potentially, this simple effect is true, but not important relative to other observations, e.g., the effect of intervention strategy dwarfs everything else.

Causal Reasoning vs. System Dynamics Research

The previous sections focused on causal reasoning research tasks that employ discrete system in their test environments. Research that uses continuous system (e.g., system dynamics) also demonstrates that people struggle to control system outputs with longer delay intervals, and are affected by/recognize differences in the underlying parameters of the system (e.g., contingency). However, fundamental differences between causal reasoning and system dynamics research considering, including dependencies

between trials and dissimilar dependent measures, make it difficult to align and compare findings.

Dependent measures. The causal reasoning literature focuses on whether people can explicitly identify causal relationships given a set of observations. These measures ask participants to either identify a system's causal structure (e.g., Rottmann & Keil, 2012), assign causal attributions (e.g., Shanks et al., 1989), or identify the probability that the same outcome would occur in a set number of instances (e.g., Buehner et al., 2003; Greville & Buehner, 2010). Conversely, the system dynamics literature focuses on the manner in which participants *manage* a system with causal relationships over time. Moreover, people participating in system dynamics studies are commonly informed of details relating to the system's underlying causal structure prior to and when interacting with the system representation (e.g., Sterman, 1989; Jensen & Brehmer, 2003). This design property of system dynamics studies arises from the intent to measure whether participants can manipulate a system to a point of equilibrium and what inputs they make to do so, rather than ascertaining if participants can explicitly identify the system's causal relationships. The use of divergent performance measures and paradigms limit the ability to characterize the relationship between understanding and behavior.

Dependencies between trials. In defining the difference between decision making with discrete and continuous system, Edwards (1962) describes dynamic decision making tasks as sequential sets of decisions with each decision producing an effect contingent on previous decisions as well as behavior of the system. In closed looped systems, feedback loops inform these sequential decisions providing participants with insight regarding their status in relation to their goal. Thus, past decisions influence

participants' future decisions as they attempted to accomplish the goal of the task. Feedback loops can improve learning or performance over a series of trials (Paich & Sterman, 1993), as long as the gleaned information is correctly perceived and interpreted (Sterman, 1989).

The methodological assumption of trial independence within the causal reasoning literature precludes examination of continuous systems with feedback. Most causal reasoning research compares people's assessments with outputs from mathematical models, which designate a correct assessment based on the rules of probability (e.g., Power PC theory, Rescorla-Wagner model). Often these models stipulate that an observation of an effect given a cause must be independent from other observations of the effect. Thus, the effect returns to a baseline state after each observation and without the application of an effective causal variable, the effect is absent or assumed to be zero. This assumption of independence is inconsistent with the closed loop systems considered within the system dynamics literature, in which previous interventions might still be acting on the effect or have yet to influence the effect when subsequent causal variables activate. Consequently, adherence to the constraint of independent observations precludes system conditions that are continuous in nature.

Causal Reasoning with Continuous Systems

Recently, a couple of conventional causal reasoning researchers (Hagmayer et al., 2010; Rottman & Keil, 2012) removed the constraint of trial independence to understand if and how people reason with interdependent trials. Both Hagmayer et al. and Rottman and Keil measured causal structure identification, rather than causal attribution. Neither investigated the influence of delay on people's causal reasoning tendencies. Only

Rottman and Keil incorporated contingency into their experimental design, but did so as a confounding variable rather than a manipulation. Lastly, although Hagmayer et al. allowed participants to freely interact with the system, Rottman and Keil controlled the learning environment by exposing participants to a predetermined set of observations. Together, their findings suggest that previous system states (Rottman & Keil, 2012) and intervention strategy (Hagmayer et al., 2010) influence the manner in which people reason, which in turn raises questions about the validity of outcomes associated with experimental designs that limit free-operant learning and assume trial independence.

Hagmayer et al. Hagmayer et al. (2010) asked participants to consider the influence of three casual variables on an effect (i.e., three light rays, all of which positively influenced the effect, to produce an optimal transmitter level in order to kill cancerous areas in a rat's brain). Participants freely administered interventions (light rays) over the span of 40 trials. Hagmayer et al. manipulated the gain associated with each of the variables rather than contingency, meaning a positive effect always followed an intervention, but the degree to which each causal variable reduced the cancerous area differed. Following the learning phase, Hagmayer et al. measured whether participants could identify the correct causal structure from two options (termed the demonstrative task) and assessed whether participants could successfully select combinations of the causal variables to reach various target values. Next, Hagmayer et al. mapped participants' success in these demonstrations to their selection strategies in the preceding learning task. Not only did they discover that participants adapted their strategy based on the system underlying parameters, but that selection strategy in the learning task affected their success in the demonstrative tasks.

Rottman and Keil. Rottman and Keil (2012) presented participants with either two or three variables. No variable had an identifiable role, such that any variable could influence or be influenced by another variable. This is a departure from the traditional causal reasoning paradigm, which typically excludes causal discovery by defining the roles of the variables, where the only issue is the presence or absence of causality. The number of variables determined the number and type of causal structure manipulations considered with each experiment. After observing a series of non-randomized trials arranged to prompt conclusions indicative of either trial independence or carry-over, Rottman and Keil measured participants' interpretations of the underlying causal structure. These interpretations suggest that people integrate information presented across trials when reasoning, rather than assume that trials are independent. Also, Rottman and Keil demonstrated that exposure to trials inconsistent with the underlying causal structure (i.e., an unstable condition with less contingency associated with causal variables or more contingency associated with non-causal variables) decreased participants' ability to correctly identify causal structures. Rottman and Keil (2012, experiment 1a, p. 103) commented that participants "endorsed links that do not exist" in the unstable condition, which suggests that people are more willing to error on the side of causality than discounting a variable's causal influence entirely.

Summary

I have discussed findings involving factors related to hypertension management (i.e., contingency, delay, contextual cues, and intervention) in isolation and highlighted the primary differences between research considering discrete and continuous systems. People tend to infer causality when the relationship between variables frequently occurs

in a stable manner and the cause is proximally near the effect. A violation of these assumptions affects causal attribution and structure identification, as well as system control (e.g. Brehmer, 1989; Diehl & Sterman, 1995). However, certain manipulations including contextual expectations, and most importantly for the present study, learning through intervention can moderate the negative influence of instability and extended delay intervals.

Although the system dynamics literature has considered the influence of delay on systems and intervention strategy, these studies focused on whether people can control a system without assessing how effectively managing a system relates to an explicit understanding of the system. Because the system dynamics literature has focused on the ability to control pre-specified systems and the causal reasoning literature has focused on the ability to discover and identify causality, both have failed to address the relationship between understanding and behavior, and unfortunately it is not feasible to draw comparisons between the two literatures. The two examples of causal reasoning with continuous systems only emphasizes the need for causal reasoning research that considers experimental designs with interrelated trials that takes into account the relationship between intervention strategy and reasoning outcome. Moreover, neither the causal reasoning nor the system dynamics literature has assessed differences in causal attribution as a function of strategy with a discrete or continuous system condition with various underlying parameters (e.g., levels of contingency and amounts of delay). Consequently, both areas of research are incomplete relative to the problem scenario of how a person diagnosed with hypertension determines how, and what treatments affect their chronic condition.

Study Objectives and Predictions

This study considered two different system conditions to better understand how people reason about a potential causal variable (proposed treatments) and a known effect variable (blood pressure level) in a contextual scenario involving the control of hypertension. Causal reasoning researchers have commonly used two variable environments when assessing the influence of delay (e.g., Greville & Buehner, 2010; Shanks et al., 1989) and various degrees of contingency (e.g., Buehner et al., 2003; Perales & Shanks, 2003; Wasserman et al., 1993) on causal attribution. So, I intentionally chose to explore how people reason with two variables, rather than exploring environments with additional variables in order to relate my findings back to this basic research. Allowing participants to freely interact with test environments with different system dynamics allowed me to consider the implications of system condition on selection strategy, and selection strategy on causal attribution.

The discrete system condition is consistent with existing causal reasoning research, such that the effect of a treatment applied in one trial is independent from the participant's treatment applications in other trials. This condition connects my study paradigm to the existing causal reasoning research by examining the influence of delay and contingency with a discrete system condition. The second condition extends the causal reasoning literature by examining the same delay and contingency conditions explored in the first condition, but in conjunction with a continuous system condition. The continuous system condition is analogous to those found within system dynamics research, as there were interdependencies between trials such that participants' actions on previous trials influence the effect of future treatment applications. A comparison of the

two conditions allows us to align results from the disparate methodological approaches used within causal reasoning and system dynamics literatures.

The problem space of hypertension management did require a combination of methods previously used in experimental research in a novel paradigm. The potential implications of specific design implementations are discussed in detail in the method section. Below, I present my hypotheses and research questions related to the experimental manipulations of contingency, delay, and their interaction. Additionally, I identify pertinent research questions related to the discrete and continuous system conditions, and intervention strategy.

Contingency and delay. There is conflicting evidence related to how people might respond to the interaction between contingency and delay. Findings from a subset of studies suggest people ascribe more causality to non-causal variables when confronted with uncertainty (e.g., Gao, Nitzany, & Edelman, 2012; Lagnado & Speekenbrink, 2010; Rottman & Keil, 2012). These findings support a prediction that participants' causal attributions will increase for variables with lower contingencies (i.e., lower ΔP values) with increased delay intervals. Conversely, reasoning theories (e.g., Occam's razor, Take the Best) suggest that situations with more uncertainty might assist participants in identifying potential causal variables with lower ΔP values in favor of parsimonious and definite explanations of causality (or no causality). In these instances, participants may use unknown causal variables to explain inconsistent observations (Rottman & Ahn, 2011), rather than attribute causality to less contingent causal variables. A decrease in participants' causal attributions for manipulations with increased amounts delay and lower ΔP values will support this perspective. Given findings relating to the influence of

delay across conditioning, system dynamics, and causal reasoning literatures, I expect that increases in delay will result in greater differences in causal attributions regardless of its direction.

In the absence of a significant interaction between the delay and contingency experimental manipulations, I expected causal attributions to reflect the differences in the ΔP values between the contingency manipulations, such that lower ΔP values would correspond to lower causal attributions (Buehner et al., 2003). Also, I anticipated that delay alone would generate outcomes inconsistent with participants' expectations. With this inconsistency, current research suggests that participants would struggle to adequately manage the system (e.g., Brehmer, 1989) and be less likely to attribute the causal variable with the effect (e.g., Greville & Buehner, 2010).

Discrete vs. continuous system. To predict whether or not participants will perform better or worse in the continuous system condition as opposed to the discrete system condition is largely dependent on which paradigm one adheres. System dynamics literature would argue that simplifying the problem space by removing the interdependencies between trials renders it devoid of essential contextual cues (and makes the findings rather meaningless when relating the findings back to systems that are continuous in nature). If this is the case, participants will be better equipped to correctly identify the underlying contingency with the continuous system condition than with the discrete system condition. Contrary to this, the causal reasoning literature would argue that a discrete problem space simplifies the reasoning task, so is not only a logical place to initiate research, but is also more likely to facilitate causal attribution. Thus, this study

addressed the research question of how system condition affects outcome measures common to the respective literatures within a consistent reasoning framework.

Intervention strategy. Although free-operant learning environments are more characteristic of the environments experienced by hypertensive patients applying new treatment regimens, they add an unpredictable amount of variability that may reduce the affect of other, controlled manipulations applied by the experimenter. This is why the number and type of interventions participants apply and subsequently observe when learning about causal scenarios is a contextual factor often controlled or ignored by causal reasoning researchers.

An analysis of the standard error of the difference between proportions illustrates that the underlying contingency value affects what intervention strategies will provide the most information regarding the influence (or lack of) of the causal variable on the effect outcome (Appendix B). Unbiased intervention strategies, with which the probability that the causal variable is applied is equivalent to the probability that it is withheld (i.e., $P(c) = P(\sim c)$), provide the most information in scenarios with no prior assumptions about a variable's efficacy. But the application of unbiased intervention strategies is difficult to rationalize (and is less informative) if one assumes that a particular variable or intervention (e.g., a treatment recommended by your trusted physician, or specific action relative to the contextual scenario, as with diet to hypertension) will have a positive effect.

Additionally, the underlying contingency value directs intervention strategy if people react to specific observations, but has no bearing on selections made independently of a system's outputs (Appendix C). For example, a strategy that only

applies the treatment when the blood pressure is elevated provides less opportunity to test outcomes associated with the treatment as the contingency value increases. This is because increasing the combined contingency, or increasing the likelihood of an effect will occur regardless of the source (i.e., treatment or ~treatment), increases the frequency that the blood pressure observation is in the normal range. For this reason, differences in system control are related to both intervention strategy and contingency. Exclusive application of the causal variable (i.e., a treatment-biased intervention strategy) will yield the lowest possible blood pressure levels when the probability of an effect given the cause is greater than the probability of the effect in the absence of that cause ($P(e|c) > P(e|\sim c)$). Moreover, the converse, lower blood pressure levels when refraining from applying the causal variable when the probability of an effect given the cause is less than the probability of the effect in the absence of that cause ($P(e|c) < P(e|\sim c)$). When the probability of an effect given the presence and absence of the cause is equivalent ($P(e|c) = P(e|\sim c)$), intervention strategy will not effect the system control measure, or the probability that the effect is observed.

An analysis of observed treatment intervention selections presents an opportunity to discover common reasoning strategies. Further, a review of the relationships between the experimental manipulations and treatment intervention strategy, and treatment intervention strategy and causal attribution will provide a clearer direction for future research. If intervention strategy is unrelated to both the experimental manipulations and causal attributions, this will provide justification for restricted reasoning paradigms and ultimately deemphasize the need to employ free-operant paradigms and consider intervention strategy moving forward. Conversely, an increased emphasis on measures

that reflect both behavior and understanding is justified if intervention strategy is related to either the experimental manipulations or reasoning outcomes as suggested by Hagmayer et al.'s (2010) findings.

II. METHOD

Participants

A total of 96 students enrolled in an undergraduate psychology course at Wright State University were recruited to participate in this study for course credit. The data from 24 participants were excluded because they did not meet the English proficiency or conscientious responder criteria (i.e., attempted to apply the treatment). When a participant failed to meet these criteria, I excluded their data and re-ran the condition with a new participant.

Table 1 summarizes the demographic distribution of the 72 participants whose data were considered across analyses. About half of the 72 participants were females (n = 35), whom were evenly distributed between the discrete (n = 18) and continuous (n = 17) system condition by happenstance. Participants were between 18 and 45 years old (M = 20.25, SD = 4.41) and had normal or corrected normal vision. A total of 18 participants reported that English was not their native language. In addition to basic demographic information, participants provided information related to their health beliefs and experiences managing their own or someone else's medical condition. I did not exclude participants based on this information as preconceived notions and experiences with specific medical treatments and conditions are arguably realistic in natural settings. Over half of the participants (n = 42) reported managing their own or another person's medical condition. These conditions ranged from acute medical conditions (e.g., cold and flu) to chronic diseases.

I performed one-way ANOVAs for a subset of the demographic questionnaire items (i.e., age, sex, grade level, and ESL) to determine whether the randomly assigned groups of participants were equivalent. Also, I compared the distribution of participants between the system conditions using independent samples *t*-tests. Participants were evenly distributed between experimental manipulations. These analyses are presented in Appendix D, along with summaries of participants' responses to the Health Belief Questionnaire measures and participants' performance on the English proficiency measure.

Table 1
Participant Demographic Distribution separated by System Condition

	Discrete System	Continuous System
Age		
< = 21	32	30
> 21	4	6
Sex		
Female	17	18
Male	19	18
Grade		
Freshman	26	24
Sophomore	9	8
Junior	1	4
Senior	0	0
ESL	9	9
College Affiliation		
Business	2	1
Education	2	2
Engineering and Computer Science	12	10
Liberal Arts	3	7
Nursing	2	3
Math and Science	6	3
University College	9	10
Health Management Experience*	18	25

Note. *Have you personally managed your own or someone else's medical / condition?

Manipulations and Conditions

Contingency. I employed three contingency manipulations to explore differences relative to the value associated with the probability that the causal variable produces the effect and the probability that the effect occurs in the absence of the causal variable. I started with the contingency manipulation, $P(e|c) = .75$, $P(e|\sim c) = .00$ as it is frequently used within the causal reasoning literature by researchers considering the influence of various contingency values, as well as those measuring the influence of delay on causal attributions (e.g., Buehner et al., 2003; Greville & Buehner, 2010; Shanks et al., 1989). The second manipulation alters the probability that an effect will occur in the absence of the causal variable, so that the difference between the probabilistic values is zero (i.e., $P(e|c) = .75$, $P(e|\sim c) = .75$; and $P(e|c) - P(e|\sim c) = \Delta P = 0$). The third manipulation lowers the probability that the effect will occur in the presence of the causal variable, but retains the same value as the first manipulation relative to the probability that the effect will occur in the absence of the casual variable (i.e., $P(e|c) = .25$, $P(e|\sim c) = .00$). Buehner et al. (2003, experiment 1) also evaluated these last two manipulations in their assessment of several contingency manipulations.[2] For the remainder of the document, I will refer to these contingency manipulations as 75|0, 75|75, and 25|0.

The two methods used for implementing contingency are random (Allan, Tangen, Wood, & Shah, 2003; Buehner et al., 2003; Buehner & May, 2003; Jenkins & Ward,

[2] Buehner et al. (2003, experiment 1) assessed 15 contingency conditions using a within subjects design. The learning task was presented using a discrete system, in which participants were not allowed to freely apply the causal variable and the influence of delay was not considered.

1965; Perales & Shanks, 2003; Shank et al., 1989; Wasserman et al., 1993) and fixed assignment (Rottman & Keil, 2012). With random assignment, researchers randomly generate and display a sequence of observations at the rate of contingency for each participant, whereas this sequence of observations is specified a-priori and remains consistent between participants with fixed assignment. In combination with a free-operant learning task, the advantage of the fixed assignment approach is that there is considerably less variation between the actual, observed contingency values across all of the possible intervention distributions (i.e., strategies) as compared to the fixed assignment implementation. However, these actual, observed contingency values are equal to the contingency specification only with specific intervention distributions and otherwise, will be greater than the specified contingency value. Appendix E contains a more complete analysis that considers the fixed and random assignment approach in the specification of contingency.

Although random assignment as the method for contingency specification is more common within the causal reasoning literature, the implementation of contingency here was linked to participants' hypertensive treatment interventions (i.e., fixed assignment), rather than randomly generated (i.e., random assignment) in order to retain some consistency between learning experiences. With manipulations in which the $P(e) = .75$, the null effect (~*e*) was observed in combination with the participants' 2, 6, 10, and 14th selection of the associated treatment option (either the application of the treatment and/or the absence of the treatment depending on the manipulation). Similarly, the effect occurred on participants' 1, 5, 9, and 13th selections of the apply treatment option when $P(e|c) = .25$. Table 2 presents the actual contingency values for each of the contingency

manipulations averaged across participants' intervention strategies and the delay manipulations.

Table 2
Actual Contingency Values relative to Intended Contingency Values for the Discrete and Continuous System Conditions

	Intended		Discrete System		Continuous System	
	P(e|c)	*P(e|~c)*	*P(e|c)*	*P(e|~c)*	*P(e|c)*	*P(e|~c)*
75|0	.75	.00	.77 (.03)	.00 (.00)	.76 (.05)	.00 (.00)
75|75	.75	.75	.78 (.05)	.76 (.04)	.78 (.06)	.77 (.12)
25|0	.25	.00	.37 (.06)	.00 (.00)	.38 (.07)	.00 (.00)

Note. Values represent the mean contingency value across participants' intervention strategies. *SDs* shown in parentheses.

Delay. Researchers exploiting a causal reasoning paradigm explore the implications of delay with duration-based based learning tasks (Allan et al., 2003; Buehner & May, 2003; Shanks et al., 1989; Wasserman & Neunabber, 1986). Actual time determines participants' exposure to the learning environment (e.g., two minutes) and specification of the delay interval is in increments of seconds. This duration-based approach provides the greatest opportunity for control of the delay interval, given delay is a factor of time. However, applied situations requiring causal assessments and management extend beyond these abbreviated durations, as is the case with hypertension management. Although running experiments over longer durations is possible (e.g., Shiffrin & Schneider, 1977), the general absence of such studies insinuates that there are disadvantageous to such endeavors, such as cost, participant retention, and inclusion of uncontrollable variables.

Trial-based learning tasks are the alternative. The use of trial-based learning environments is customary in the system dynamics literature, as well as within areas of the causal reasoning literature that control for delay. The temporal ambiguousness of trials allows researchers to assign timeframes contextually appropriate to their

experimental scenarios, which arguably increases external validity. However, there are several implications of a trial-based approach to the implementation of delay. First, the speed at which participants progress through the learning task will determine the actual time between trials, which hampers comparison to duration-based research. Second, the use of trials to specify the delay interval prohibits the exploration of a no-delay manipulation, which serves as the control condition within duration-based research. However, the use of a one-trial delay condition is a justifiable, if not better, alternative given some delay is inherent to most complex, continuous systems. Third, implementing delay within a trial-based framework necessitates carry-over between trials. This overlap between action-outcome pairings makes the system categorically less discrete, but Rottman & Keil's (2012) findings suggest that people inherently interrelate information presented on contiguous trials in discrete-trial based tasks. Also, overlaps between action and outcome pairings already exist within duration-based, free-operant learning tasks that consider delay (e.g., Buehner & May, 2003, experiment 2; Greville and Buehner, 2010). Lastly, delay/time becomes a symbolic, rather direct representation. Although the inability for trials to adequately represent time is a concern, the use of trials to represent time is not novel to this experiment.

I used three delay manipulations to study the influence of delay between the administration of the treatment intervention and its effect on the patient's blood pressure: one-trial delay, two-trial delay, and three-trial delay. With the one-trial delay manipulation, the outcome associated with the treatment intervention was shown at the start of the next trial (i.e., if you take treatment on trial 1, you see the applicable effect on trial 2). With the remaining delay manipulations, the impact of treatment interventions

occurred on subsequent trials, but not on the next trial. With the two-trial delay manipulation, the outcome associated with the administered treatment intervention skipped a trial. With the three-day delay manipulation, the outcome was presented three trials after the application of the intervention (skipped two trials). The blood pressure observation associated with the null-effect (elevated blood pressure level with the discrete system and +2 with the continuous system) was not shown to participants until the delay manipulation allowed the administered treatment intervention to influence the effect (i.e., one trial delay/trial 2, two trial delay/trial 3, three trial delay/trial 4).

Figures 5 and 6 depict the delay manipulations in combination with the 75|0 contingency manipulation for both the discrete and continuous system conditions. For the purposes of illustration, in both instances, the apply treatment intervention was selected on trials 1 through 8 and the withhold treatment intervention was selected for the remaining trials. Appendix F presents the systems' behavior (blood pressure levels) in relation to additional treatment intervention scenarios. Table 3 contains the approximate delay durations associated with each of the trial-based delay manipulations observed in this study. Quantifying the actual time associated with each trial is necessary to compare the observed findings with research that uses duration-based, rather than trial-based learning tasks.

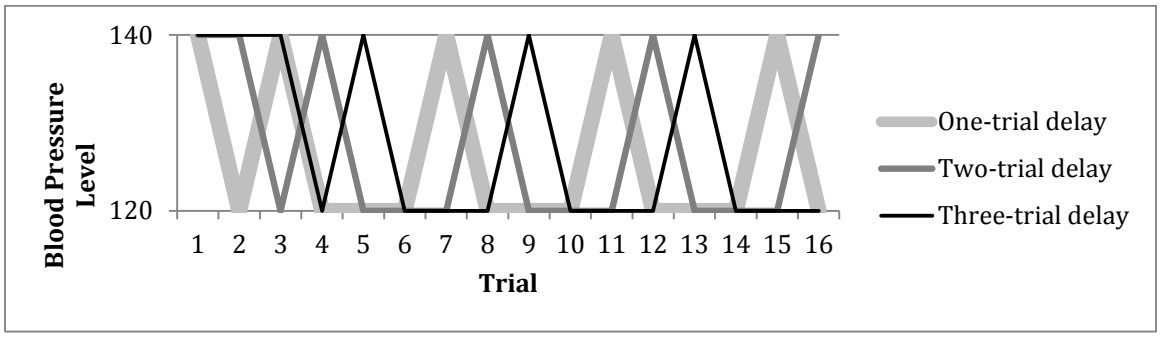


Figure 5. Delay manipulations with the discrete system condition with the 75|0 contingency manipulation. 140 and 120 represent an elevated and normal blood pressure levels respectively. Application of the treatment invention occurs on trials 1-8.

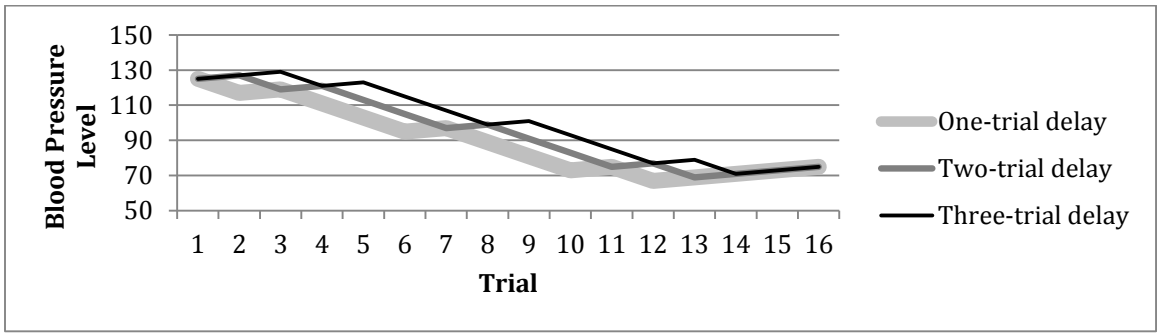


Figure 6. Delay manipulations with the continuous system condition with 75|0 contingency manipulation. Application of the treatment intervention occurs on trials 1-8.

Table 3
Approximate Duration associated with each Delay Manipulation in Seconds for the Discrete and Continuous System Conditions

Delay Manipulation	Discrete System	Continuous System
One-trial delay	3.90	3.82
Two-trial delay	7.80	7.63
Three-trial delay	11.70	11.45

Note. For the discrete system condition, the amount of time spent on each trial was just under four seconds across all contingency * delay treatment combinations (M = 3.90, SD = 1.16). For the continuous system condition, the amount of time spent on each trial was just under four seconds across all contingency * delay treatment combinations (M = 3.82, SD = 0.90). Approximate duration was calculated by multiplying the mean trial time by the number of trials of delay.

System conditions. The contigency and delay manipulations were applied to two system conditions: discrete and continuous. Half of the study's participants worked with each condition. With the discrete system condition, a patient's blood pressure level was

depicted as either an elevated or a normal state. Patients' blood pressure level started at and returned to the elevated state after each trial unless acted upon according to the contingency manipulation. As such, the treatment interventions worked on the effect in accordance to experimental manipulations without influence of prior treatment interventions. Although each intervention affected the system separately, there was carry-over between trials due to the trial-based implementation of the delay manipulation making this a partially-discrete condition. This partially-discrete condition is not unique to this paradigm and is consistent with Greville and Buehner's (2010) experimental design. Greville and Buehner allowed carry-over between trials to overcome the challenges (i.e., differences in contingency, differences in learning task duration) of studying delay with a completely discrete system.

With the continuous system, the trials were dependent such that an application of a treatment option on previous trials influenced the impact of the treatment application on subsequent trials (e.g., Edwards, 1962). The same carry-over between trials applies to the continuous system condition due to the delay manipulation as mentioned with my discrete(-like) system condition. However, the blood pressure measurements were continuous between trials as well, meaning the output on each trial was representative of the cumulative impact of the participant's actions. For this reason, the patients' blood pressure levels were presented on a continuous, 100-interval scale. The left and right halves of the scale were depicted as ranges containing normal (color coded as green) and elevated (color coded as red) blood pressure levels, respectively. Patients' blood pressure level started at the center of the elevated range. If the treatment influenced the patients' blood pressure according to the underlying contingency manipulation, the blood pressure

indication would shift 8 points to the left (i.e., toward the normal range), otherwise the blood pressure indication would shift 2 points to the right (i.e., toward the elevated range). This movement toward the elevated range (i.e., the gain value associated with the null effect) is representative of treatment decay or a patient's natural tendency toward blood pressure elevation. Appendix G provides further rationale for the selection of this -8:+2 effect to null effect gain ratio. If the level reached either bound of the scale, the blood pressure level would remain at the bound until patient's selection behavior in combination with the underlying contingency and delay manipulations warrented movement.

Measures

As discussed previously, a major difference between the causal reasoning and system dynamics literatures is the dependent measure. Causal reasoning focuses on the ability of people to explicitly describe either the direction or magnitude of the effect on a system, whereas the system dynamics research measures behavior, that is peoples' interaction with and ability to control a system. Both measures are appropriate to a hypertension management context. Peoples' interventional approach informs the evolution of their beliefs, whereas their overall success in controlling the system is potentially indicative of their cumulative understanding of the system. However, neither of these measures requires people to explicitly express this understanding, which is the benefit of attribution measures characteristic of the causal reasoning paradigm. In addition to these outcome measures, I collected data using a series of demographic and debrief questions. For brevity, the following section will only present the measures considered in the body of this document. Appendix H contains the descriptions of

additional measures such as selection diversity, selections based on observations, selections based on changes in observations, selections based on unexpected changes, and response times.

Causal attribution measures. I used participants' causal attributions from the judgment task to assess their explicit understanding of the situation (i.e., the causality of the treatment option). In the causal reasoning literature, researchers either specify causal direction (i.e., preventative or generative) (e.g., Buehner et al., 2003, experiment 1; Jenkins & Ward, 1965; Wasserman et al., 1993) or leave the directional nature of the scenario ambiguous (e.g., Allen et al., 2003, experiment 1; Perales & Shanks, 2003; Wasserman & Neunabber, 1986). Specifying a causal direction within the context of hypertension management (or in any reasoning domain) seems presumptive and unnecessary. First, medical treatments can exacerbate or generate symptoms despite their preventive intention (e.g., flu vaccines), so using a directional scale preemptively narrows the range of reasoning outcomes. Also, if reasoning outcomes are strictly preventative or generative, reasoners should be capable of making this distinction. Further, Wasserman and Neunaber (1986) found participants were less sensitive to differences in contingency and altered their judgment strategy with directionally minimized scales. For these reasons, I applied a measure that assesses both direction and magnitude of the potential causal relationship with an ordinal and ratio measure, respectively.

Ordinal measure. The ordinal measure describes participants' overall impression of how the treatments influenced a patient's outcome and was collected in the first part of the judgment task. Participants selected whether the treatment raised, lowered, or had no effect on the patients' blood pressure level.

Ratio measure. If participants selected 'no-effect' with the ordinal measure in the first part of the judgment task, I coded their responses as zero. If participants determined that a treatment 'lowered' or 'raised' their patient's blood pressure, they were asked to determine the degree of the treatments influence on a 100-point scale. To distinguish between the lowered and raised causal attributions, I inverted the causal attributions associated with 'raised' responses on a zero (never raised patients' blood pressure) to -100 (always raised patient's blood pressure) scale. No conversions were made to the causal attributions associated with a 'lowered' assessment.

System control measure. The amount of blood pressure decline (as opposed to elevation) is a system control measure that is compatible with both the discrete and continuous system conditions and is meaningful to the contextual scenario of hypertension management (see Appendix I for a comparison of the two options system control measures that I considered for this study). I measured the decline in blood pressure after 13 treatment interventions, rather than 15 in order to assess the same number of treatment interventions across the delay conditions. So, I compared the cumulative blood pressure values associated with trial 14, 15, and 16 with one-trial, two-trial, and three-trial delay manipulations, respectively. Although participants were not shown numeric blood pressure values, a numeric scale underlay the continuous system outcome graphic. I transferred participants' treatment intervention selections made in the discrete system condition to the continuous systems' numeric scale to allow for a system comparison. Importantly, differences in contingency are expected with this outcome measure. For example, patients' blood pressure levels readily lower with the 75|75 contingency manipulation as compared to the 25|0 contingency manipulation.

Intervention strategy measures. Participants' experiences with the learning tasks will vary based on the order and frequency that they administered the treatment intervention options (i.e., apply, withhold). The computer application recorded each of the participants' treatment interventions. I used these inputs to assess whether participants applied identifiable patterns of exploration (i.e., strategies) consistently within and between patient scenarios and considered whether these strategies were related to the experimental manipulations and treatment type (e.g., exercise). I also considered whether treatment intervention strategy affected participants' causal attributions and system control. I used two treatment intervention measures when categorizing participants' strategic behavior: the number of treatment interventions applied and the number of runs.

Number of treatment inventions applied. The number of treatment interventions applied tallies the number of times that a participant applied the treatment intervention option as opposed to withholding the treatment across the 16 trials. This measure was used to identify conscientious responders, as well. If participants withheld the treatment intervention option on all 16 trials, I excluded their data and re-ran the condition with a new participant as it is impossible to assess something's influence without ever sampling the phenomena.

Number of runs. The number of runs measure leverages part of the runs test that is used to detect non-randomness (Bradley, 1968). A run is defined as a series of the same input/value - in the present study the input is the treatment intervention option. So, the number of runs is the number of times the input selection changes. For example, let's pretend that I flipped a coin ten times and got the following output: HHHHTHHTTT with

H and *T* representing a heads and tails outcome, respectively. The number of runs for this sequence is four. The first run has a length of four (HHHH). The second run has a length of one (T). The third run has a length of two (HH) and the fourth run has a length of three (TTT). Generally, with runs test analyses, fewer runs are associated with a non-random process.

Demographic questions. In addition to providing their age, gender and major, participants answered questions that assessed their English proficiency, and their current knowledge and expectations related to medical treatments (see Appendix J for complete list of questions). I used these responses to verify that participants' language proficiency was adequate to understand the written instructions and to ensure that the participants randomly assigned to each condition were equivalent.

English proficiency. I presented participants with four definitions and asked them to identify the corresponding word from a set of six options. This method of measuring English proficiency is used on a language learning website (i.e., Level test) to ascertain readers' language ability in order to recommend appropriate instructional materials. The four definitions presented to participants were randomly selected from a set of 12 definition to word pairings. Half of these definition to word pairings were categorized as identifiable by readers at an intermediate level, while the other half an upper intermediate level. Given a large portion of the protocol involved reading in English, participants that answered two or more of these four questions incorrectly were excluded from the final data analysis and their condition was re-run with a new participant.

Medical experience. I asked participants a series of questions about their experiences managing medical conditions, treatment expectations, and beliefs related to medical treatments. Specifically, I asked about their expectations related to treatment onset and offset durations, as well as the percentage of treatment interventions that should positively act on a medical condition to consider the cumulative treatment experience as effective. To better understand participants' general beliefs about medicine, I asked participants to rate a subset of items from the Beliefs about Medicine Questionnaire (BMQ) (Horne, Weinman, & Hankins, 1999): if doctors had more time with patients, they would prescribe fewer medicines, doctors use too many medicines, medicines do more harm than good, and people who take medicines should stop their treatment for a while every now and again. These items consider medication overuse and harm, and were selected based on Horne et al.'s (1999, p.10) factor analysis results. Horne et al. (1999) suggests that people who feel more strongly about these items will be more inclined to 'seek alternative methods of treatments.' Relative to this study, participants with higher ratings on these items may apply the treatment less frequently when determining how the treatment options influence their patients' blood pressure.

Debrief questions. I asked participants to provide feedback pertaining to their reasoning experience with their last hypothetical patient. This series of questions (also shown in Appendix J) asked participants to discuss the strategy that they used during the learning and judgment tasks, their expectation of how the treatment option may have influenced the patient's blood pressure level at the start of the learning task, and to explain the underlying construct/system behavior. These questions specifically targeted a single patient scenario because I expected some participants to modify their behavior

according to the experimental conditions. Further, the last patient scenario was selected to minimize the amount of time between the participants' experience and their recollection of this experience.

Design

All participants learned about hypothetical patients' responses to a series of administered treatment interventions (i.e., potential causal variables) in a learning task and rated whether these treatment interventions affect the patient's blood pressure level (i.e., the effect) in a judgment task. Between patient scenarios, I revised elements of the cover story to highlight this transition and reinforce to participants that they were interacting with a new patient. The computer application randomly generated a name for each hypothetical patient by selecting from six options (i.e., Mr. Smith, Mr. Jones, Mr. Green, Mr. Nelson, Mr. White, and Mr. Tyler). Once used, the system did not present this name to that participant on subsequent tasks (i.e., sampling without replacement). I used the same technique to revise the names and illustrations representing the hypertensive treatment options (i.e., medication, healthy diet, and exercise) presented with each patient scenario. I manipulated the underlying contingency (a. 25|0, b. 75|0, c. 75|75) and the amount of delay (a. one-trial delay, b. two-trial delay, c. three-trial delay) to form nine treatment combinations (e.g., hypothetical patient scenarios). Table 4 presents these nine treatment combinations. I considered how these nine treatment combinations influenced peoples' reasoning ability separately with a discrete and a continuous system condition as a between subjects effect, and used participants' casual attributions, treatment intervention behavior, system control, and time as dependent measures.

Table 4
Learning Task: 3 (Delay) x 3 (Contingency) Design

	Contingency Manipulation		
Delay Manipulation	0 – 75|0	1 - 75|75	2 - 25|0
0 - one-trial delay	00	10	20
1 - two-trial delay	01	11	21
2 - three-trial delay	02	12	22

Note. 75|0, 75|75, and 25|0 represent the $P(e|c) = .75$, $P(e|\sim c) = .00$, $P(e|c) = .75$, $P(e|\sim c) = .75$, and $P(e|c) = .25$, $P(e|\sim c) = .00$) conditions respectively.

I assigned treatment combinations using a 3 x 3 random block partially confounded factorial design (RBPF) for each system condition (i.e., continuous and discrete). The confounded design allocated 1/3 of the treatment combinations (i.e., 3 patient scenarios) to each participant. This design avoids between subjects comparisons for the independent variables in question, and the potential for carryover with repetitions of the same value for an independent variable across the session. While this design allows for the assessment of main effects using a within subject error term, it also reduces the number of treatment combinations assigned to each participant to reduce fatigue. However, only half of the 4 *df* interactions in the confounded factorial are within subjects. A replication accounted for within this design switches the between and within subjects contrasts so that all contrasts can be tested with within subjects power, albeit for only half of the data (see Table 5). To eliminate the possibility of an ordering effect, I counterbalanced the treatment order. The three treatment combinations in each block can be presented in six different orders. Table 6 illustrates the orders for the treatment combinations in Block 1. I randomly assigned six participants in each block to view the treatment combinations in each of the six possible presentation orders. Figure 7 shows the distribution of the 72 participants amongst the system conditions, the six blocks and further divided into the six possible treatment orders.

Table 5
RBPF – 3^2 Design

	Treatment Combination (*Delay$_j$Contingency$_k$*)		
	1	2	3
Block 1	00	12	21
Block 2	01	10	22
Block 3	02	11	20
Block 4	00	11	22
Block 5	02	10	21
Block 6	01	12	20

Note. *Delay*$_0$, *Delay*$_1$, and *Delay*$_2$ represent the one-trial delay, two-trial delay, and three trial manipulations respectively. *Contingency*$_0$, *Contingency*$_1$, and *Contingency*$_2$ represent the *P(e|c)* = .75 *P(e|~c)* = .00, *P(e|c)* = .75, *P(e|~c)* = .75, and *P(e|c)* = .25, *P(e|~c)* = .00 manipulations.

Table 6
Presentation Orders for the Treatment Combinations in Block 1

	First	Second	Third
1	00	12	21
2	00	21	12
3	12	00	21
4	12	21	00
5	21	00	12
6	21	12	00

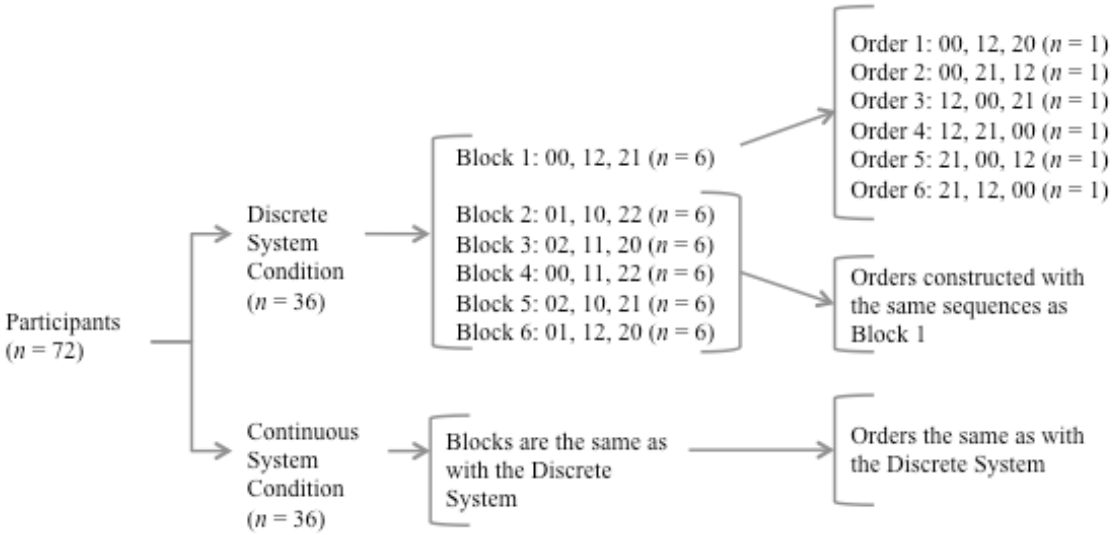


Figure 7. Breakdown of Participants into Conditions, Treatment Blocks, and Order of Treatments

Procedure

I conducted this study in sessions with multiple participants[3] individually working on laptop computers. Participants reviewed a consent form and offered their consent to participate before the start of the study (see Appendix K for the consent form). Participants were randomly assigned to a set of experimental treatment combinations and were asked to enter this treatment combination into the computer application. After entering this information, the computer application guided participants through an introduction familiarizing them with the cover story (i.e., physician managing patients blood pressure levels), as well as the upcoming learning and judgment tasks (see Appendix L). The final screen of this introduction (shown in Figure 8) presented participants with two task goals: a) figure out how the treatment is influencing the patient's blood pressure level and b) lower the patient's blood pressure as much as possible. In addition to presenting the task goals, this screen reinforced that each of the three patient scenarios (experimental treatment combinations) should be considered independent from one another, and insinuated the presence of temporal variability and external influences. I reiterated this information between patient scenarios, as well as changed the patient's name and hypertension treatment option (e.g., medicine, diet, and exercise) to emphasize the change in the learning task scenario. Participants were encouraged to ask questions during this introduction phase, as they were not given feedback from the facilitator (me) after this point.

[3] The final session only contained one participant. The remaining 16 sessions included between two and eight participants.

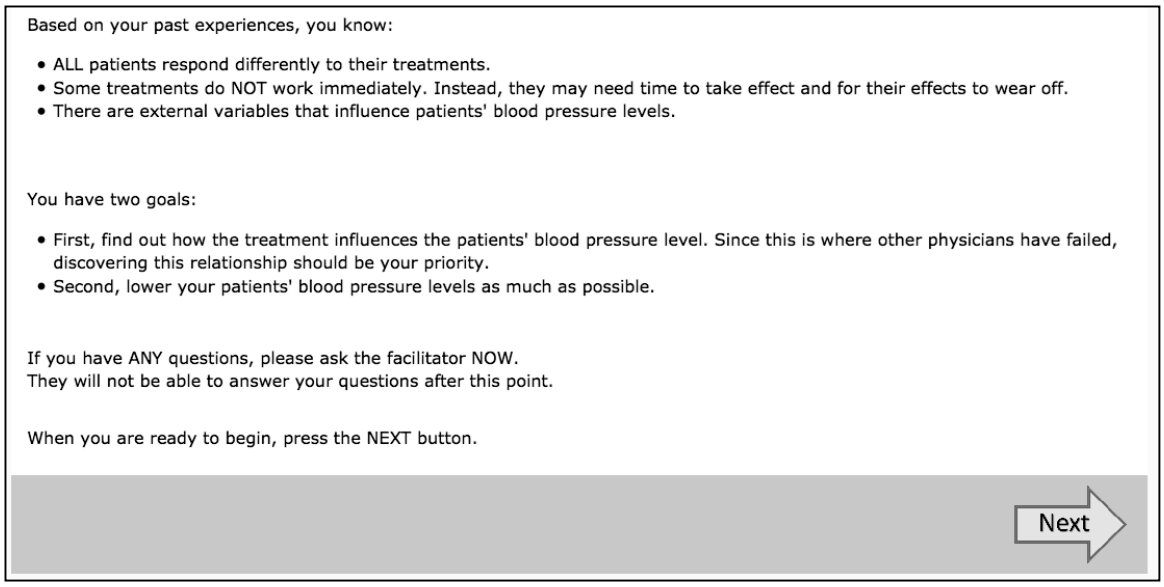


Figure 8. Task Goals and Contextual Information Screen

Following the introduction, participants continued to use the computer application to learn about their three hypothetical patients, again with three different experimental treatment combinations. Participants applied the randomly assigned hypertension treatment options (medicine, diet, and exercise) and monitored their patients' status over a series of trials. After each learning task, I asked participants to determine how the hypertension treatment influenced their patient's blood pressure level. Participants received feedback concerning the accuracy of their responses after they completed the judgment task related to their last hypothetical patient. This feedback screen also directed participants to an online demographic and debrief questionnaire. Participants self-progressed through their three patient scenarios and the questionnaire with no time limitation. The vast majority of participants completed both of these activities in less than 30 minutes.

Tasks

Learning task. Participants monitored their patients' status while administering 16 days (i.e., 16 trials) worth of treatment interventions that behaved in accordance with the experimental conditions. This number of trials is consistent with past studies that have used eight trials per intervention option (Buehner et al., 2003; Rottman & Keil, 2012). The user interface presented participants with two intervention options: apply treatment or do not apply treatment (Figure 9). In so doing, participants determined the frequency, as well as the order in which they applied (or chose not to apply) each treatment intervention. So, for each trial, the participant needed to determine whether or not to administer the treatment to the patient. Once the participant selected an intervention option, a next button activated allowing the participant to move to the next trial.

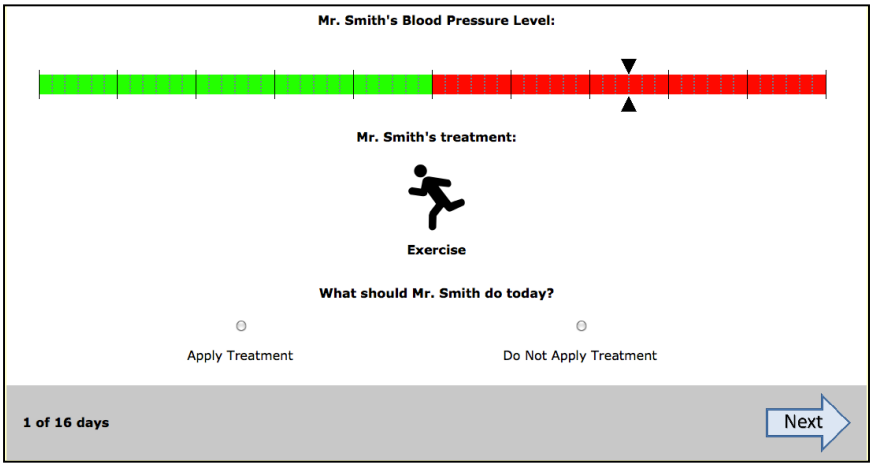


Figure 9. Learning Task Interface

With each new trial, the patient's status changed according to the participant's treatment intervention select and the underlying experimental manipulations of the patient scenario (i.e., contingency value, the amount of delay, and system condition). For the discrete system condition, I used a dichotomous widget to represent the patients'

blood pressure status (e.g., Novick & Cheng, 2004). This icon depicted the patient's blood pressure level as either normal (middle of the green range) or elevated (middle of the red range) (see Figure 10). I used the same widget in the continuous condition, but there were multiple intervals within the normal and elevated blood pressure ranges (see Figure 11). With both system conditions, the patients' status on the first trial began at the midpoint of the elevated blood pressure range. In addition to the patients' status, the indication of the number of remaining trials for each patient scenario updated with each trial. Participants could not move backwards or alter their intervention selections once they moved to a new trial.

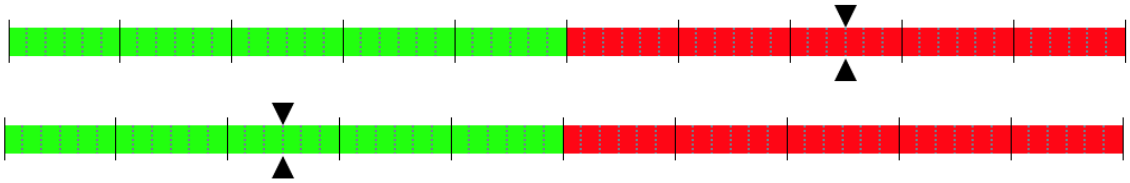


Figure 10. Elevated and Normal Blood Pressure Level Indication in the Discrete System Condition. The status indication is always positioned in the middle of the respective ranges.

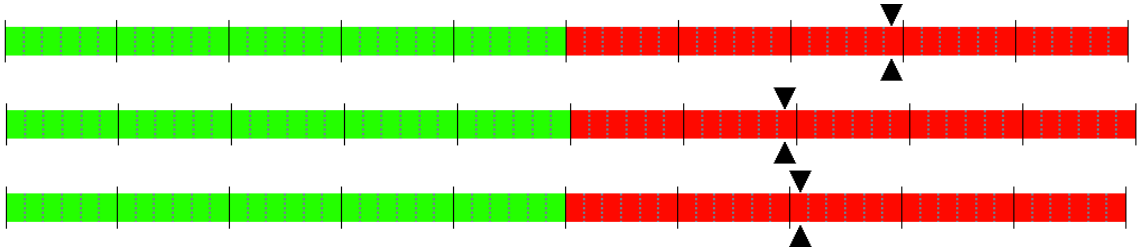


Figure 11. Elevated Blood Pressure Level Indications in the Continuous System Condition. The elevated status indication can have various states. The same is true for the normal level in the continuous state.

Judgment task. The judgment task followed each learning task. I asked participants to assess how the treatment option (i.e., the cause) affected the patients' blood pressure level (i.e., the effect) in a two-part task (e.g., Buehner et al., 2003, experiments 2 and 3). First, participants decided whether the treatment affected the

patients' blood pressure level by selecting one of three ordinal categories (i.e., the treatment lowered the patient's blood pressure, the treatment had no effect on the patient's blood pressure, or the treatment raised the patients blood pressure). If the participants determined that the treatment either lowered or raised the patient's blood pressure level, they were asked to rate the extent to which the treatment affects the patient's condition. This measurement was collected using a graphic rating scale ranging from zero (the treatment had no effect on the patient's blood pressure level) to 100 (the treatment always [raised or lowered depending on the previous selection] the patient's blood pressure level), with continuous whole numbers (e.g., Buehner et al., 2003; Greville & Buehner, 2010). Figures 12 and 13 show the interfaces related to first and second part of the judgment task.

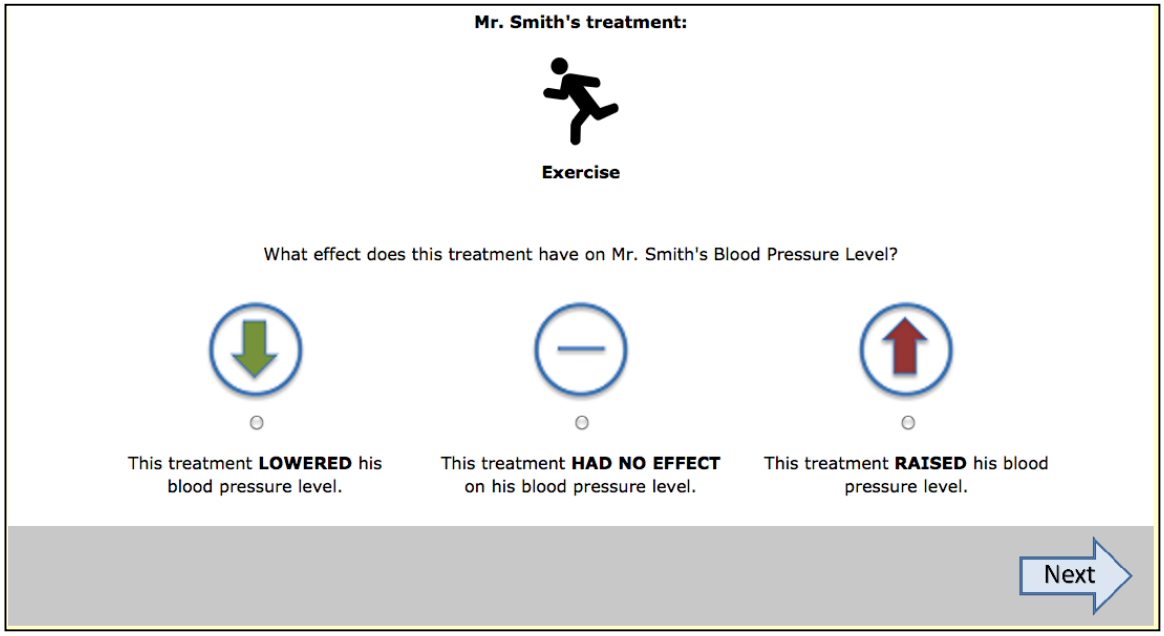


Figure 12. Judgment Task – Selection of Ordinal Category Interface

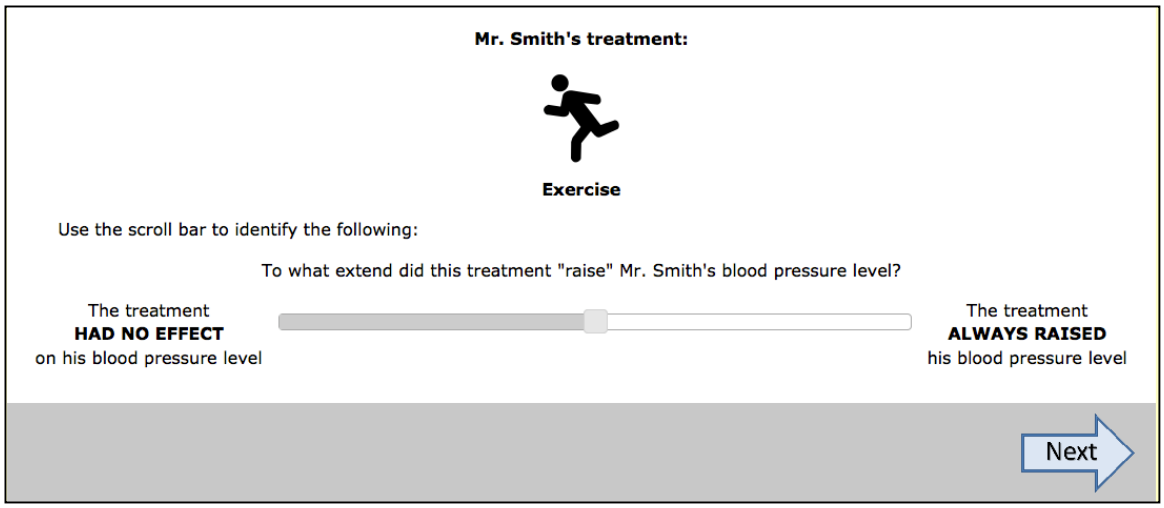


Figure 13. Judgment Task – Ratio Scale Selection

After completing the judgment task related to their third and final hypothetical patient, participants received feedback regarding the accuracy of their assessments across the session (Figure 14). The purpose of this feedback was to provide participants with some insight into their performance. However, I intentionally did not present this feedback throughout the session in order to reduce carryover effects, and to limit potential frustration and task abandonment that may accompany incorrect responses. Response accuracy was determined by participants' initial, ordinal assessment regarding the influence of the treatment on the patient's blood pressure relative to the underlying contingency manipulation. Specifically, the correct response for the conditions in which the probability of an effect given the cause ($P(e|c)$) was greater than the probability of an effect in the absence of the cause ($P(e|\sim c)$) (i.e., $\Delta P > 0$ as was the case with 75|0 and 25|0 contingency manipulations) was that the treatment *lowered the patient's blood pressure*. Further, *had no effect on the patient's blood pressure* was considered correct response when the probability of an effect given the cause ($P(e|c)$) was equal to the probability of an effect in the absence of the cause ($P(e|\sim c)$) (i.e., $\Delta P = 0$, the 75|75 contingency manipulation). The user interface indicated whether the participants'

responses were correct or incorrect, but did not provide rationale for this assessment to ensure participants remained ignorant of the experimental manipulations.

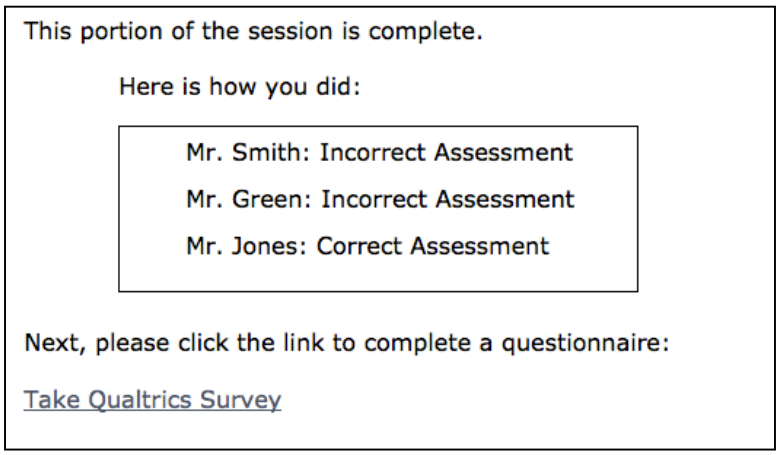


Figure 14. Performance Feedback Screen

III. RESULTS

I assessed differences in reasoning outcomes (i.e., causal attributions and system control) according to the experimental manipulations, the relationship between these reasoning outcomes, differences in intervention strategies, and the relationship between reasoning strategy and causal attributions. Additionally, I compared the findings from the present study to existing experimental research and reasoning theory, and reviewed participants' self-reported data. For quantitative analyses, I used an alpha level of .05, and performed normality (Appendix M) and homogeneity of variance (Appendix N) assessments.

To summarize, the omnibus analyses revealed differences in participants' causal attributions (Figure 15), but not success in controlling the system (Figure 16). Both system condition and contingency (Figure 15b), but not delay (Figure 15c), influenced participants' causal attributions. The observed pattern of causal attributions was consistent between the system conditions (25|0 < 75|0 and 75|75, 75|0 = 75|75), but the attributions of participants that interacted with the continuous system condition were more extreme. Participants' causal attributions negatively correlated with their system control with each level of delay in both system conditions (Figure 17). That is, higher attributions were associated with lower blood pressure values. Yet, this negative relationship only persisted with the 75|0 contingency manipulation in the discrete system condition, when the outcome data was separated by contingency and system.

Although neither system condition nor delay directly affected the system control measure (Figure 16), my consideration of intervention strategy did reveal behavioral differences between the system conditions, specifically relative to the strategic dimension of observation dependence. The use of reactionary, observation-dependent strategies was more prevalent with the discrete system condition with most of the contingency and delay manipulations (Figure 19). In contrast, differences in the distribution of the strategic dimension of intervention bias were not tied to the experimental manipulations (Figure 18). Yet, causal attributions differed according to both strategic dimensions with observation-independent and treatment-biased strategies leading to higher causal attributions (Figure 20).

Lastly, the observed causal attributions generally reflect the existing experimental research. The exception is participants' conclusion that the treatment associated with the 75|75 contingency manipulation was extremely efficacious. Participants' self reported assessment technique and expectations concerning treatment efficacy provide some rationale for this disconnect. But, causal attributions failed to consistently correspond to theoretical predictions, even when intervention strategy was taken into account, limiting the ability to assess what information contributed to participants' understanding of causality.

The Omnibus Analyses

I used a combination of *t*-tests and ANOVAs to assess the differences between the experimental manipulations on the dependent measures. I used paired samples *t*-tests to compare the outcomes that were confounded by the respective system conditions (essentially, I ran two separate experiments). I performed a separate paired t-test for each

level of contingency, and each level of delay to ensure data independence in the repeated measures design. Thus, each of these analyses considered 36 outcomes associated with each system condition. Using these outcomes, I obtained difference scores by yoking the causal attributions of participants assigned to the same set of contingency and delay combinations in the same presentation order, but from the different system conditions.

Then, analyzing the data separately according to the system condition with the data collapsed over presentation order, I used a series of ANOVAs with a 3 x 3 randomized block partially confounded factorial design (RBPF-3^2) to assess the differences associated with the delay (a. one-trial delay, b. two-trial delay, c. three-trial delay) and contingency (a. 25|0, b. 75|0, c. 75|75) manipulations. If there were no significant interactions, I established whether there were significant delay and contingency main effects, and when applicable, performed pairwise comparisons using Tukey HSD analyses to establish which manipulations statistically differed. Below, I will review the results pertaining to participants' causal attributions and system control, as well as the relationship between these measures. The complete set of analyses, including those that use response time and other strategy-oriented dependent measures are presented in Appendix O.

Causal Attributions. With causal attributions as the outcome measure, there were significant effects reflecting both system and contingency differences, but not differences in delay. Table 7 depicts participants' abstract, ordinal attributions (i.e., whether the treatment raised, lowered or had no effect on patients' blood pressure levels). Figure 15 graphs participants' more detailed, causal attributions on a ratio scale (i.e., the extent to which the treatment affected the patients' blood pressure level) for each of the

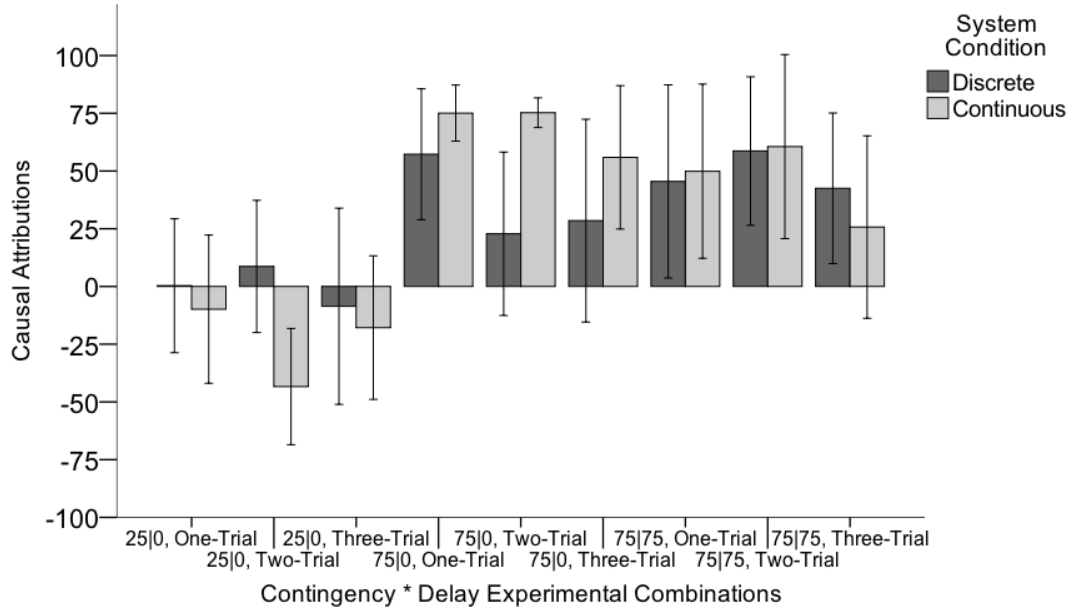
system conditions. Figure 15a separates the data by the contingency and delay treatment combinations, but collapsed over presentation order. Figures 15b and 15c separates the data by levels of contingency and by levels of delay, again collapsing over presentation order. Appendix P contains alternative conversions and transformations of the raw causal attributions, as well as the outcomes of analyses considering the revised data sets. These analyses were conducted to verify that scale and data transformation were not responsible for the findings associated with causal attributions, given the variety of measurement scales used within the causal reasoning literature.

Table 7
Frequency of Abstract, Ordinal Attributions separated by the Experimental Manipulations

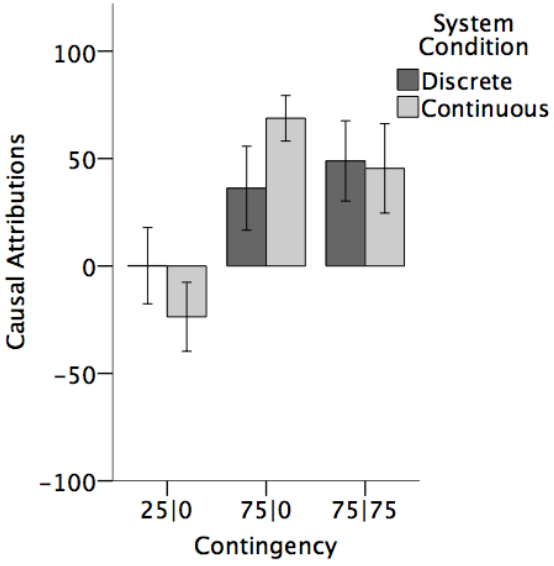
	Raised		No effect		Lowered	
	D	C	D	C	D	C
25|0 and						
One-trial delay	3	3	6	5	3	4
Two-trial delay	2	7	6	4	4	1
Three-trial delay	5	5	2	5	5	2
75|0 and						
One-trial delay	1	0	1	0	10	12
Two-trial delay	2	0	4	0	6	12
Three-trial delay	3	1	1	0	8	11
75|75 and						
One-trial delay	2	1	2	0	8	11
Two-trial delay	1	2	0	0	11	10
Three-trial delay	1	2	3	2	8	8

Note. Raised = ordinal selection that the treatment raised the patient's blood pressure level, No effect = ordinal selection that the treatment had no effect on the patient's blood pressure level, Lowered = ordinal selection that the treatment lowered the patient's blood pressure level. D = Discrete system condition. C = Continuous system condition.

A) Causal Attributions separated by the Contingency and Delay Combinations



B) Causal Attributions separated by Levels of Contingency



C) Causal Attributions separated by Levels of Delay

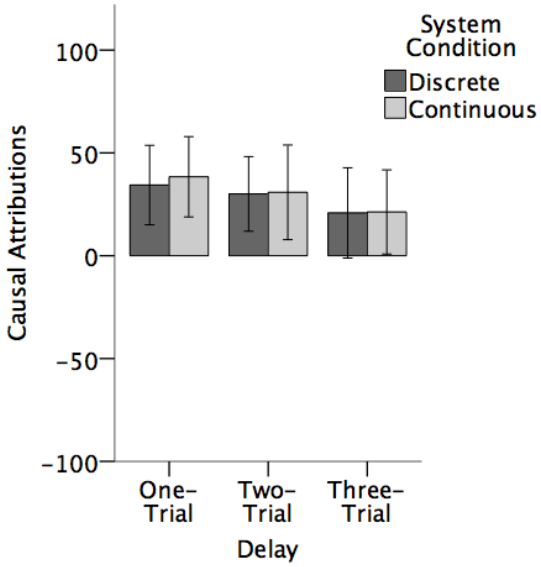


Figure 15. Causal attributions separated by the experimental manipulations, collapsing over presentation order. Causal attributions between 100 and 0 indicated that the treatment lowered the patient's blood pressure, a causal attribution of zero indicated that the treatment had no effect on the patient's blood pressure, and causal attributions between 0 and -100 indicate that the treatment raised the patient's blood pressure levels. Error bars illustrate 95% confidence interval.

Effect of system condition on causal attributions. Causal attributions were lower and higher in the continuous system condition with the 25|0 and 75|0 contingency

manipulations, respectively. The causal attributions were equivalent between the system conditions when the data were separated by delay, as well as with the 75|75 contingency manipulation. Table 8 contains the outcomes of the series of paired samples *t*-tests for each level of contingency, and each level of delay to ensure data independence in the repeated measures design.

Table 8
Paired Samples t-tests - Effect of System on Causal Attributions

	Discrete		Continuous				*95% CI*		
	M	*SD*	*M*	*SD*	*t(35)*	*p*	*LL*	*UL*	*d*
Contingency									
25|0	0.14	52.44	-23.67	47.50	2.27	0.03	2.47	45.14	0.38
75|0	36.19	57.75	68.75	31.33	-3.47	< .01	-51.61	-13.51	-0.58
75|75	48.89	55.18	45.42	61.48	0.24	0.81	-25.50	32.44	0.04
Delay									
One-trial	34.36	57.10	38.39	57.70	-0.36	0.72	-26.80	18.74	-0.06
Two-trial	30.06	53.62	30.83	68.03	-0.06	0.95	-26.77	25.23	-0.01
Three-trial	20.81	64.89	21.28	60.50	-0.04	0.97	-26.04	25.10	-0.01

Note. Discrete = Discrete System Condition. Continuous = Continuous System Condition. CI = Confidence Interval; LL = lower limit; UL = upper limit. I conducted paired samples *t*-tests for separating the data by each level of contingency and each level of delay. Participants from each system condition were yoked by their respective contingency by delay combinations and presentation order. The participant's outcome from the continuous system condition was subtracted from the participant's outcome from the discrete system condition. So, a positive *t*-value suggests causal attributions were higher with the discrete system condition. All tests two-tailed.

Effect of delay and contingency on causal attributions by system. I expected longer delay intervals to interact with the contingency manipulation, such that delay would influence causal attributions more in conjunction with the less probable contingency manipulations (i.e., 75|75 and 25|0). Surprisingly, there was not a significant within block interaction between contingency and delay using causal attributions as the outcome dependent measure, relative to either the discrete ($F(4, 64) = 1.52$, $p = .21$) or continuous ($F(4, 64) = 0.69$, $p = .60$) system conditions.

Effect of contingency on causal attributions by system. In the absence of a significant interaction between the delay and contingency experimental manipulations, I expected causal attributions to reflect the differences in the ΔP values (i.e., $P(e|c) - P(e|\sim c)$) between the contingency manipulations. The findings partially support this hypothesis, with inconsistencies with both system conditions regarding the effect of contingency, particularly regarding the ability to distinguish between the 75|0 and 75|75 contingency manipulations.

With the discrete system condition, there was a significant contingency main effect with causal attributions as the outcome dependent measure ($F(2, 64) = 7.55$, $p < .01$). As predicted, a Tukey HSD pairwise comparison indicated causal attributions were higher with the 75|0 contingency manipulation ($M = 36.19$, $SD = 57.75$) than with the 25|0 contingency manipulation ($M = 0.14$, $SD = 52.44$), $p < .05$. Inconsistent with my predictions were the relationships between the 25|0 and 75|75, and 75|0 and 75|75 contingency manipulations. Causal attributions were higher with the 75|75 contingency manipulation ($M = 48.89$, $SD = 55.18$) than with the 25|0 contingency manipulation, $p < .05$, and there was no significant difference in causal attributions between 75|0 and 75|75 contingency manipulations, $p > .05$.

With the continuous system, the contingency manipulations differed in the same manner relative to causal attributions as with the discrete system condition. Again, there was a significant contingency main effect ($F(2, 64) = 38.72$, $p < .001$). Causal attributions with both the 75|0 ($M = 68.75$, $SD = 31.33$) and 75|75 ($M = 45.42$, $SD = 61.48$) contingency manipulations were higher than with the 25|0 contingency manipulation ($M = -23.67$, $SD = 47.50$), $p < .01$ for both Tukey HSD pairwise

comparisons. The 75|0 and 75|75 contingency manipulations did not differ statistically by causal attribution, $p > .05$.

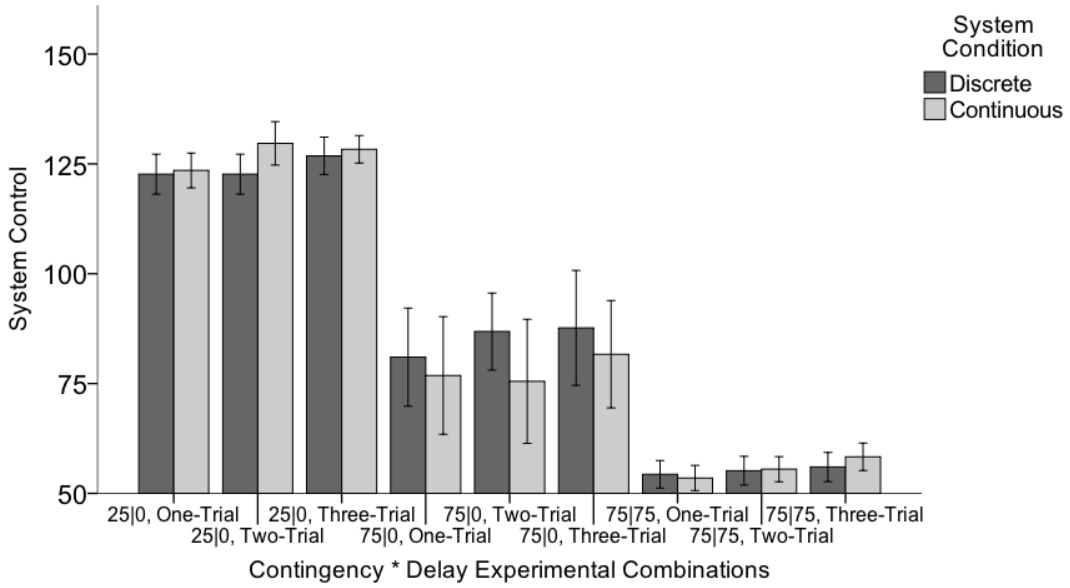
Effect of delay on causal attributions by system. Similar to my prediction involving contingency, I anticipated that delay alone would affect causal attributions. Specifically, I anticipated that causal attributions would decrease as the delay interval increased. The findings do not support this hypothesis.

With the discrete system condition, there was no significant difference between the delay manipulations with causal attributions as the dependent measure ($F(2, 64) = 0.57$, $p = .57$). The mean values of the causal attributions associated with the one-trial, two-trial, and three-trial delay manipulations were 34.36 ($SD = 57.10$), 30.06 ($SD = 53.62$), and 20.81 ($SD = 64.89$), respectively. Similarly for the continuous system, there was no significant difference between the delay manipulations with causal attributions: $F(2, 64) = 1.23$, $p = .30$. The mean values of the causal attributions were 38.39 ($SD = 57.70$), 30.83 ($SD = 68.03$), and 21.28 ($SD = 60.50$) for the one-trial, two-trial, and three-trial delay manipulations.

System Control. The system condition did not affect participants' ability to control the system. Also, there were no significant interactions between contingency and delay, or main effects as a function of the delay with the system control outcome measure. These findings are inconsistent with my expectation that increases in delay would hinder participants' ability to lower blood pressure levels. I did not consider the contingency main effect, as contingency, rather than participants' reasoning behavior, directly determines the outcome of system control (e.g., blood pressure values will decline more with a 75|0 contingency value than with a 25|0 contingency value). Figure

16 depicts the system control outcomes associated with each system condition. Figure 16a separates the data by the contingency by delay treatment combinations, but collapsed over presentation order. Figure 16b separates the data by levels of delay, again collapsing over presentation order. Recall that lower system control values reflect better performance.

A) System Control separated by the Contingency and Delay Combinations



B) System Control separated by Levels of Delay

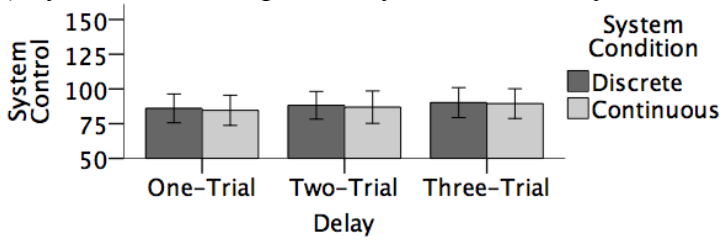


Figure 16. System control separated by the experimental manipulations. Error bars illustrate 95% confidence interval. The system control scale ranged from 50 (lowest possible blood pressure observation) to 150 (highest possible blood pressure observation).

Effect of system condition on system control. Participants' system control was equivalent between the system conditions. Table 9 contains the outcomes of the series of paired samples t-tests used to compare the system conditions. Again, each paired

samples t-test considered a level or either contingency or delay to ensure data independence in the repeated measures design.

Table 9
Paired Samples t-tests - Effect of System on System Control

	Discrete		Continuous				95% CI		
	M	*SD*	*M*	*SD*	*t(35)*	*p*	*LL*	*UL*	*d*
Contingency									
25|0	124.06	7.10	127.17	6.78	-1.99	.06	-6.29	0.07	-.33
75|0	85.17	17.30	78.00	20.45	1.80	.08	-0.90	15.23	.30
75|75	55.17	5.00	55.78	4.95	-0.56	.58	-2.82	1.59	-.09
Delay									
One-trial	86.00	30.57	84.61	32.09	0.61	.55	-3.27	6.05	.10
Two-trial	88.22	29.43	86.89	34.50	0.43	.67	-4.93	7.60	.07
Three-trial	90.17	31.93	89.44	31.66	0.28	.78	-4.36	5.81	.05

Note. Discrete = Discrete System Condition. Continuous = Continuous System Condition. CI = Confidence Interval; LL= lower limit; UL=upper limit. I conducted paired samples *t*-tests for separating the data by each level of contingency and each level of delay. Participants from each system condition were yoked by their respective contingency by delay treatment combinations and presentation order. The participant's outcome from the continuous system condition was subtracted from the participant's outcome from the discrete system condition. So, a positive *t*-value suggests system control outcomes were higher with the discrete system condition. Higher values are indicative of worse performance with the system control measure. All tests two-tailed.

Effect of delay and contingency on system control by system. I expected longer delay intervals to interact with the contingency manipulation, such that delay would influence system control more combination with the 75|0 contingency manipulation given the possible range of system control differences was greater with this contingency manipulation, than the 25|0 or 75|75 contingency manipulations. (Possible differences in system control are presented Appendix Q.) However, there was not a significant within block interaction between contingency and delay, relative to either the discrete ($F(4, 64) = 0.66$, $p = .62$) or continuous ($F(4, 64) = 0.45$, $p = .77$) system conditions, with system control as the dependent measure.

Effect of delay on system control by system. In the absence of a contingency by delay interaction, I anticipated that participants' ability to control the system would decline with increases in delay. The findings do not support this hypothesis. With the discrete system condition, the delay main effect was not significant with system control as the dependent measure ($F(2, 64) = 1.26$, $p = .29$). With the discrete system condition, the mean values of the system control outcomes associated with the one-trial, two-trial, and three-trial delay manipulations were 86.00 ($SD = 30.57$), 88.22 ($SD = 29.43$), and 90.17 ($SD = 31.93$), respectively. Likewise, there was no significant difference between the delay manipulations relative to system control with the continuous system: $F(2, 64) = 1.33$, $p = .27$. With the continuous system condition, the mean system control values were 84.61 ($SD = 32.09$), 86.89 ($SD = 34.50$), and 89.44 ($SD = 31.66$) for the one-trial, two-trial, and three-trial delay manipulations, respectively.

Correlations between causal attribution and system control. I performed a series of Pearson product-moment correlational analyses to assess the relationship between system control and causal attribution.[4] The data associated with these analyses are plotted in Figure 17 and the outcomes are shown in Table 10. I separated the data by system and contingency collapsing across levels of delay, and by system and delay collapsing across levels of contingency to maintain data point independence. (See Appendix S for the analysis that separates the data by system, contingency, and delay.)

[4] Appendix R contains the comprehensive set of results. Appendix S focuses specifically on the measures used to define strategy relative to causal attributions and system control. Appendix T contains a series of regression analyses assessing the influence of system control on causal attribution that control for delay.

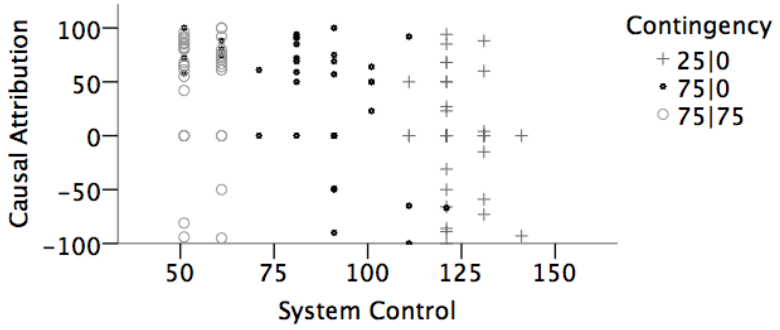
With the 75|0 contingency manipulation in the discrete system condition, better system control (i.e., the more declines in blood pressure observations) correlated with the higher causal attributions. This relationship was not observed in combination with the other contingency manipulations within the discrete system or with the continuous system condition when outcomes were separated by contingency. However, higher causal attributions were associated with greater declines in blood pressure observations across all of the delay manipulations with both system conditions.

Table 10
Correlations between the Causal Attributions and System Control

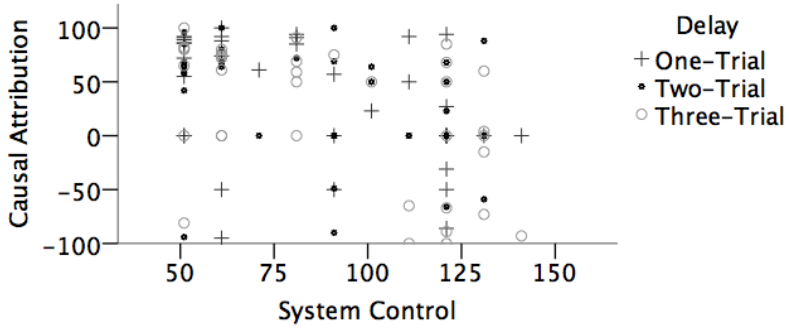
	Discrete System			Continuous System		
	r	r^2	*p*	*r*	r^2	*p*
Contingency						
25|0	- .178	.03	.30	- .296	.09	.08
75|0	- .509	.26	< .01	- .109	.01	.53
75|75	- .023	< .01	.89	- .157	.02	.36
Delay						
One-Trial	- .412	.17	.01	- .470	.22	< .01
Two-Trial	- .359	.13	.03	- .675	.46	< .01
Three-Trial	- .497	.25	< .01	- .412	.17	.01

Note. All tests two tailed.

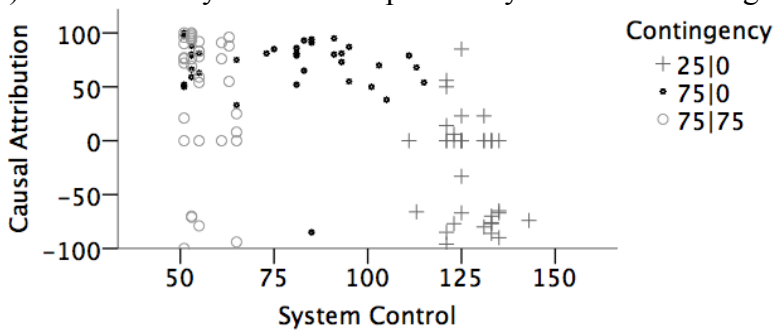
A) Discrete System Data Separated by Levels of Contingency



B) Discrete System Data Separated by Levels of Delay



C) Continuous System Data Separated by Levels of Contingency



D) Continuous System Data Separated by Levels of Delay

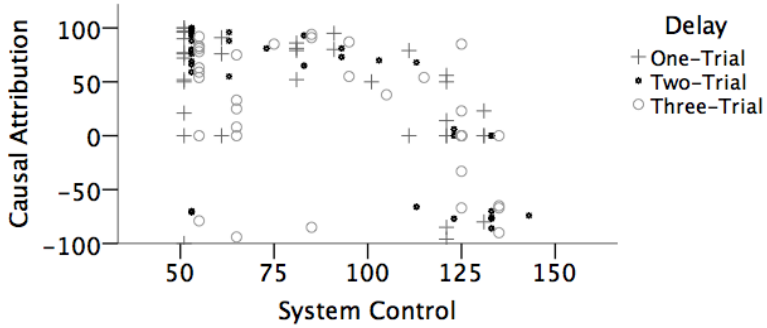


Figure 17. Relationships between the causal attribution and system control dependent measures. System condition, and either levels of delay or levels of contingency separate the data.

Strategic Differences

I conducted a strategy analysis to (better) identify and categorize patterns of treatment intervention behavior. I sorted participants' treatment intervention behavior on two dimensions of strategy: intervention bias and observation dependence. There was no evidence to suggest that participants changed their behavior in the midst of individual patient scenarios (Appendix U) or between patient scenarios (Appendix V), and participants' self-reported behavior corroborates these findings (see Appendix W). However, there were significant system effects on intervention strategy, specifically with the strategic dimension of observation dependence (Figure 19). Additionally, I considered the effect of intervention strategy on causal attribution. In general, increases in causal attributions were observed following treatment-biased and observation-independent strategies (Figure 20).

Strategy categorizations. I identified the treatment intervention strategies of intervention bias and observation dependence using two dependent measures: the total number of treatment interventions applied and the number of context specific runs. I performed analyses considering the strategic dimensions of intervention bias and observation dependence separately, as the classification schemes are not orthogonal, meaning it is not possible to distinguish between or relate the findings between these separate analyses. However, the dependent measures used to form these categorizations do appear to be negatively related with both system conditions[5], such that applying more

[5] The findings from correlational analyses from discrete ($r = -.58$, $p < .01$) and continuous ($r = -.50$, $p < .01$) system conditions indicate that the strategic dimensions are negatively correlated. However, it is important to note that these analyses violate the assumption of

treatment interventions lessens the degree with which participants switch between applying and withholding the treatment (i.e., number of runs).

Intervention bias. To establish whether participants were biased to a particular treatment intervention, I separated the intervention selection data into categories using the total number of treatment interventions applied. I calculated the 95% confidence interval for the unbiased strategy. Using the upper and lower confidence interval limits, I divided participants into three categories: *treatment-biased* (participants that applied the treatment on 12 or more trials), *unbiased* (participants that applied the treatment across 5 to 11 trials), and *withhold-biased* (participants that applied the treatment on four or less trials). Participants rarely had a bias toward withholding the treatment with both the discrete (n = 4) and continuous (n = 3) system conditions, so data within this categorization were excluded from analyses reviewing differences in intervention bias. The distribution of unbiased (n = 114) and treatment-biased (n = 95) intervention strategies was split across the experimental manipulations. However, unbiased intervention strategies in which *P(treatment) = P(~treatment*) = 8 (e.g., the intervention strategy typically used by causal reasoning researchers when controlling the learning experience) were only applied in 17 of 216 learning tasks.

Observation dependence. Next, I determined whether participants formed their treatment intervention strategies independent of or dependent on their patients' blood pressure observation shown on each trial. I identified *observation-dependent* and

---

independence as they collapsed data across the experimental manipulation combinations, so each individual contributed three data points for each measure. Appendix R contains correlations between all of the dependent measures, including those mentioned here.

observation-independent strategies using the number of runs measure. The number of runs describes the number of times participants switched between applying and withholding the treatment. I scored behavior according to consistency with the observed state (i.e., whether the treatment selection was applied after observing an increase in blood pressure and withheld after an observing a decrease in blood pressure). I determined the number of expected runs associated with an observation-dependent treatment intervention strategy independently for each combination of experimental manipulations (e.g., 75|0 contingency with one-trial delay in the discrete condition). Using these values as the expected value input, I conducted a series of runs tests[6] to establish if the number of observed runs significantly varied from the number of runs associated with an observation-dependent intervention strategy. Using the resulting *z*-statistics, I characterized participants' treatment intervention behavior as either observation-dependent ($p >= .05$) or observation-independent ($p < .05$). A total of 119 observation-independent and 97 observation-dependent intervention strategies were observed.

Effect of the system condition on strategy. I used a series of Fisher's Exact Tests to assess whether participants' use of treatment intervention strategies was contingent on the system condition. I separated the data by levels of contingency or levels of delay to ensure data independence. The results of these analyses indicate that

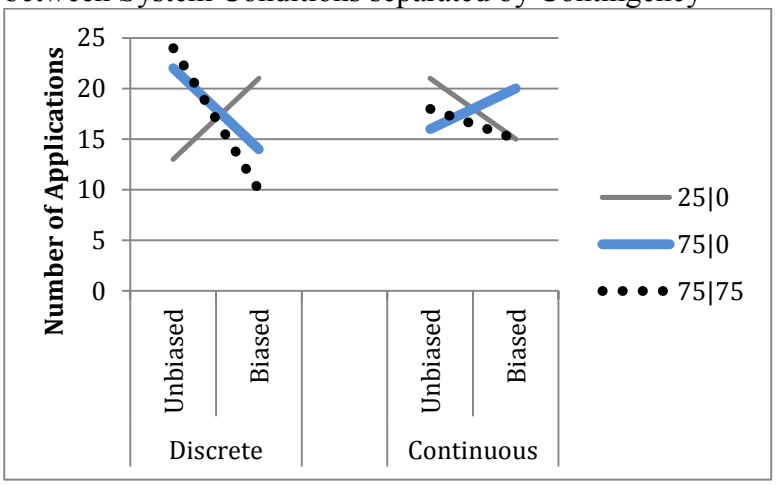
[6] This series runs tests differs from the traditional Wald-Wolfowitz runs test for randomness. Although the statistical formula and procedure is consistent, I compared the observed number of runs to the expected number assuming a particular intervention strategy.

the system condition generally did not affect participants' intervention bias, but did affect their observation dependence.

Effect of system condition on intervention bias. The distribution of intervention bias was consistent between the discrete and continuous system condition. The only difference in intervention bias was observed with the discrete system condition; participants applied more unbiased intervention strategies with the 75|75 contingency manipulation in comparison to the 25|0 contingency manipulation ($p < .01$). The same pattern of behavior is visually apparent between the 75|0 and 25|0 contingency manipulations, but the difference was not significant ($p = .054$).

Figure 18 depicts the number of applications of unbiased and treatment-biased intervention strategies applied in each system condition separated by A) levels of contingency, and B) levels of delay. Table 11 contains the distribution of intervention strategies separated by intervention bias, alongside the Fisher's Exact Tests *p* values used to assess the differences in the distribution of intervention bias between the system conditions. Appendix X contains the comprehensive set of analyses that includes analyses considering the differences in the distribution of intervention bias according to either contingency or delay, confounded by the system condition.

A) Differences in the Distribution of Intervention Bias between System Conditions separated by Contingency



B) Differences in the Distribution of Intervention Bias between System Conditions separated by Delay

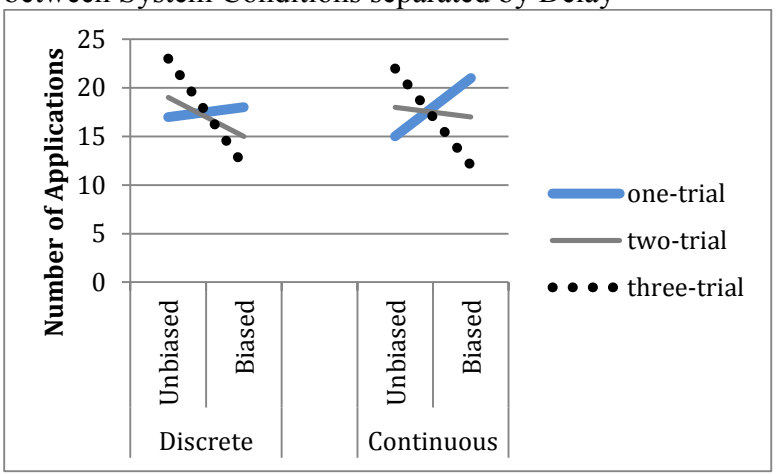


Figure 18. Differences in the Distribution of Intervention Bias between System Conditions. A) Data separated by contingency. B) Data separated by delay. Asterisks identify significant differences. Unbiased = Unbiased Treatment Intervention Strategy. Biased = Treatment-biased Intervention Strategy. Discrete = Discrete System Condition. Continuous = Continuous System Condition

Table 11
Distribution of the Strategic Dimension of Intervention Bias and Outcomes of Fisher's Exact Tests assessing differences between the System Conditions

	Discrete System		Continuous System		
	Unbiased	Biased	Unbiased	Biased	*p*
Contingency					
25|0	13	21	21	15	.10
75|0	22	14	16	20	.23
75|75	24	10	18	15	.21
Delay					
One-trial	17	18	15	21	.64
Two-trial	19	15	18	17	.81
Three-trial	23	12	22	12	1.00

Note. Biased = Treatment-biased intervention strategy. All tests two-tailed.

Effect of system condition on observation dependence. Unlike the strategic dimension of intervention bias, there was a system effect on the strategic dimension of observation dependence. The system condition interacted with both contingency (75|0 and 75|75) and delay (two-trial and three-trial), with more participants applying observation-dependent strategies with the discrete system condition than with the continuous system condition. Yet, the distribution of observation dependence was independent from system condition with the 25|0 contingency and one-trial delay manipulations.

Within the continuous system condition, observation-independent strategies were used more in combination with the 75|0 ($p < .01$) and 75|75 ($p < .01$) contingency manipulations relative to the 25|0 contingency manipulation. This explains why there was no between system effect on observation dependence pertaining to the 25|0 contingency manipulation. Yet, there were no significant differences in the distribution of observation dependence with respect to delay within either system condition to explain why the between system interaction was only observed with longer intervals of delay.

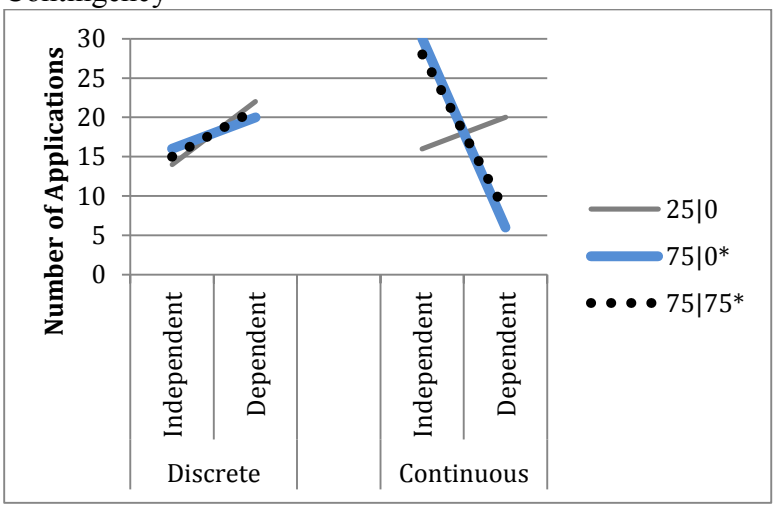
Table 12 contains the distribution of intervention strategies separated by observation dependence, alongside the Fisher's Exact Tests *p* values that illustrate the use of more observation-dependent strategies with the discrete system condition. Figure 19 depicts the number of applications observation-independent and observation-dependent strategies applied with each system condition separated by levels of A) contingency, and B) delay. Again, Appendix X contains the comprehensive set of analyses that consider the differences in the distribution of observation dependence as a function of either levels of contingency or levels of delay, confounded by each system condition.

Table 12
Distribution of the Strategic Dimension of Observation Dependence and Outcomes of Fisher's Exact Tests assessing differences in the Distribution of Observation Dependence between the System Conditions

	Discrete System		Continuous System		
	Independent	Dependent	Independent	Dependent	*p*
Contingency					
25|0	14	22	16	20	.81
75|0	16	20	30	6	< .01
75|75	15	21	28	8	< .01
Delay					
One-trial	19	17	25	11	.23
Two-trial	11	25	23	13	< .01
Three-trial	15	21	26	10	.02

Note. Independent = Observation-independent intervention strategy; Dependent = Observation-dependent intervention strategy. All tests two-tailed.

A) Differences in the Distribution of Observation Dependence between System Conditions separated by Contingency



B) Differences in the Distribution of Observation Dependence between System Conditions separated by Delay

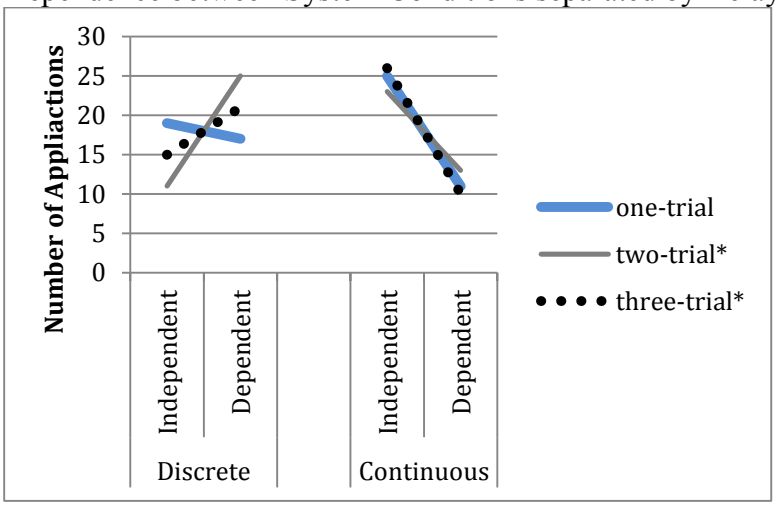


Figure 19. Differences in the Distribution of Observation Dependence between System Conditions. A) Data separated by levels of contingency. B) Data separated by levels of delay. Asterisks identify significant differences. Independent = Observation-Independent Intervention Strategy. Dependent = Observation-Dependent Intervention Strategy. Discrete = Discrete System Condition. Continuous = Continuous System Condition

Effect of intervention strategy on causal attributions. Next, I assessed whether intervention strategy was related to participants' causal attributions.[7] I performed a series of 2 (intervention strategy) by 2 (system condition) ANOVAs using the causal attribution as the dependent measure. I performed separate ANOVAs for the two strategic dimensions (intervention bias, observation dependence) as well as for each level of contingency (25|0, 75|0, 75|75)[8] collapsing over the delay manipulations to maintain subject/data point independence. Below, I focus on the simple effects with respect to intervention strategy and contingency as that only one of twelve interactions between the system conditions and the strategic dimensions was significant (see Appendix AA). Also, the results pertaining to the system condition main effects concur with the paired samples *t*-test findings discussed within the omnibus analysis section (and included in Appendix O), so will not be repeated here.

---

[7] Differences in system control are inherent with the contingency manipulation (Appendix C). Although I did consider the differences in system control as a function of contingency and delay, and intervention strategy, the results reflect the distribution of contingency values underlying each intervention strategy, rather than the effect of intervention strategy (Appendix Y).

[8] Appendix Z contains a similar set of analyses with the data separated by delay. These analyses are not presented in the document because the distribution of contingency values was not equivalent between the intervention strategy conditions, so the effect of contingency on causal attributions realized with the omnibus ANOVAs confounds these results. The data is separable by contingency because delay did not affect causal attributions.

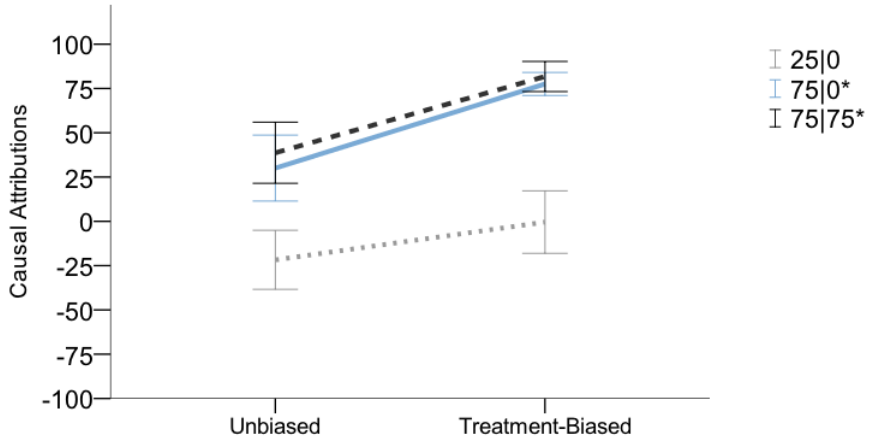
In general, causal attributions appear higher with treatment-biased and observation-independent strategies. However, this tendency was only significant with a subset of the contingency manipulations. Relative to intervention bias (Figure 20a), higher causal attributions followed treatment-biased intervention strategies as compared to unbiased intervention strategies with the 75|0 and 75|75 contingency manipulations. Yet, this same relationship was not observed with the 25|0 contingency manipulation. Similarly, observation-independent intervention strategies were associated with higher causal attributions, but only with the 25|0 and 75|0 contingency manipulations (Figure 20b). Table 13 presents the simple effect findings. Figure 20 depicts the relationships between causal attributions and intervention strategies separated by levels of contingency, collapsing over system and delay.

Table 13
Effect of Strategy by Contingency Level on Causal Attributions (simple effects, between subjects)

	Unbiased			Treatment-biased						
Contingency	*n*	*M*	*SD*	*n*	*M*	*SD*	*F*	*df*	*p*	η_p^2
25|0	34	-21.74	47.77	36	-0.42	52.20	1.93	1, 67	.17	.03
75|0	38	30.05	56.55	34	77.53	18.57	19.00	1, 69	< .01	.22
75|75	42	38.69	55.40	25	81.76	20.66	13.59	1, 64	< .01	.18
	Observation-independent			Observation-dependent						
Contingency	*n*	*M*	*SD*	*n*	*M*	*SD*	*F*	*df*	*p*	η_p^2
25|0	30	3.73	45.80	42	-22.83	52.32	5.85	1, 69	.02	.08
75|0	46	70.52	30.99	26	20.54	58.45	14.43	1, 69	< .01	.17
75|75	43	50.23	59.62	29	42.59	56.29	0.47	1, 69	.50	.01

Note. η_p^2 = partial eta squared.

A) Intervention Bias



B) Observation Dependence

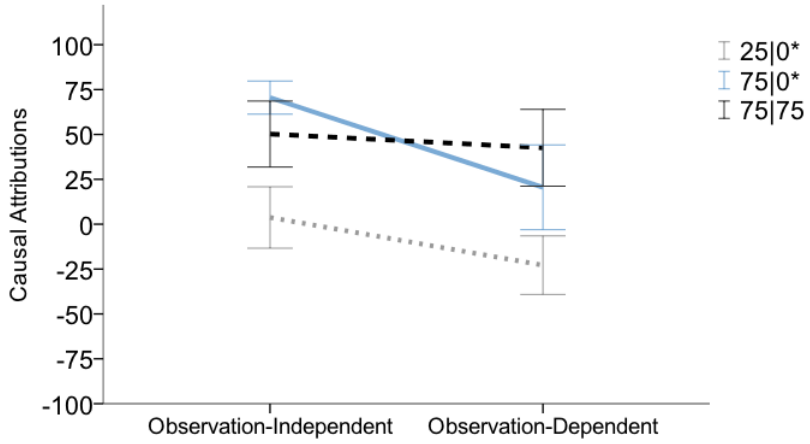


Figure 20. Relationships between participants' causal attributions and intervention strategy separated by levels of contingency collapsing over system condition and delay. Error bars illustrate 95% confidence interval. Asterisks identify significant differences. A) Relationship between causal attributions and the strategic dimension of intervention bias. B) Relationship between causal attributions and the strategic dimension of observation dependence.

Comparison to Previous Experimental Research

I compared the present data to existing experimental research using slopes rather than the actual values associated with causal attributions to account for differences in the experimental manipulations (e.g., 1 second to 3 second delay vs. 4 to 8 second delay) and procedural approach (e.g., open/free-operant vs. directed learning environment). These

slopes describe the change in causal attribution given the change in experimental manipulation. I generated the slopes using the mean causal attribution value, rather than individual data-points, because raw data sets were unavailable for the referenced experimental research. Also, when generating the mean values for the present study, I excluded portions of the complete data set to maximize similarity between the rating scales (i.e., outcome measure) and learning environments to justify a comparison. Specifically, I only considered data from the discrete system condition and data associated with *no effect* and *lowered* blood pressure assessments.

Using these slopes, I performed independent samples *t*-tests to assess whether the relationships observed with past experimental research were consistent with the relationships observed in the present study; these *t*-tests considered the respective sample sizes, slopes, and standard error values. My findings partially replicate findings from previous experimental research. Slopes containing causal attributions associated with the 75|75 contingency manipulation significantly differed between the data sets. However, the remaining contingency relationship (25|0 to 75|75), and the relationships between the delay manipulations were consistent with the empirical literature.

Comparison to existing experimental research with contingency. I expected the relationships between contingency manipulations to be equivalent to the changes in *P(e|c)* and *P(e|~c)* found within Buehner et al. (2003, experiment 1). I compared changes in contingency relative to changes in causal attributions, rather than using the mean of the causal attribution values associated with each contingency value, to account for experimental paradigm and manipulation differences. Initially, I only included data pertaining to the one-trial delay manipulation in the discrete system condition because

this condition is the most similar to Buehner et al.'s experimental paradigm, which did not consider the influence of delay. However, I the changes in causal attributions given contingency across the delay manipulations as well, as both the contingency by delay interaction, and the delay main effect were not significant, but again, only with data from the discrete system condition.

Table 14 contains the results of the independent samples *t*-tests comparing the slopes associated Buehner et al. to the slopes generated from the mean causal attribution values (i.e., 25|0 relative to 75|0 and 75|0 relative to 75|75). The change in participants' causal attributions between the 25|0 and 75|0 contingency manipulations replicates the positive linear relationship observed by Buehner et al. When only the one-trial delay data are included in the analysis, the slope between the causal attributions for the 25|0 and 75|0 contingency manipulations exceeds the magnitude observed by Buehner et al. However, the slopes are equivalent when causal attributions from all of the delay manipulations are included in the analysis. The slope of the differences in causal attributions between the 75|0 and 75|75 contingency manipulations is positive as well. This observation directly contradicts Buehner et al.'s findings (again that uses a discrete learning paradigm with no delay) that suggest causal attributions should decrease as the ΔP *(i.e., P(e|c) -P(e|~c))* decreases. A visual comparison of slopes from Buehner et al. (also shown in Figure 2) and the observed data are depicted in Figures 21 and 22.

Table 14
Contingency Comparison to Buehner et al.'s (2003, experiment 1) Findings

	n	*SE*	*β*	*t*	*df*	*p*	*d*
25|0 to 75|0							
One-trial delay	20	7.10	116.54	5.41	276	< .01	0.65
All delay manipulations	56	4.51	80.84	0.73	312	.47	0.08
75|0 to 75|75							
One-trial delay	21	7.51	2.57	4.80	266	< .01	0.59
All delay manipulations	62	4.13	9.08	9.72	225	< .01	1.30

Note. I excluded negative causal attributions associated with 'raised blood pressure' assessments from these analyses. All tests two-tailed.

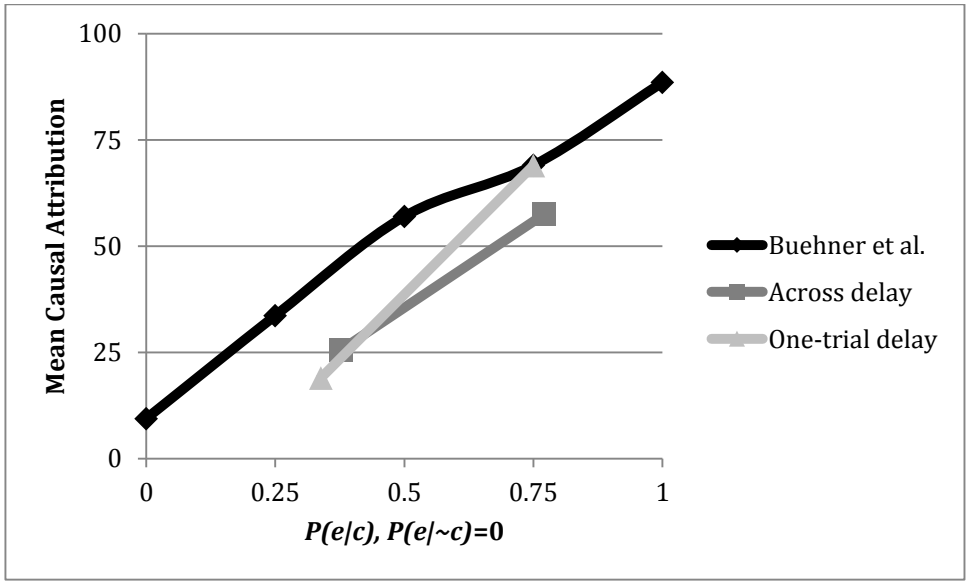


Figure 21. Relationships between causal attributions associated with 25|0 and 75|0 contingency manipulations. Data from Buehner et al. (2003, experiment 1) and a subset of the data from the present study (across delay manipulations in the discrete system condition and only the one-trial delay in the discrete system condition).

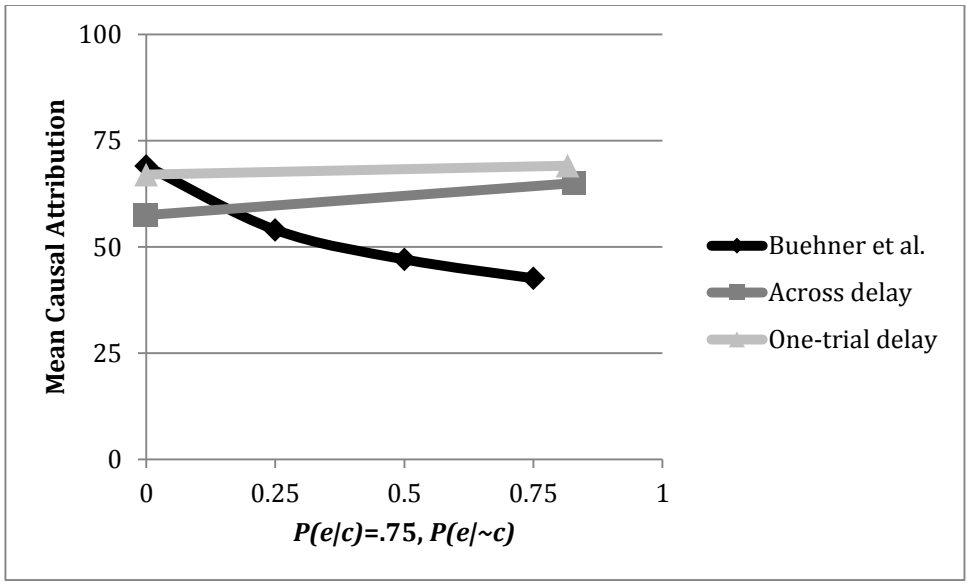


Figure 22. Relationship between causal attributions associated with 75|0 and 75|75 contingency manipulations. Data from Buehner et al. (2003, experiment 1) and a subset of the data from the present study (across delay manipulations in the discrete system condition and only the one-trial delay in the discrete system condition).

Comparison to existing experimental research with delay. I expected the observed findings to correspond to findings from existing experimental research illustrating decreases in causal attribution with increases delay intervals. I compared causal attributions from the 75|0 contingency manipulation in the discrete system condition with findings from Greville and Buehner (2010, experiment 1 and 2) as this data segmentation was the most equivalent. The implemented delay intervals (i.e., $M_{\text{one-trial delay}}$ = 3.92 seconds, $M_{\text{two-trial delay}}$ = 7.95 seconds, $M_{\text{three-trial delay}}$ = 11.42 seconds)[9] differed from those used by Greville and Buehner (i.e., 2 seconds and 4 seconds, 3 and 6

[9] These times were calculated based on the data associated with the 75|0 contingency manipulations in the discrete system condition. The average trial time in the learning task across the contingency manipulations was 3.90, 7.80, 11.70 seconds for the one-trial, two-trial, and three-trial delay manipulations respectively.

seconds for experiments 1 and 2, respectively). Given this discrepancy, I compared changes in causal attribution relative to changes in delay rather than comparing the mean causal attribution value for specific intervals.

I excluded the data associated with the three-trial delay manipulation because Greville and Buehner's experiments only considered two shorter duration time intervals. Table 15 lists the number of participants, standard errors, and slopes associated with Greville and Buehner's findings and the results of the independent samples *t*-tests comparing the delay decrements with the 75|0 contingency by discrete system condition. Figure 23 depicts the negative relationship between causal attributions with the one-trial and two-trial delay manipulations (β = -6.38, *SE* = 7.46) alongside Greville and Buehner's findings. The absence of statistical differences confirms that the observed reductions in causal attributions were consistent with Greville and Buehner's various experimental conditions.

Table 15
One-trial and Two-trial Delay Comparison to Greville and Buehner's (2010) Findings

	n	*SE*	β	*t*	*df*	*p*	*d*
Experiment 1, fixed delay	30	3.26	-12.72	0.78	77	.44	0.18
Experiment 1, distributed delay	30	2.94	-7.15	0.10	77	.92	0.02
Experiment 2, fixed delay	60	2.85	-2.57	0.48	137	.63	0.08
Experiment 2, distributed delay	60	2.49	-3.61	0.35	137	.73	0.06

Note. Comparison to the causal attributions collected with the 75|0 contingency manipulation with the discrete system condition. All tests two-tailed.

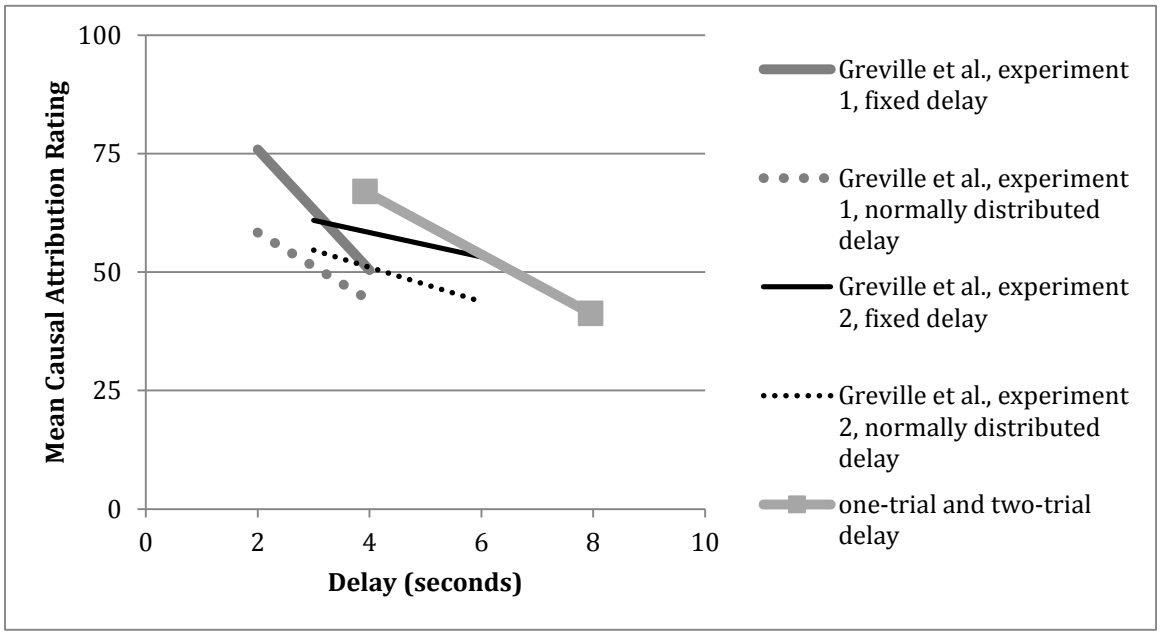


Figure 23. Relationship between causal attributions associated with the one-trial and two-trial delay manipulations to the findings associated with Greville and Buehner's (2010) various experimental conditions.

In addition to these analyses, I reviewed the slope of the causal attributions associated with all three of the delay manipulations implemented in this study. The mean causal attribution value associated with the three-trial delay manipulation was greater than the two-trial delay manipulation creating a curvilinear relationship (shown in Figure 24). This curvilinear relationship suggests that different causal attributions may emerge when longer delay intervals are considered. However, it is not appropriate to suggest or compare the complete data set with Greville and Buehner's outcomes (2010, experiment 1 and 2), as these experiments did not assess the influence of longer delay intervals.

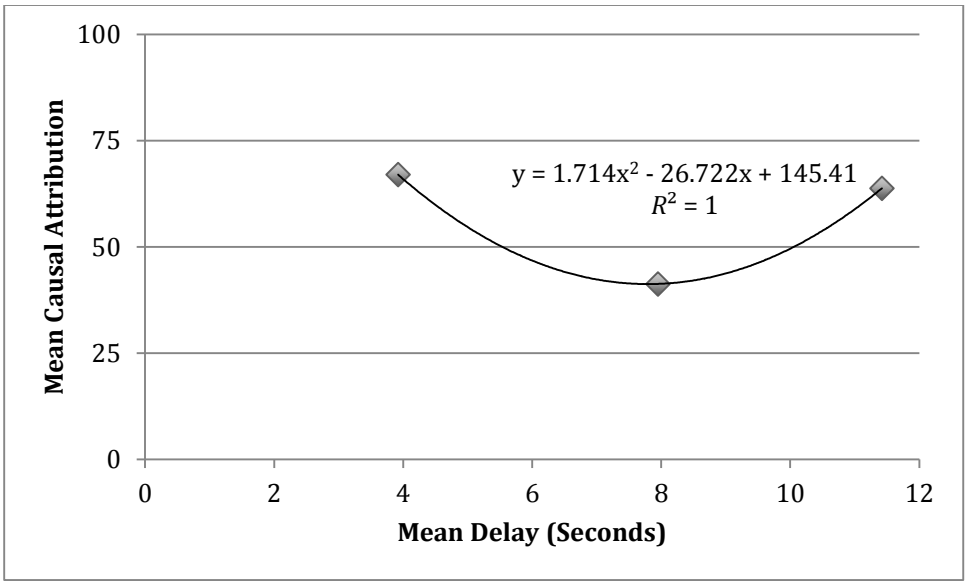


Figure 24. Relationship between causal attributions associated with the one-trial, two-trial, and three-trial delay manipulations.

Comparison to Reasoning Theory

I used one-sample *t*-tests to compare the observed causal attributions to theoretical predictions separated by treatment intervention strategy (i.e., treatment bias, observation dependence) and contingency. Although the observed main effect trends are generally consistent with existing experimental research, a single reasoning theory does not successfully predict the observed causal attributions across the contingency manipulations for either system condition (see Appendix AB for the comprehensive set of analyses). This finding is consistent with the observed differences in causal attributions (contingency effect, system effect), the effect of reasoning strategy on causal attributions, and the differences in participants' self-reported assessment technique in the judgment task (discussed in the next section).

To summarize, the theoretical predictions accounted for more causal attributions associated with the 25|0 contingency manipulation in the discrete system condition (which were less extreme) and an equivalent amount of causal attributions with the 75|0

contingency manipulation (despite the more extreme causal attributions with the continuous system condition). Also, the pattern of findings was similar between the discrete and continuous systems with the 75|75 contingency manipulation. Overall, the theoretical predictions that left out probabilistic information, including base rates, were more successful at predicting the observed causal attributions.

Debrief Questionnaire Insights

In addition to collecting demographic data, the debrief questionnaire attempted to elicit participants' expectations of and adherence to elements of the cover story including task goals, hypertension treatment options, and treatment effectiveness. Questions concerning the experiment, as opposed to general expectations (e.g., treatment effectiveness) directed participants to reflect and respond based on their third patient scenario only. This feedback indicated that more participants prioritized discovering the treatment's efficacy over lowering their patients' blood pressure. Participants' expectations concerning the three hypertensive treatment cover story variations (medicine, diet, and exercise) were consistent and their threshold for treatment efficacy was well above the 25|0 contingency manipulation. Also, participants' descriptions of their assessment technique in the judgment task confirm their use of different procedures in forming attributions in the judgment task.

For qualitative questions, I coded and categorized responses, and then used chi-square tests to establish categorical differences. For quantitative questions, I used independent samples *t*-tests to assess differences. When participants' provided a range of values, rather than a specific value as their response, I converted their response to the

median value of the range. Appendix AE contains analyses pertaining to questionnaire material not covered and referenced in this section or elsewhere in the document.

Use of task goals. To understand participants' adherence to and prioritization of task goals, I asked participants to explain why they applied their specific intervention strategy (that they used with their third patient scenario). In most instances, these descriptions mentioned one of the task goals: lowering the patient's blood pressure (n = 24) and realizing whether the treatment was an effective solution for the patient (n = 43). However, some participants did mention alternative motivators, including their beliefs about the treatment (n = 11) and their experiences with the previous patient scenarios (n = 2). Also, a handful of participants provided a non-interpretable response (n = 5). Figure 25 summarizes participants' use of the task goals with the data separated by system condition.

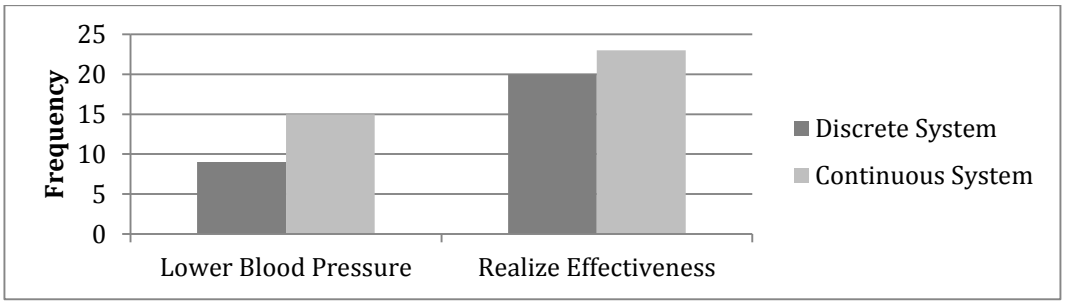


Figure 25. Use of task goals. Only considers behavior with the third patient scenario.

A Fisher's Exact Test indicates that participants' adherence to either of the task goals was unrelated to their system condition (p = .40). With the data collapsed over system conditions, more participants' claimed to focus on realizing the effectiveness of the treatment than lowering the patient's blood pressure level (X^2 (1, n = 55) = 6.56, p = .01, *Cramer's V* = .35). Participants that mentioned using both task goals (n = 6) were excluded from these analyses to maintain data independence.

Equivalence between cover stories. Given the within subjects experimental design, I used different cover stories (i.e., name of patient, type of hypertensive treatment) to help participants distinguish between their three patient scenarios. I asked participants to discuss their expectations of the hypertension treatment that they administered in their third patient scenario to confirm whether participants' bias toward the hypertension treatment options (i.e., medication, diet, and exercise) were comparable. The number of responses concerning each of the hypertension treatments differed as the treatment options were randomly assigned without replacement across the three patient scenarios. So, 21, 26, and 25 participants provided their expectations concerning the implications of a healthy diet, exercise, and medication, respectively.

Most participants expected that these treatments to lower blood pressure levels (55 of 57 participants). The results of chi-square tests indicate that this expectation was evenly distributed across the three hypertension treatment options (X^2 (2, n = 52) = 0.62, p = .74, *Cramer's V* = .08) and that this expectation was true regardless of system conditions (X^2 (2, n = 52) = 3.16, p = .21, *Cramer's V* = .17).

A small number of participants did convey *alternative* biases. For example, there was mention of alternative directional influences on blood pressure levels (i.e., stabilize (n = 1), raise (n = 2), and have no effect (n = 2)). Also, a small number of participants indicated that their treatment expectations were based on their experiences with previous patient scenarios (n = 3) or that actual clinical research had proven a treatment's ability to reduce blood pressure (n = 2). Although I did want to mention these alternative expectations, it is unlikely that they influenced my findings given the relatively small number of participants in this category and the diversity of their expectations. The

remaining ten participants did not clearly communicate their expectations or stated that they were unsure of their expectation (i.e., unknown). Figure 26 summarizes participants' expectations relative to each of the hypertension treatment options.

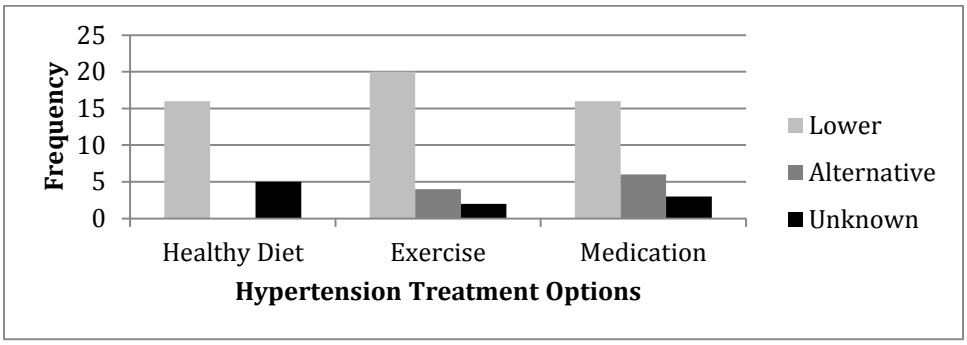


Figure 26. Treatment expectation by hypertension treatment option. N = 72; treatments were randomly assigned.

Treatment effectiveness. I asked participants to quantify the percentage of treatment applications that should influence the patient's condition for that treatment to be considered effective to better understand their prior biases toward treatment effectiveness, as well as to contextualize their responses in the judgment task. Most participants felt that treatment applications should produce the desired outcome at least 50% of the time for the treatment considered effective (X^2 (1, n = 72) = 29.39, p < .01, *Cramer's V* = .64). Participants' expectations for treatment efficacy were consistent between the discrete and continuous system condition according to an independent samples t-test: $t(70) = -0.16$, $p = .87$, 95% *CI* [-0.10, 0.09], $d = -.04$. Figure 27 depicts the skewed distribution of expected treatment efficacy separated by system condition.

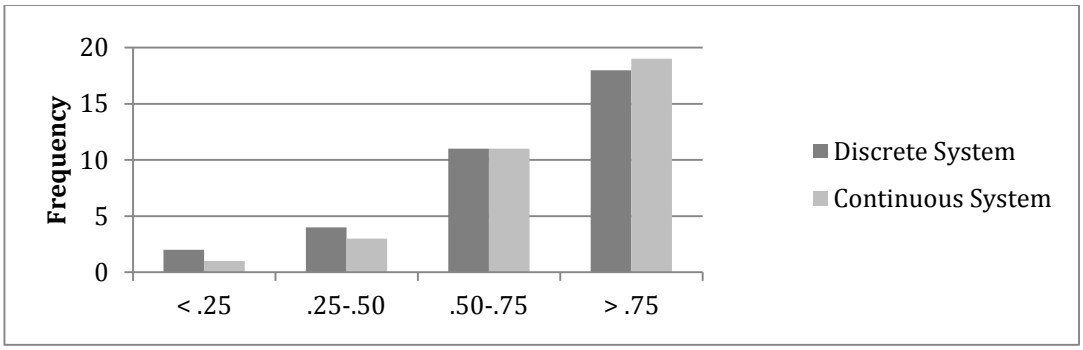


Figure 27. Expectation for treatment efficacy. Response distribution of the percentage of treatment applications that should influence a patient's condition for that treatment to be considered effective.

Self-reported assessment technique in the judgment task. Participants' self-reported responses emphasize that participants employed different approaches when completing the judgment task. Their responses referenced relying on specific observations (e.g., the two times I clicked, when I applied…, each day), using a tally of successful applications (e.g., most days, not all of the time but most of the time, constantly, how many times), and consideration of the overall influence of the treatment (e.g., before and after, after a while, stayed within the same area). Additionally, some participants mentioned relying on their intuition or general beliefs when considering the efficacy of the treatment (e.g., what I thought, what I believed was correct, I guessed).

Figure 28 shows the distribution of responses separated by system condition. More participants stated that they used information from specific trials when assessing a treatment's causality with the continuous than with the discrete system (X^2 (1, n = 12) = 5.33, p = .02, *Cramer's V* = .67). This finding contradicts the observation that participants in the discrete system condition were more reliant on specific system states when forming their intervention strategy. None of the other assessment techniques significantly varied between the system conditions (Appendix AF). Also, more

participants reported using explicit (specific, tally, and overall) over implicit (intuition and unknown) assessment techniques (X^2 (1, n = 72) = 4.50, p = .03, *Cramer's V* = .25).

I performed a series of ANOVAs to assess whether the differences in the explicitness affected causal attributions. I categorized the specific, tally, and overall assessment techniques as explicit, and intuitions and unknown as implicit assessment techniques. There were no significant interactions between explicitness and intervention strategy on causal attribution. Also, there were no significant interactions between explicitness and level of contingency, and explicitness and levels of delay, with causal attributions as the dependent measure. (See Appendix AF for the comprehensive analyses.)

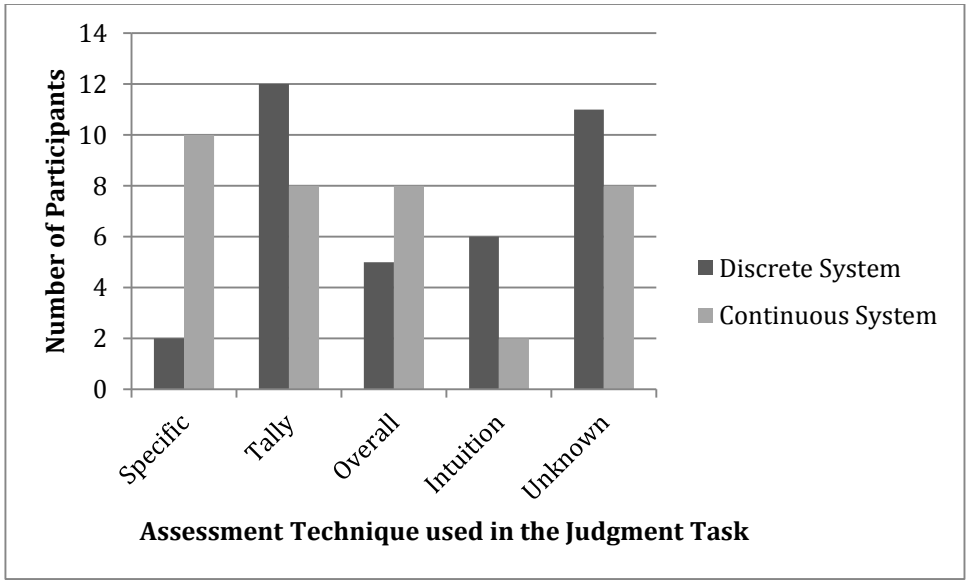


Figure 28. Approach to judgment task. N = 72. Distribution of participants self-reported assessment techniques

IV. DISCUSSION

My hypotheses and research questions involved the relationships between contingency, delay, and system manipulations, as well as the role of intervention strategy in a paradigm that permits the joint examination of casual understanding and system control. The results indicate that causal attributions differed according to both dimensions of intervention strategy. The relationships between these outcome measures present an opportunity to unify, and improve the causal reasoning and system control experimental paradigms. Neither reasoning behavior nor understanding appears to successfully measure or present a cohesive understanding of the reasoning process, in isolation.

Concluding that performance was better with one system condition over the other is speculative. However, an increased reliance on blood pressure outcomes with the discrete system condition highlights a difference in information accessibility between the system conditions. Differences in reasoning behavior and outcomes, as well as the feasibility of studying time with truly discrete system conditions question the utility of research paradigms that employ these discrete systems, as the extensibility of the findings, although valid, is limited.

There was limited support for the hypotheses concerning the contingency and delay manipulations. The unexpectedly high causal attributions associated with the 75|75 contingency manipulation differed from those previously observed in experimental research, but the consideration of participants' expectations and observed outcomes

rationalize the observed outcomes.

Finally, the absence of findings associated with the delay manipulation as measured by system control and causal attribution suggests a potential issue with the length of the learning task, as well as the use of trials as a representation of time.

Intervention Bias

The free operant learning paradigm enabled the consideration strategic differences and their affect on reasoning outcomes. Researchers routinely control the learning environment (e.g., Buehner et al., 2003; Lagnado & Speekenbrink, 2010; Perales & Shanks, 200; Rottman & Keil, 2012). Specifying the number interventions controls the learning environment to assess variables other than strategic approach. This is an effective method as long as 1) the forced strategy is ecologically valid or 2) outcomes do not differ according to strategy. Yet, my findings challenge both of these assumptions.

Overwhelmingly biased. Most participants applied intervention strategies that favored the application of the treatment (this includes intervention strategies that I categorized as unbiased). This overwhelming bias toward the treatment is a departure from research that controls the learning environment by presenting participants with an equal number of each intervention option. Such an intervention strategy was rarely observed (i.e., only in 17 of the 216, or 8% of the observed strategies). Yet, being biased toward the treatment is reasonable with the contextual scenario of hypertension management. There was no reason to suspect a placebo condition at the onset of the experiment and one of the task goals was to lower blood pressure as much as possible. Moreover, participants' bias toward the treatment was reinforced given the lowest possible blood pressure levels were observed with the exclusive application of the

treatment intervention because the probability of an effect in the presence of the causal variable was greater or equal to the probability of an effect in the absence of the causal variable ($P(e|c) \geq P(e|\sim c)$) with all of the contingency manipulations. So, why withhold the treatment?

Higher causal attributions with bias. Treatment-biased strategies led to higher causal attributions with the 75|0 and 75|75 contingency manipulations. Participants expected the treatment to effectively lower their patients' blood pressure levels at least 50% of the time. So, when they applied treatment-biased intervention strategies with the 75|0 and 75|75 contingency manipulations, and observed many efficacious, positive results (Appendix AG contains observed contingencies), they attributed these positive results to the treatment intervention. This finding is representative of confirmation bias. Conversely, the use of unbiased intervention strategies with the 75|0 contingency manipulation produced fewer positive results leading to lower causal attributions. With the 75|75 contingency manipulation, unbiased intervention strategies provided the necessary, and optimal information according to standard error (Appendix B) to realize that both applying and withholding the treatment yielded the same result. This understandably decreased causal attributions. The observation of positive results was rare and well below participants' efficacy expectations with the 25|0 contingency manipulation, regardless of intervention strategy. The inability to produce a recognizable number of positive results with either unbiased or treatment-biased interventions strategies explains why intervention bias does not account for the variance in participants' causal attributions with 25|0 contingency manipulation.

The differences in causal attributions according intervention strategies reinforce Lagnado and Sloman's (2002) belief, and Hagmayer et al.'s (2010) findings that learning experience, in addition to (or possibly instead of) the act of intervening in itself, affects reasoning outcomes. Additionally, the predominant use of treatment-biased intervention strategies questions whether the outcomes associated with controlled learning environments will transfer to contexts where applying a particular intervention option is favorable, as with hypertension management.

Observation Dependence

The results convey important differences between the system conditions relative to participants' observation dependence and the effect of observation dependence on causal attributions. The prevalent use of observation-dependent strategies with the discrete system in comparison to the continuous system condition, suggests that participants detect this difference and are reacting to the available information (or lack there of). Instead of simplifying or reducing the complexity of the reasoning task, the discrete system condition actually appears to have increased the complexity relative to the continuous system condition. If the system condition affects reasoning behavior and outcomes, as my findings suggest, the presence of system differences challenges the assumption that conclusions from examining discrete, or simplified, deconstructed systems can be applied to predict and control behavior with more complex system representations and environments.

Discrete system generates observation dependence. Differences in the prevalence of observation-dependent strategies distinguish between the two system conditions. I defined the strategy of withholding the treatment only when the outcome

was within tolerance (i.e., in the normal blood pressure range) as observation-dependent. I suspect that the information presented with each system condition is responsible for these differences. The cumulative output of the participants' interventions, as well as the implication of each individual intervention selection was updated on each trial with the continuous system condition. Essentially, this cumulative output serves as a data repository in that it aggregates and retains participants' collective set of interventions. With the discrete system condition, the current state on each trial only conveyed information pertaining to each individual intervention selection, which transferred the burden of calculating and maintaining the cumulative statistics of the administered treatments to the observer.

The need to retain multiple events, which would grow accordingly with the delay interval, strains working memory capacity. The prevalence of observation-dependent strategies with the discrete over continuous system condition with the two-trial and three-trial delay manipulations conditions reflects this additional strain. In the discrete system condition, participants' likely directed their attention toward tracking individual intervention to outcome pairings to ease the burden on working memory. Observation-dependent strategies limit the act of intervention selection to a reaction, eliminating the need to expend additional effort on determining which intervention selection to apply with each trial. With the continuous system, the cumulative output appears to have mitigated the effect of delay, as participants' observation dependence was equivalent over levels of delay.

When the data were parsed by levels of contingency, more participants used observation-dependent strategies with the discrete over the continuous system condition

in combination with the 75|0 and 75|75 contingency manipulations. By design, exclusive application of the treatment produced the most efficacious observations across the contingency manipulations. With the 75|0 and 75|75 contingency manipulations, the prevalence of an effect actually decreases the number of treatment interventions applied with observation-dependent strategies (apply only when symptomatic) in comparison to observation-independent strategies. This is because the system's current state (i.e., the presence or absence of the effect) directs intervention behavior with observation-dependent strategies. The continuous system's cumulative output likely emphasized the positive results associated with applying the treatment with the 75|0 and 75|75 contingency manipulations. Subsequently, the accessibility of this information provide reinforcement to apply the treatment irrespective of whether the blood pressure level was in the elevated range (i.e., apply observation-independent strategies) in order to satisfy the goal of lowering the patient's blood pressure level.

With the 25|0 contingency manipulation, the decrease in observation-independence with the continuous system further reinforces the argument that participants were aware of the cumulative output and were focused on lowering their patients' blood pressure. Observation-independent strategies were associated with less treatment interventions with this contingency manipulation because the low rate of a positive effect increased the number of treatment interventions applied with observation-dependent strategies. Moreover, the equivalent use of observation-dependent strategies between the system conditions with the 25|0 contingency manipulation simply identifies a situation in which the use of different information provoked analogous strategic behavior. In sum, the dissimilar, rather than similar use of observation-dependent strategies

between the system conditions suggests participants lost information with the discrete system condition and were attempting to compensate for it with strategy or were shifting their strategy to reflect the information that was available.

Lower causal attributions with dependence. Observation-dependent intervention strategies were associated with lower causal attributions with the 25|0 and 75|0 contingency manipulations. With the 75|75 contingency manipulation, I suspect that the overwhelming prevalence of positive outcome nullified observable differences between observation-independent and observation-dependent intervention strategies, leading participants to conclude that the treatment was effective regardless of strategy. As discussed in the previous section, observation-dependent strategies are a behavioral response to the current system state, rather than the cumulative outcome of one's intervention selections. Focusing on specific intervention to outcome pairings should have increased participants' awareness of inconsistences attributable to both contingency and more delay between participants' expectations (i.e., the treatment will lower the patient's blood pressure) and their observations. Therefore, I suspect that unmet efficacy expectations (i.e., the treatment will be effective at least 50% of the time) and the observation of inconsistencies account for the lower causal attributions that followed observation-dependent strategies with the 25|0 contingency manipulation. The observation of inconsistencies, but not unmet efficacy expectation, may account for the lower causal attributions associated with observation-dependent intervention strategies with the 75|0 contingency manipulation, as well. Additionally, the observation of less efficacious results following observation-dependent strategies with the 75|0 contingency

manipulation (discussed in the previous section) supports the observed decrease causal attributions.

As with intervention bias, the observed findings indicate that differences in observation dependence impacted participants' understanding of the system's underlying environmental parameters. In the previous section, I identified that differences in the prevalence of observation-dependent strategies distinguish between the two system conditions. The reliance on observation-dependent strategies with the discrete system suggests participants focused on changing their patient's blood pressure on the present trial (i.e., experiencing tunnel vision imposed by the available information) rather than the fully considering the changes in their patients' blood pressure over time. An appropriate analogy appears in the context of driving where a person might be aware of the speed that they are driving at various points in time, but unaware of the relationship between these instances (i.e., acceleration or deceleration). Ultimately, information differences tied to system condition direct people's intervention strategy, which subsequently influences their reasoning outcome.

Causal Attributions

Independent of intervention strategy, the results indicate that both system condition and contingency (but not delay) affected causal attributions. The effect of system on causal attribution reflects the informational differences discussed relative to observation dependence. The higher than expected causal attributions observed with the 75|75 contingency manipulation is a departure from previous experimental research. Yet, it seems quite reasonable for participants to conclude that the treatments associated with the 75|0 and 75|75 contingency manipulations were equally as effective given similar

declines in blood pressure. The equivalence of causal attributions over levels of delay identifies limitations of the implemented experimental paradigm, despite replicating patterns observed in previous experimental research.

More extreme with continuous system. Causal attributions of participants that interacted with the continuous system condition were more extreme than observed with the discrete system condition with the 25|0 and 75|0 contingency manipulations. With the 75|0 contingency manipulation, higher causal attributions in the continuous system condition exemplify the affect of observation dependence on causal attributions. As I have discussed above, with this particular contingency manipulation, observation-independent intervention strategies were associated with higher causal attributions and more participants applied observation-independent strategies with the continuous system condition. However, with the 25|0 contingency manipulation, intervention strategy is not a viable explanation for the system difference in causal attributions as observation dependence and intervention bias was equivalent between the system conditions.

Instead, I suspect the effect to null-effect gain ratio and the display of the cumulative blood pressure level are responsible for lowering causal attributions with this contingency manipulation in the continuous system condition. With the continuous system condition, the confounded gain ratio decreased patients' blood pressure by eight intervals with the observation of an effect and increased patients' blood pressure by two intervals with the observation of a null-effect on a continuous scale. (See Appendix I for a review on the selection of this -8: +2 effect to null effect ratio.) Because of the limited probability of an effect and prevalence of the null-effect with the 25|0 contingency manipulation, the blood pressure output progressively increased with the continuous

system condition. The mean value of the cumulative blood pressure level on the 16th trial that was 1.94 points greater than the starting point.[10]

In contrast, participants in the discrete system condition observed binary outputs (i.e., effect = normal, null effect/starting point = elevated). With the 25|0 manipulation, the blood pressure output was almost always in the elevated state. Observing a gradual elevation over time, and an outcome greater than the starting (in the continuous system condition) as opposed to viewing a consistently elevated, but stagnant output (in the discrete system condition), explains why causal attributions associated with the continuous system condition were lower with the 25|0 contingency manipulation. Although the explanations for the shift in causal attributions between systems condition differed with respect to each contingency manipulation (i.e., observation dependence with 75|0 contingency manipulation, and the effect to null-effect gain ratio with the 25|0 contingency manipulation), both reflect participants' sensitivity to informational differences (i.e., cumulative, interval-based display vs. binary output) between the system conditions.

Higher than expected attributions with 75|75 contingency. I expected causal attributions to reflect the differences between causal attributions observed in previous experimental research (Buehner et al., 2003, experiment 1). The direction of change in causal attributions between the 25|0 and 75|0 contingency manipulations was consistent with increases in the probability of an effect given the causal variable (*P(e|c)*), when the probability of an effect in the absence of the causal variable was held constant at zero

[10] For reference, the cumulative output was on average 48.89 and 70.86 points less than the starting point with the 75|0 and 75|75 contingency manipulations, respectively.

($P(e|\sim c) = 0$) observed previously with experimental research (Buehner et al., 2003, experiment 1). In contrast, the higher than expected causal attributions associated with the 75|75 contingency manipulation, and their equivalence to the 75|0 contingency manipulation, were inconsistent with the previously observed decline in causal attribution in combination with increases in the probability of an effect in the absence of the causal variable ($P(e|\sim c)$), when the probability of the effect given the causal variable was held constant at zero ($P(e|c) = 0$).

The predominant use of treatment-biased intervention strategies explains why the observed causal attributions were greater with the 75|75 contingency manipulation than found by Buehner et al (2003, experiment 1), as Buehner et al. presented an equal number of each intervention in their learning task. Treatment-bias intervention strategies provided less opportunity to realize that withholding the treatment was equally as effective as administering the treatment with the 75|75 contingency manipulation, as compared to an unbiased, evenly distributed strategy (Appendix B). Additionally, limiting withhold treatment observations likely contributed to, and even strengthened participants' beliefs that the treatment was accountable for lowering their patients' blood pressure levels. The significantly lower causal attributions of participants that did apply unbiased intervention strategies with the 75|75 contingency manipulation supports this explanation. Although participants' causal attributions only partially supported the predictions, their conclusions are reasonable given their prior biases toward the hypertension treatment options and treatment efficacy, and their observations in the reasoning tasks.

No delay effect. I expected the causal attributions of participants to differ according to magnitude of delay, as well as in combination with the underlying contingency. Increasing the amount of delay between the administration of the causal variable and onset of effect generates outcomes inconsistent with participants' expectations. With this inconsistency, current research suggests that participants will be less likely to attribute the causal variable with the effect (e.g., Greville & Buehner, 2010).

The apparent reductions in causal attributions between the one-trial and two-trial delay manipulations were consistent with Greville and Buehner's (2010) various conditions. Yet, there were no significant differences in causal attributions as a function of delay, in isolation, or in combination with the contingency manipulations when the entire data set was taken into account. The interaction between delay and the system condition to produce differences in observation dependence challenges explanations of either ceiling or floor effects. Nonetheless, the failure to support the hypothesized differences in causal attribution raises concerns about the number of trials in the learning task and the use of trials as a representation of time, which I will discuss within the limitations section.

System Control

I anticipated system control would suffer with increases in delay, as such deficits have been observed in experimental research (Diehl and Sterman, 1995; Sterman, 1989). Instead, no differences in system control were observed as a function of contingency and delay, the effect of delay, or system condition. A post-hoc review of possible differences in system control given contingency, the number of trials in the learning task, and intervention strategy revealed that the experimental design limited differences with the

system control measure. Considerable differences in system control performance were only feasible with the 75|0 contingency manipulation with 16 trials. So, the results suggesting that system control success (lower blood pressure values) negatively relates to causal attributions are likely an artifact of the contingency manipulations. I will discuss the range of potential system control outcomes further in the limitations section.

Theoretical Implications

Previous studies have demonstrated the influence of delay and contingency in isolation on causal attributions within discrete systems (e.g., Greville & Buehner, 2010; Shanks et al., 1989) and on system control with continuous systems (e.g., Brehmer, 1989, 1995; Diehl & Sterman, 1995). Within the causal reasoning literature, studies that have focused on systems that are continuous in nature have focused on decision-making rather than reasoning (Gonzalez & Dutt, 2011) or have maintained a constant temporal relationship between cause and effect (e.g., Hagmayer et al., 2010; Rottman & Keil, 2012). This study extends this research by providing a comparative analysis that considers both causal attributions and system control relative to intervention strategy with both a discrete and a continuous system condition. The findings provide insight into how participants' reasoning strategy relates to their fundamental understanding of causality, in addition to highlighting a difference in information accessibility between the system conditions. Discrepancies between the observed findings and those in the causal reasoning literature challenge the use of particular methods as opposed to the observed findings, particularly the exclusive study of unbiased intervention strategies in controlled learning environments. In general, I am doubtful that causal reasoning theories and results transfer to my paradigm.

Free-operance. This study allowed participants to freely form and test their hypotheses, hypothetically providing them with the greatest amount of information from which to make assessments. The results stress the role of intervention strategy in the formation of causal assessments. This, in combination with participants' bias toward the treatment intervention challenges the extensibility of findings associated with unbiased intervention strategies to this particular contextual scenario. Conceivably, a better understanding of reasoning strategy is necessary prior to focusing on outcomes only associated with a single, possibly irrelevant intervention strategy.

Convergence of causal attribution and control measures. System control does appear to play a role in causal attribution, as the two measures were negatively correlated over levels of delay and with the 75|0 contingency manipulation. I suspect that this relationship failed to persist with the 25|0 and 75|75 contingency manipulations because of the limited possibility of differences in system control with these conditions (discussed in the limitations section). Although success in controlling the system (lowering patients' blood pressure levels) did correspond to higher causal attributions, these attributions did not always reflect the underlying contingency value. This finding suggests that system control is possible without a detailed understanding of the system's underlying parameters and is consistent with research outside the system dynamics and causal reasoning literature that suggests participants' ability to control a system may not correlate with their ability to identify causal variables. Specifically, in their research considering the role of expertise in diabetes management, Lippa, Klein, and Shalin (2008) found that participants' actions or applied knowledge did not always correspond to their declarative knowledge about their condition.

In favor of continuous environments. The prediction of whether or not participants would perform better or worse in the continuous condition as opposed to the discrete condition depends on the paradigm. The causal reasoning literature deconstructs the reasoning environment to focus on how specific factors, like contingency and delay, influence reasoning. Simplifying the problem space should a) facilitate causal attribution, and b) produce outcomes that are transferable to complex reasoning scenarios. The system dynamics literature challenges these assumptions and argues that such simplification a) removes essential contextual cues hindering causal attribution, and b) produces uninformative findings relative to more complex and continuous systems.

Behavioral shifts related to differences in information accessibility provide evidence that simplified system environments and tasks may not be valid precursors to more complex scenarios. Researchers have tried to reduce the burden on working memory and improve the accessibility of information necessary in calculating probabilistic values (specifically $P(e|c)$) by using simultaneous, rather than sequential presentation formats within discrete system environments (Buehner et al., 2003; Cheng & Novick, 1990). This shift in approach is valid given a) the simultaneous presentation of material appears to facilitate learning over sequential presentation of the same information (Imhof, Scheiter, & Gerjets, 2011; Imhof, Scheiter, Edelmann, & Gerjets, 2012) and b) people are more successful in remembering more attributes with fewer variables (e.g., trials), than fewer attributes with more variables (Yntema, 1963). However, contextually, reasoning about a group of events between individuals, or having all the evidence on hand at the same time, may differ from reasoning about a series of events within an individual with evidence accruing over time. The continuous system

framework provides the same binding element (cumulative output) that aids in information accessibly as the simultaneous presentation, yet preserves the reasoning scenario.

Delay in discrete system environments. In the method section, I highlighted several implications of using a trial-based over a duration-based learning task including carry-over between trials, making the discrete system more continuous in nature. Yet, carry-over between trials is not unique to this study as Rottman and Keil (2012) found that people naturally interrelate information presented on separate trials and that carry-over already exists within duration-based, free-operant learning tasks that consider delay (e.g., Buehner & May, 2003, experiment 2; Greville and Buehner, 2010). Duration-based, free-operant studies that have avoided carry-over to maintain the discrete system framework have other experimental design issues related to the segmentation of the overall learning period into smaller units (e.g., Shanks et al., 1989). For example, the system responds to whether an action is produced in the underlying segment, not how many actions are produced, which reduces the observed contingency. Also, the underlying segment is generally determined by the delay specification, so the length of each learning task and segment differs between conditions.

Delay is a kind of carryover, but when implemented in the discrete paradigm, it lacks the full properties of cumulative outcome. This is quite literally an artificial construct that would not occur in nature, which questions the ability to and the utility of studying delay with discrete system conditions.

Practical Implications

People routinely use strategies to reason about causal relationships (e.g., Cheng & Holyoak, 1985; Kelley, 1973). Yet, as we observed, these strategies do not always result in outcomes that align with formal reasoning logic (e.g., Lave, 1988; Tversky & Kahneman, 1974). Perhaps forcing a system 2 measure (explicit measure of attributions) on a system 1 learning task provides an explanation as to why normative models have failed to universally account for causal attributions observed with the present study, as well as consistently within the causal reasoning literature. The research of Tversky and Kahneman (1974), and Kahneman (2003) emphasizes the failure of normative predictions to align with decision-making outcomes associated with system 1 thinking. According to dual process theory, an inability to explicitly describe ones' behavior, as well as fast, automatic, and random behavior is indicative of system 1 thinking. In contrast, system 2 thinking is effortful, slow, calculated, and conscious.

I classified participants self-reported assessment techniques as implicit and explicit based on participants' ability to verbalize their assessment strategy (Figure 22), and causal attributions did not differ according to this categorization of explicitness (Appendix AF). However, my categorization of explicit responses describes heuristics as opposed to rational logic, which Kahneman (2003) classifies as system 1 rather than system 2 thinking. Thus, the insignificant difference in causal attribution according my categorization of explicitness supports Kahneman's classification scheme, suggesting that all of the observed reasoning falls into the system 1 thinking. Consequently, expecting the prescriptive analyses (e.g., ΔP, Power PC theory, Rescola-Wagner) to mathematically, or logically align with participants' system 1 reasoning outcomes

appears unreasonable. Yet, this is what is conventionally done in causal reasoning research that focus' on people's ability to detect and accurately describe underlying contingency values. The analysis of these data provides two encouraging indications of reasonable reasoning: 1) intervention strategies are sensitive to changes in information accessibility (i.e., adaption to the environment) and 2) causal attributions reflect experience.

Limitations

The primary limitation of this study was the number of trials implemented in the learning task. I selected the number of trials for the learning task based on experimental designs in the causal reasoning with contingency literature that use eight trials per intervention option (Buehner et al., 2003; Rottman & Ahn, 2009; Rottman & Keil, 2012). I felt that limiting the number of trials was both ecologically valid (approximately, two weeks) and would foster task vigilance. Yet, these referent experimental designs did not consider delay and controlled the learning environment. Other research paradigms use more trials in the learning task, including causal reasoning with delay (Allan et al., 2003; Greville & Buehner, 2010), system dynamics (Diehl and Sterman, 1995; Sterman, 1989), non-human conditioning (e.g., Claflin, Garrett, & Buffington, 2005), and as do other designs that focus solely on contingency (Jenkins & Ward, 1965). In retrospect, sixteen trials may not have been long enough for participants to recognize the delay interval, as well as for participants' to detect differences between the treatment options, or to generate differences in system control.

Detecting a difference between proportions. Peoples' ability to establish conclusive differences between apply and withhold treatment observations differs

according to contingency. (Appendix A contains the analysis assessing the plausibility of detecting differences between proportions.) With the 16 trial learning task, only the 75|0 contingency manipulation should have provided sufficiently conclusive evidence of a difference between the probability of an effect given a treatment/causal variable ($P(e|c)$) and the probability of an effect in the absence of that treatment/causal variable ($P(e|\sim c)$), in order to attribute the blood pressure observations (e) to the treatment invention (c). The confidence intervals associated with 25|0 and 75|75 contingency manipulations both include zero, which in turn challenges whether 16 trials in the learning task provided enough evidence to conclusively determine that the effect is attributable to the treatment regardless of intervention strategy (optimal or not according to standard error).

Yet, even when the learning task is lengthened to 56 trials, this analysis suggests that differences in people's ability to detect a difference will be limited to specific intervention strategies with the 25|0 contingency manipulation. Further, this analysis posits that detecting a difference between the two proportions with the 75|75 contingency manipulation (or any contingency that $\Delta P = 0$) is prohibitive regardless of the observed number of trials because in actuality there is no difference between the two proportions. The theoretical ability to detect differences between proportions durations beyond 16 trials (i.e., 32 and 56 trials) suggests that the length of the learning task is a limitation that extends beyond this study.

Range of system control outcomes. The range of possible system control outcomes based on the intervention distributions extremes (only applying one treatment intervention and applying the treatment intervention on all of the trials) compliments the difference between proportions assessment. With the 16 trial learning task, only the 75|0

contingency manipulation allowed for prominent differences in system control performance. This is especially true given the system control measure only considered 13 treatment interventions, rather than 16, in order to assess the same number of treatment interventions across the delay manipulations. Table 16 presents the range of possible system control outcomes in relation to three contingency manipulations (25|0, 75|0, and 75|75) and three learning task durations (16, 32, and 56 trials). The upper and lower bound of the range of outcomes is defined using intervention strategy extremes (i.e., when only one treatment is applied and when the treatment is applied on all trials).

Table 16
Range of System Control Outcomes with Extreme Intervention Distributions for each Contingency Manipulation

	Intervention Distribution				
	Treatment	~Treatment	25|0	75|0	75|75
16 trials	1	12	141	141	50
	13	0	111	50	50
32 trials	1	28	150	150	50
	29	0	103	50	50
56 trials	1	52	150	150	50
	53	0	91	50	50

Note. Assumes a starting blood pressure level of 125, with a maximum level of 150 and minimum level of 50. Applying treatment applications reduced blood pressure level by 8 and withholding the treatment increased the blood pressure level by 2. The number of interventions considered by the system control measure is less than the total number of trials so that the same number of interventions is considered across the delay manipulations. For this examination, I assumed a three-trial delay condition, so all of the learning durations were reduced by three trials.

Increasing the number of trials expands the range of possible system control outcomes to the greatest extent with the 25|0 contingency manipulation. Yet, the lower bound (50) is not reached, even with the 56 trial duration, because of the low, overall probability of an effect. In contrast, increasing the number of trials in the learning task does not affect the range of possible system control outcomes with the 75|75 contingency

manipulation because applying and withholding the treatment intervention have the same outcome. With the 75|0 contingency manipulation, the maximum potential range is reached with 32 trials. Perhaps, increasing the number of trials would allow for more variation in intervention strategy, as well, which in turn may be highlight delay effects that are unnoticeable given the current experimental design.

Trials as a representation of time. Even with an expanded number of trials, the potential implications of using trials as a representation of time, rather than actual elapsed time, remain. Although there are examples of trial-based learning environments in the system dynamics literature that consider delay (Brehmer, 1989; Diehl and Sterman; 1995), reasoning about a representation of time may not generate the same effect as directly experiencing and reasoning in real time. Moreover, variation in the actual time between trials and between participants is a consequence of using trials to specify delay, and these differences in the actual time likely nullified the effect of delay on causal attributions. Alternatively, the influence of delay on reasoning outcomes, especially causal attributions may be non-linear or plateau with longer intervals. Shanks et al. (1989) observed non-significant differences with delay intervals comparable to the implemented one-trial and two-trial delay manipulations, as well as with a much longer 16-second manipulation.

I formed my expectation that increases in delay would negatively influence participants' reasoning outcomes without knowing the length or the variability in the actual time associated with the delay intervals. Also, the use of trials to specify the delay interval prohibited the exploration of a no-delay, base-line condition, so establishing whether delay had no effect on participants' causal attributions, or whether this effect was

equivalent across the manipulations is not plausible. Future research should reconsider the use of trials to represent time and recast the experiment to accomplish a better mapping between dwell time in experiment and the cause/effect model.

Future Research

Despite the duration of the learning task and the use of trials to represent time as limitations, this study provides valuable insights to guide future research. Differences in intervention strategy and their influence on reasoning outcomes suggest a natural focal point. Although we have an initial understanding of the relationship between strategy, system control, and causal attribution, more can be learned about this triad of measures. Simplifying the experimental design will allow better identification of the factors that motivate strategic differences. Exploration of additional system parameters (e.g., contingency and gain) will only contribute to the understanding of what information people leverage from different contextual environments.

For example, consideration of a 25|25 contingency may have provided clarity concerning the causal attributions associated with the 25|0 and 75|75 contingency manipulations. If causal attributions were equivalent between the 25|0 and a 25|25 contingency manipulation, as was the case with the 75|0 and 75|75 contingency manipulations, this would suggest that outcomes associated with treatment applications were more salient. Otherwise, the equivalence between the 75|0 and 75|75 contingency manipulations is more attributable to the overall probability of an effect.

Also, consideration of alternative cover stories and task goals, as well as expertise would provide a better understanding of the extensibility of these findings. Lastly, use of discrete system conditions should be avoided moving forward, in favor of continuous

system conditions, given issues accommodating time-based manipulations, decreases in information accessibly, and limited external validity.

V. CONCLUSIONS

This study served as a starting point for future research into how people reasoning about a new treatment when diagnosed with a chronic medical condition. In order to examine reasoning under conditions comparable to hypertension management, I used a novel reasoning paradigm, modeled from the causal reasoning literature, to vary the amount of delay and contingency between actions and outcomes within discrete and continuous system conditions. I asked participants to consider causal scenarios involving a single causal variable (treatment option) that acted in accordance to these underlying experimental manipulations and collected data pertaining to the separate outcome measures employed by the causal reasoning and system dynamics literatures, causal attributions and system control, respectively.

Although the observed findings mostly concur with existing experimental research focused on contingency and delay in isolation with discrete system conditions, the influence of system condition and intervention strategy on causal attributions emphasizes an alternative research focus. Participants' application of observation-independent, as opposed to observation-dependent intervention strategies with the continuous system condition reveals a difference in information accessibility between the system conditions, as do the differences in causal attributions with the 25|0 and 75|0 contingency manipulations. This system effect emphasizes the need to explore continuous, over discrete systems within the causal reasoning literature. Additionally,

allowing people to freely intervene in the learning task highlighted the role of intervention strategy in shaping causal attributions, which challenges the utility of research that focuses on controlled learning experiences founded in theory rather than observation.

Admittedly, there were several limitations related to the experimental design that constrained what analyses could be performed, which in turn restricted the conclusions. Still, the data do provide an initial understanding of the relationship between the reasoning environments and the dependent measures employed separately by the causal reasoning and system dynamics literatures. Yes, there is still a lot to learn with respect to how people reason when confronted with a chronic medical condition, like hypertension. However, my initial findings suggest that people's reasoning outcomes, although inconsistent with normative theory, are actually quite reasonable once the various elements that are involved in everyday cognition, including prior bias, information salience, and reasoning strategy are taken into account.

APPENDIX A

Detecting a Difference between Two Proportions

To understand how different contingency specifications may affect people's ability to detect a difference between *P(e|c)* and *P(e|~c)*, I calculated the confidence interval for the difference between two proportions. I assumed a null hypothesis that there is no difference between the proportions or $\Delta P = P(effect|treatment)-P(effect|\sim treatment) = 0$. Confidence intervals that exclude zero indicate that it is plausible for people reject the null and conclude that the presence of the causal variable is more efficacious than its absence. I considered three contingency values (25|0, 75|0, 75|75[11]) across three intervention distributions (apply treatment-biased, unbiased, and withhold treatment-biased) assuming a 16 trial, 32 trial, and 56 trial learning task.[12]

Table A1 contains the 95% confidence intervals for each contingency by intervention distribution combination**.** Confidence intervals that exclude zero are

---

[11] 25|0 = *P(effect|treatment)* = .25, *P(effect|~treatment)* = .00, 75|0 = *P(effect|treatment)* = .75, *P(effect|~treatment)* = .00, 75|75 = *P(effect|treatment)* = .75, *P(effect|~treatment)* = .75.

[12] This analysis could not be conducted across all of the intervention distributions because the contingency fluctuates on a trial-by-trial basis (e.g., can not have .75 contingency with only one observation).

referenced in bold font. This analysis indicates that peoples' ability to establish conclusive differences between the observations associated with apply and withhold treatment interventions differed according to contingency.[13] Specifically, only the 75|0 contingency manipulation should have provided conclusive enough evidence of a difference between the *P(effect|treatment) and P(effect|~treatment)* to attribute the blood pressure observations to the treatment invention with a 16 trial learning task. The confidence intervals associated with 25|0 and 75|75 contingency manipulations all include zero, challenging whether 16 trials in a learning task provides enough evidence to conclusively determine the influence of *P(effect|treatment)* regardless of intervention strategy (optimal or not according to standard error).

[13] This insight is also apparent with the analysis considering standard error of the differences between proportions presented in Appendix B.

Table A1

Difference of Proportions Confidence Intervals for a subset of Intervention Distributions and Contingency Values

	Intervention Distribution		*95% CI*					
			25|0		75|0		75|75	
	Treatment	~Treatment	*LL*	*UL*	*LL*	*UL*	*LL*	*UL*
16 trials	4	12	- .13	.78	**.14**	**.99**	- .56	.40
	8	8	- .20	.64	**.19**	**.96**	- .44	.44
	12	4	- .38	.57	**.06**	**.93**	- .40	.56
32 trials	4	28	- .03	.78	**.19**	**.98**	- .55	.31
	16	16	- .04	.53	**.38**	**.92**	- .32	.32
	28	4	- .37	.45	**.11**	**.89**	- .31	.55
56 trials	4	52	- .00[1]	.78	**.21**	**.99**	- .54	.28
	28	28	**.05**	**.45**	**.50**	**.89**	- .24	.24
	52	4	- .36	.39	**.13**	**.86**	- .28	.54

Note. *Treatment* = total number of applications of the treatment. *~Treatment* = total number of instances that the treatment was withheld. *CI* = Confidence Interval of a Proportion, including continuity correction. *LL* = Lower Limit. *UL* = Upper Limit. 25|0 = *P(effect|treatment)* = .25, *P(effect|~treatment)* = .00, 75|0 = *P(effect|treatment)* = .75, *P(effect|~treatment)* = .00. 75|75 = *P(effect|treatment)* = .75, *P(effect|~treatment)* = .75.

[1] -0.0018

APPENDIX B

Standard Error of the Difference between Two Proportions

Standard error is the degree to which the sampled value reflects the population value. Smaller standard error measurements place smaller confidence intervals around the sampled value, providing a more precise estimate. Using the concept of standard error, I examined all of the possible treatment intervention distributions assuming a 16 trial learning task to determine the likelihood that each experience (the sample) would allow for the realization of different underlying contingencies (the population value). I considered these treatment intervention distributions across three contingency values (25|0, 75|0, 75|75[14]) that varied according to the probability that the presence and absence of the treatment generated the expected effect on the patients' blood pressure. Table B1 contains the standard error of the difference between two proportions for each contingency by intervention distribution combination. For reference, I highlighted the lowest standard error value for each contingency value with bold font and identified the evenly distributed treatment intervention strategy (i.e., *P(treatment)* = *P(~treatment)* = 8).

[14] 25|0 = *P(effect|treatment)* = .25, *P(effect|~treatment)* = .00; 75|0 = *P(effect|treatment)* = .75, *P(effect|~treatment)* = .00; 75|75 = *P(effect|treatment)* = .75, *P(effect|~treatment)* = .75

Concerning the specific standard error values, when *P(effect|~treatment)* = zero the error is naturally smaller because no error is affiliated with one of the two proportions. Also, with the 75|0 and 25|0 contingency values, treatment-biased intervention distributions produce the lowest standard error outcomes. This implies that treatment intervention strategies biased toward the effective action/variable will provide the most information to the observer when the probability of the conjugate action producing an effect is zero. Conversely, an unbiased intervention strategy lowers the standard error to the greatest degree with the 75|75 contingency value. This outcome highlights the need to observe both actions equivalently when both actions are capable of producing efficacious observations, in order to obtain the most information about the underlying contingencies. Further, this analysis reinforces that an unbiased treatment intervention strategy (*P(treatment) = P(~treatment))* is not the optimal treatment intervention strategy across all contingency values. Although an evenly distributed, unbiased intervention strategy will produce the lowest standard error on average, a treatment-biased intervention strategy is a reasonable approach if there is reason to believe that the treatment option will be effective.

Table B1

Standard Error of the Difference between Two Proportions given possible Treatment Intervention Distributions assuming 16 Trials

Intervention Distribution		*SE*		
Treatment	~Treatment	25|0	75|0	75|75
1	15	0.43	0.43	0.45
2	14	0.31	0.31	0.33
3	13	0.25	0.25	0.28
4	12	0.22	0.22	0.25
5	11	0.19	0.19	0.23
6	10	0.18	0.18	**0.22**
7	9	0.16	0.16	**0.22**
8	8	0.15	0.15	**0.22**
9	7	0.14	0.14	**0.22**
10	6	0.14	0.14	**0.22**
11	5	0.13	0.13	0.23
12	4	0.13	0.13	0.25
13	3	0.12	0.12	0.28
14	2	0.12	0.12	0.33
15	1	**0.11**	**0.11**	0.45

Note. *SE* = Standard Error of the difference between two proportions. The following is the calculation used to compute SE: $\sqrt{\frac{(\square_1)(1-\square_1)}{\square_1} + \frac{(\square_2)(1-\square_2)}{\square_2}}$ Treatment = total number of applications of the treatment. ~Treatment = total number of instances that the treatment was withheld. 25|0 = *P(effect|treatment)* = .25, *P(effect|~treatment)* = .00; 75|0 = *P(effect|treatment)* = .75, *P(effect|~treatment)* = .00; 75|75 = *P(effect|treatment)* = .75, *P(effect|~treatment)* = .75.

APPENDIX C

Outcome Differences based on the Strategic Dimension of Observation Dependence

When freely intervening with a system, participants may select and apply a strategy independent of the observed outcomes. Alternatively, it is plausible for participants to select subsequent moves based on recently observed outcomes or changes in outcomes. Using this strategic concept of observation dependence, I examined three viable treatment intervention distributions (observation-dependent, observation-independent and treatment biased, and observation-independent and unbiased) and calculated the frequency that the observed blood pressure level was normal (the effect was present), as opposed to elevated (the effect was absent).[15,16]

First, I considered the observation-dependent intervention strategy of applying and withholding the treatment in response to elevated and normal blood pressure observations, respectively. Essentially, this strategy follows the decision rule of only applying the treatment when symptomatic. Notice in Table C1, that the underlying

---

[15] I used a one-trial delay manipulation in the discrete system condition separated for each of the proposed contingency manipulations. The same analysis can be performed with the continuous system condition when declines relative to increases in blood pressure are used as the frequency measure, rather than frequency in the normal range.

[16] I reviewed 15 interventions rather than 16 with this analysis because the outcome associated with the intervention applied on 16th trial is not displayed to participants.

contingency values direct the treatment intervention distribution with this, and other observation-dependent strategies. With this particular strategy, the treatment is applied more often with the 25|0 contingency value (12 times) than the 75|0 contingency value (9 times) because the lower contingency of the treatment on the effect results in more elevated blood pressure observations (no effects). When the blood pressure observation is frequently normal (the effect occurs frequently) as with the 75|75 contingency value, the number of apply treatment selections with our observation-dependent intervention strategy is further reduced to five applications.

To contrast this observation-dependent approach, I considered two observation-independent strategies: treatment-biased and unbiased. A treatment-biased approach assumes that the treatments recommended by our healthcare providers will improve our condition. This approach applies 12 treatment applications[17] with three withhold applications (to represent occasional forgetfulness) evenly interspersed. The unbiased approach evenly distributes the intervention options with eight application of the treatment (c) followed by seven instances of withholding the treatment ($\sim c$). Again, this is typically the intervention strategy used by researchers that control the learning environment. In Table C1, notice that the treatment distribution does not vary with the observation-independent intervention strategy across the explored contingencies as it did relative to the observation-dependent intervention strategy.

[17] 12 treatments was the mean number of treatments observed with both the discrete (n = 45, M = 12.36, SD = 3.91) and continuous (n = 74, M = 11.97, SD = 3.63) system condition with observation-independent strategies

Also shown in Table C1 is the frequency that the blood pressure observation is in the normal range (i.e., % Normal). This output is dependent on both the contingency value and the distribution of treatment interventions. The influence of contingency on % Normal outputs is evident with the observation-independent intervention strategies because the treatment distribution is held constant. Increasing the combined contingency, or increasing the likelihood of an effect will occur regardless of the source (i.e., treatment or ~treatment), increases the frequency that the blood pressure observation is in the normal range. Comparing the % Normal values within each contingency value illustrates that there is a range of possible observations, and that % Normal outputs are directly dependent on the distribution of treatment interventions. For example, the frequency that the blood pressure observation is in the normal range with the 75|0 contingency value decreases with the observation-dependent intervention strategy (47%) relative to the observation-independent and treatment-biased intervention strategy (60%) given the reduction in treatment applications. That said, the range of possible % Normal values does vary as a function of the contingency specification, meaning treatment intervention strategy will affect blood pressure observations with certain contingency values (e.g., 75|0) to a greater degree than others (e.g., 75|75).

Table C1

Distribution of Treatment Interventions and the Implication of Treatment Intervention Strategy on Ability to Control the System separated by Contingency

	Observation-dependent			Observation –independent					
				Treatment-biased			Unbiased		
	c	~c	% Normal	c	~c	% Normal	c	~c	% Normal
25|0	12	3	.20	12	3	.20	8	7	.13
75|0	9	6	.47	12	3	.60	8	7	.40
75|75	5	10	.73	12	3	.73	8	7	.73

Note. c = total number of applications of the treatment. ~c = total number of instances that the treatment was withheld. % Normal = the frequency the observed blood pressure level was normal, as opposed to elevated. 25|0 = *P(effect|treatment)* = .25, *P(effect|~treatment)* = .00. 75|0 = *P(effect|treatment)* = .75, *P(effect|~treatment)* = .00. 75|0 = *P(effect|treatment)* = .75, *P(effect|~treatment)* = .75. This analysis considered 15 interventions rather than 16 because the outcome associated with the intervention applied on 16th trial is not displayed to participants.

APPENDIX D

Distribution of Participants between Experimental Manipulations

I used separate one-way ANOVAs to confirm that there were no differences in participants' demographics (i.e., age, sex, grade, ESL) between the nine experimental manipulation combinations (contingency * delay). I performed separate analyses for the discrete and continuous system conditions. There were no significant differences (see Table D1 and D2). Also, I used independent samples *t*-tests to compare the system conditions using the same demographic measures. Again, there were no significant differences in the demographic measures between the system conditions (see Table D3).

Table D1

Demographic Differences between Participants in the Discrete System Condition

		Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Age	Between Groups	134.722	8	16.840	0.846	.572	.20
	Within Groups	537.500	27	19.907			
	Total	672.222	35				
Sex	Between Groups	3.000	8	.375	1.688	.147	.33
	Within Groups	6.000	27	.222			
	Total	9.000	35				
Grade	Between Groups	2.389	8	.299	1.112	.386	.25
	Within Groups	7.250	27	.269			
	Total	9.639	35				
ESL	Between Groups	2.500	8	.313	1.985	.087	.37
	Within Groups	4.250	27	.157			
	Total	6.750	35				

Table D2

Demographic Differences between Participants in the Continuous System Condition

		Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Age	Between Groups	217.222	8	27.153	1.484	.209	.31
	Within Groups	494.000	27	18.296			
	Total	711.222	35				
Sex	Between Groups	3.222	8	.403	1.891	.103	.36
	Within Groups	5.750	27	.213			
	Total	8.972	35				
Grade	Between Groups	3.389	8	.424	0.847	.571	.20
	Within Groups	13.500	27	.500			
	Total	16.889	35				
ESL	Between Groups	1.500	8	.188	0.964	.483	.22
	Within Groups	5.250	27	.194			
	Total	6.750	35				

Table D3

Demographic Differences between System Conditions

	M				95% CI		
	Discrete System	Continuous System	*t*(70)*	*p*	*LL*	*UL*	*d*
Age	20.22 (4.38)	20.28 (4.51)	-0.053	.958	-2.145	2.034	-0.012
Sex	1.50 (0.51)	1.47 (0.51)	0.233	.817	-0.210	0.266	0.056
Grade	1.31 (0.53)	1.44 (0.69)	-0.957	.342	-0.429	0.151	-0.229
ESL	1.25 (0.44)	1.25 (0.44)	0.000	1.000	-0.206	0.206	0.000

Note. $N = 72$. *SD* in parentheses. *CI* = confidence interval; *LL* = lower limit; *UL* = upper limit. Performance distribution only considers data from participants that were included in the final data analyses. Sex coding: 1 for males, 2 for females. Grade coding: 1 for freshman, 2 for sophomores, 3 for juniors, and 4 for seniors. ESL coding: 1 for native English speaker, 2 for non-native English speaker *degrees of freedom differed for the grade analysis due to heteroscadicity: *t*(65.14).

Table D4 contains performance on the English proficiency measure separated by system condition. Chi-square analyses indicated that participants' ability to select the correct the intermediate ($X^2(1, n = 144) = 1.029$, $p = .310$, *Cramer's V* = .08) or upper

intermediate ($X^2(1, n = 144) = 0.150$, $p = .699$, *Cramer's V* = .03) vocabulary words when given a definition was equivalent between the system conditions.

Table D4
Performance on the English proficiency measure

	Discrete System		Continuous System	
	Correct	Incorrect	Correct	Incorrect
Intermediate	71	1	69	3
ugly	17	0	15	0
blind	14	0	12	0
sensible	5	1	10	0
unlikely	12	0	8	2
lazy	10	0	17	0
shy	13	0	7	1
Upper intermediate	68	4	69	3
victim	13	0	11	0
litter	10	3	11	1
reward	8	0	18	0
courage	8	0	9	1
plug	15	0	12	0
currency	14	1	8	1

Note. $N = 72$. Performance distribution only considers data from participants that were included in the final data analyses

Table D5 contains the mean and standard deviation values of participants' responses to Health Belief Questionnaire items separated by system condition. Independent samples *t*-tests confirm that there were no significant differences in responses between the system conditions. Of particular interest is the last item, which states that people *should stop their treatment every now and again* as this belief may have influenced intervention strategy. Again, there was not a significant difference between the system conditions relative to this belief.

Table D5

Health Belief Questionnaire Responses

	M				95% *CI*		
	Discrete System	Continuous System	*t(70)*	*p*	*LL*	*UL*	*d*
1. Too Many	3.08 (1.03)	3.19 (1.04)	-0.457	.649	-.596	.374	-0.109
2. Natural	3.08 (1.18)	3.31 (1.01)	-0.859	.393	-.738	.294	-0.205
3. Harm	2.64 (0.99)	2.28 (0.85)	1.661	.101	-.072	.795	0.397
4. Stop	2.72 (1.03)	3.11 (1.09)	-1.555	.124	-.888	.110	-0.372

Note. $N = 72$. *SD* in parentheses. *CI* = confidence interval; *LL* = lower limit; *UL* = upper limit. All responses were collected on a 5 point scale ranging from strongly disagree (1) to strongly agree (5). All tests two-tailed. 1. Too many = Doctors use too many medicines. 2. Natural = Natural remedies are safer than medicines. 3. Harm = Medicines do more harm than good. 4. Stop = People who take medicines should stop their treatment for a while every now and again.

APPENDIX E

Fixed versus Random Contingency Specification

The two methods used for implementing contingency are random (Allen et al., 2003; Buehner et al., 2003; Buehner & May, 2003; Jenkins & Ward, 1965; Shank et al., 1989; Perales & Shanks, 2003; Wasserman et al., 1993) and fixed assignment (Rottman & Keil, 2012). With random assignment, researchers randomly generate and display a sequence of observations at the rate of contingency for each participant, whereas this sequence of observations is specified a-priori and remains consistent between participants with fixed assignment. To determine which method to use, I assessed how the actual contingency (i.e., the contingency observed by participants) may vary from the specified contingency (i.e., the parameter designated by the experimenter) within a free-operant learning environment with both the fixed and random method of contingency specification. For both conditions, I assumed a specified contingency of $P(effect|treatment) = .75$ and $P(effect|\sim treatment) = .00$. Table E1 lists representative actual $P(effect|treatment)$ contingency values generated using fixed and random assignment for the possible treatment intervention distributions assuming a 16 trial learning task.

For the fixed assignment method, I examined having three out of every four apply treatment interventions produce an effect and every fourth selection produce the null effect. Again, this assumes a contingency value of $P(effect|treatment) = .75$. By design, the actual contingency values associated with fixed assignment remain consistent between samples, so there is no need to repeat this procedure as done with the random

assignment condition. For the random assignment condition, I randomly generated a set of numbers for each of the possible treatment intervention distributions assuming a 16 trial learning task. The amount of numbers generated for each set was equal to the number of treatment applications in each intervention distributions. For instance, I generated one number for the 1:15::treatment:~treatment intervention distribution. The generated numbers ranged from 0 and 100. I paired the numbers below and above 75 with the effect and null-effect respectively, given the contingency condition of *P(effect|treatment)* = .75. I repeated this procedure three times to assess variability.

With fixed assignment method, if the treatment is applied one time (i.e., the 1:15 intervention distribution), you will see exactly one observation of the effect (100% of the apply cases). In contrast, with the random assignment method, it is possible that the one application of the treatment intervention will not produce the effect. If the treatment is applied on all 16 of the trials (i.e., the 16:0 intervention distribution) with the fixed assignment method, then .75 (12 trials) of the outcomes will display the effect and .25 (4 trials) the null effect. Notice that the actual, observed contingency is always greater than .75 when the number of treatments is not divisible by 4 with the fixed assignment method. This method requires a decision rule as to whether or not to favor the effect or null-effect, whereas with random assignment, whether the actual contingency to either more or less than the contingency specification is exactly that - random.

This analysis confirms that the actual contingency value will differ from the specified contingency value with both methods of contingency implementation. The mean value of actual contingencies is closer to the specified contingency value of .75 with random assignment (1: M = .70, SD = .24; 2: M = .80, SD = .78; 3: M = .81, SD =

.80). However, there is considerably more variation[18] between the actual contingency values across the possible intervention distributions as compared to the fixed assignment implementation (M = .83, SD = .09). The decision to use of fixed assignment over random assignment in the specification contingency reflects the benefit of consistency in experience between participants in a free-operant learning environment and less variability in actual, observed contingency with a limited number of trials.

Table E1

Representative Actual P(effect|treatment) Contingency Values after 16 trials for the Possible Treatment Intervention Distributions with Fixed and Random Distribution of Contingency

Intervention Distribution		Fixed	Random		
Treatment	~Treatment		1	2	3
1	15	1.00	.00	1.00	1.00
2	14	1.00	1.00	.50	1.00
3	13	1.00	.66	1.00	1.00
4	12	.75	.50	.75	1.00
5	11	.80	.60	.60	.60
6	10	.83	.83	.83	.67
7	9	.86	.71	.86	.71
8	8	.75	.50	.88	.75
9	7	.78	.67	.78	.78
10	6	.80	.60	1.00	.80
11	5	.82	1.00	.73	.82
12	4	.75	.75	.75	.83
13	3	.77	.92	.77	.77
14	2	.79	.93	.71	.86
15	1	.80	.73	.73	.80
16	0	.75	.75	.75	.63

Note. Treatment = total number of applications of the treatment. ~Treatment = total number of instances that the treatment was withheld. Fixed = Fixed assignment. Random = Random assignment. 1, 2, 3 represent repeated simulations of the deriving actual contingency with the random assignment approach. Specified contingency: *p(effect|treatment)* = .75, *p(effect|~treatment)* = .00.

[18] $Fmax_1(15) = 7.11$, $p < .05$; $Fmax_2(15) = 75.11$, $p < .05$; $Fmax_3(15) = 79.01$, $p < .05$

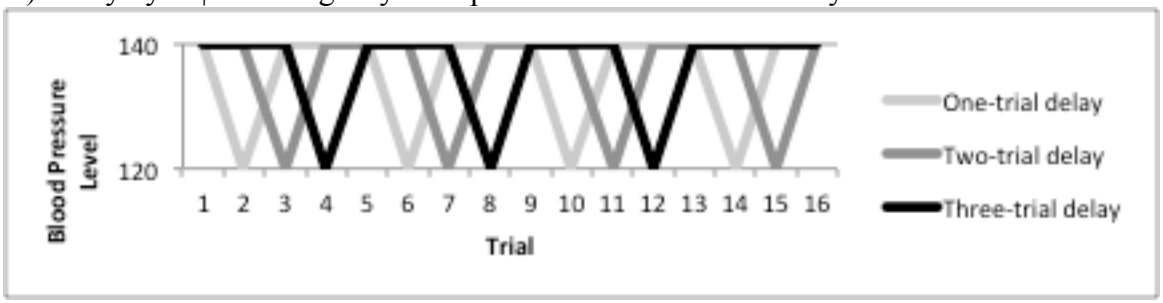
APPENDIX F

Sample of Blood Pressure Observations given Treatment Intervention Strategies, Delay, Contingency, and System Condition

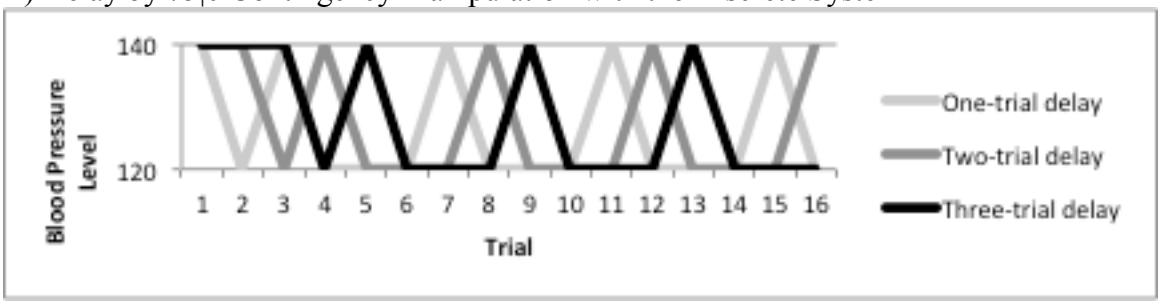
Below are figures depicting blood pressure observations with a subset of intervention strategies. Each figure presents these outcomes relative to the nine experimental treatment combinations (delay * contingency). The figures are separated by system condition, as well. Figure F1 through Figure F5 present observations in the discrete system framework. Figure F6 through Figure F10 present observations in the continuous system framework. Figures F1 and F6 present the treatment intervention strategy of applying the treatment on all of the trials. Figures F2 and F7 present the treatment invention strategy of applying the treatment only on the first trial and withholding the treatment for the remaining trials. With Figures F3 and F8, the treatment is applied on the first eight trials and withheld of subsequent trials. Figures F4 and F9 present a treatment intervention strategy that alternates applying two treatment interventions with withholding two treatment interventions. Lastly, Figures F5 and F10 present a treatment-biased intervention strategy of applying the treatment on the first 11 trials and withholding the treatment for the remaining trials.

Discrete System

A) Delay by 25|0 Contingency Manipulation with the Discrete System



B) Delay by 75|0 Contingency Manipulation with the Discrete System



C) Delay by 75|75 Contingency Manipulation with the Discrete System

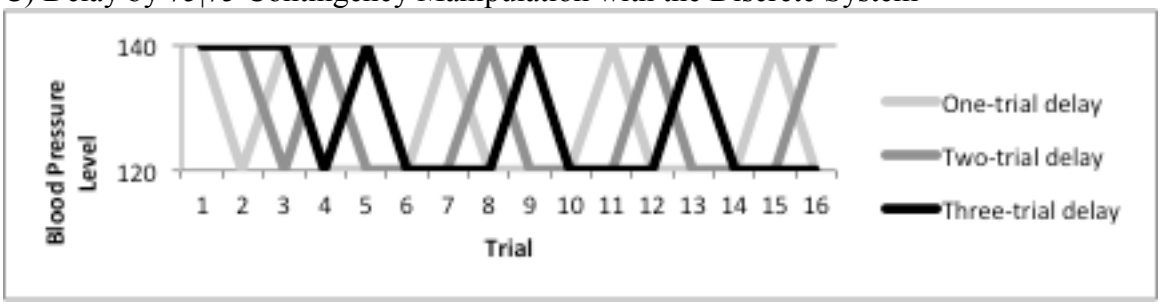
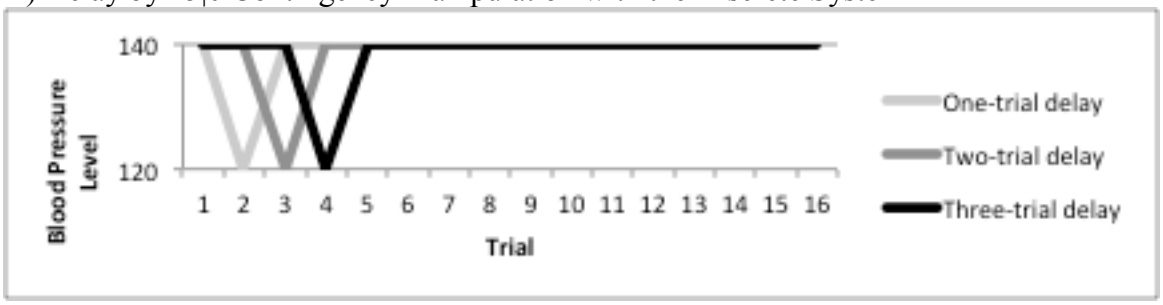
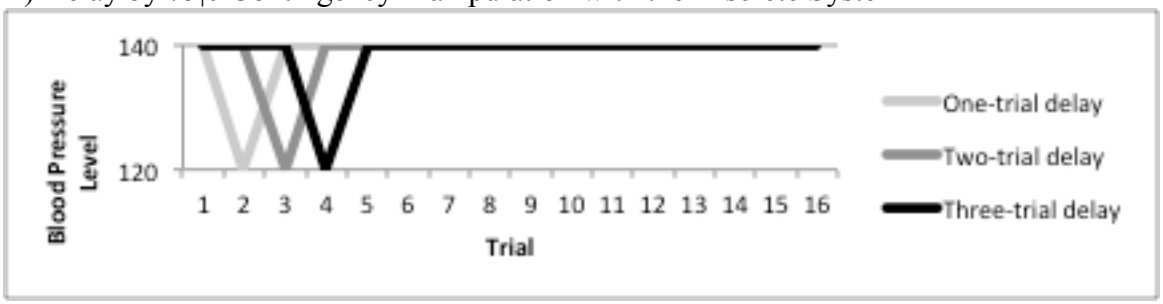


Figure F1. Illustration of blood pressure observations assuming a treatment intervention strategy of applying the treatment on all of the trials with the discrete system condition. Observations are separated by delay manipulations and the a) 25|0 contingency manipulation, b) 75|0 contingency manipulation, and c) 75|75 contingency manipulation. 140 and 120 represent an elevated and normal blood pressure levels respectively. The observations are the same for the 75|0 and 75|75 contingency manipulations.

A) Delay by 25|0 Contingency Manipulation with the Discrete System



B) Delay by 75|0 Contingency Manipulation with the Discrete System



C) Delay by 75|75 Contingency Manipulation with the Discrete System

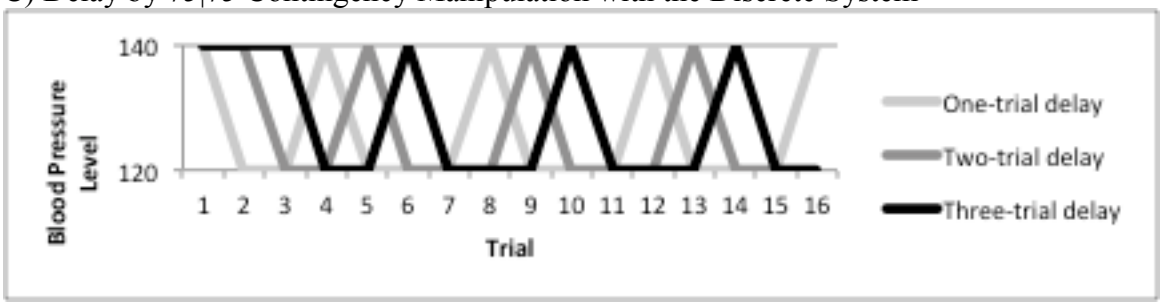
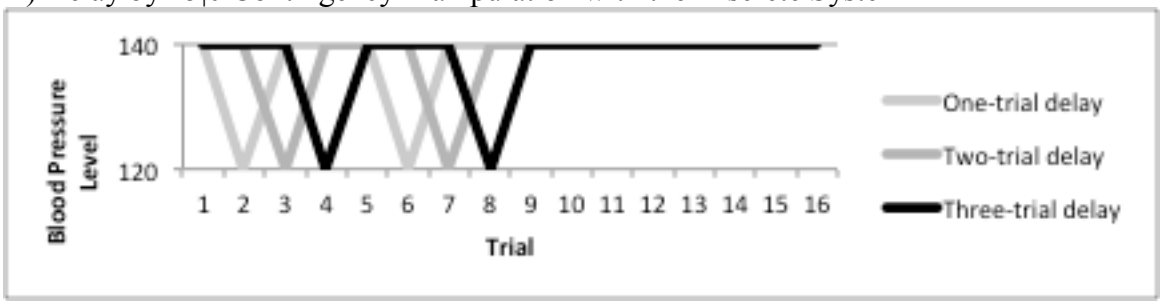
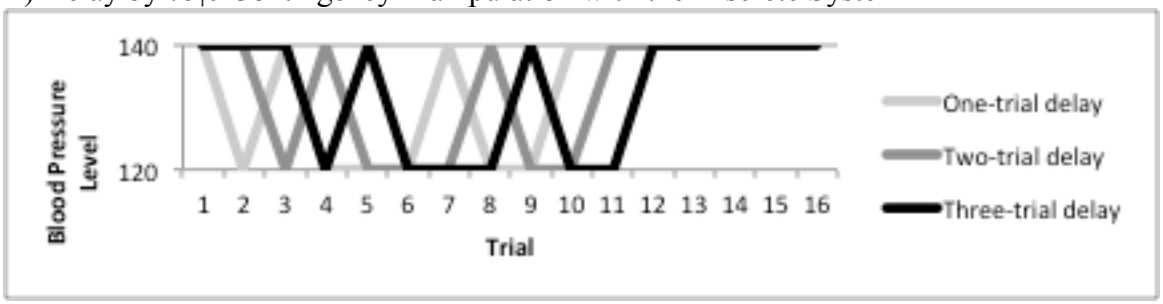


Figure F2. Illustration of blood pressure observations assuming a treatment intervention strategy of applying the treatment only on the first trial with the discrete system condition. Observations are separated by delay manipulations and the a) 25|0 contingency manipulation, b) 75|0 contingency manipulation, and c) 75|75 contingency manipulation. 140 and 120 represent an elevated and normal blood pressure levels respectively. The observations are the same for the 25|0 and 75|0 contingency manipulations.

A) Delay by 25|0 Contingency Manipulation with the Discrete System



B) Delay by 75|0 Contingency Manipulation with the Discrete System



C) Delay by 75|75 Contingency Manipulation with the Discrete System

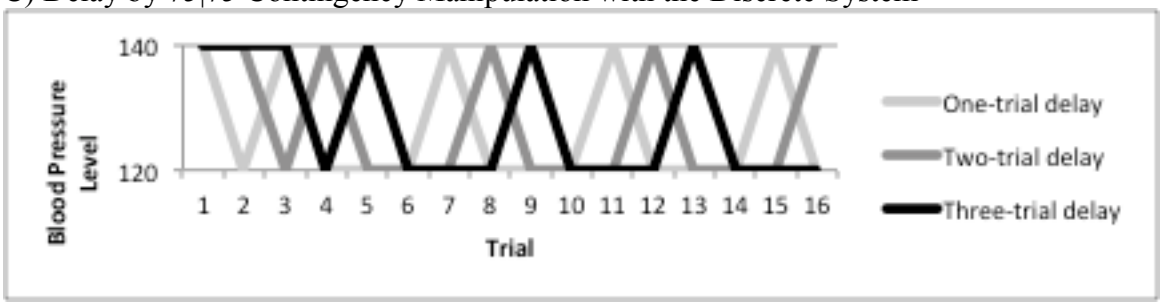
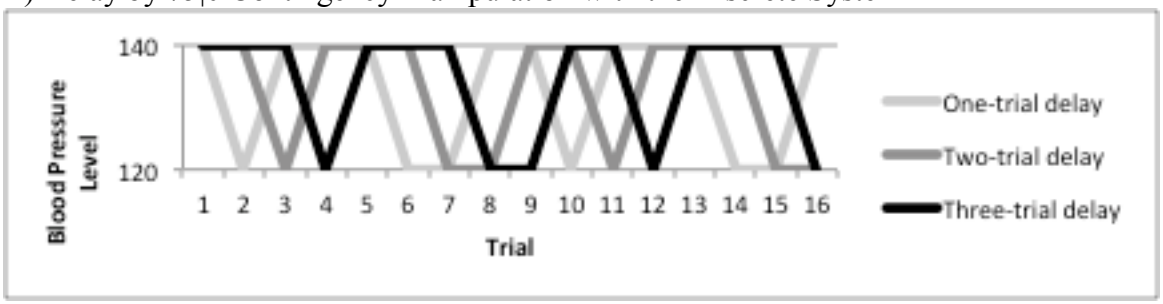


Figure F3. Illustration of blood pressure observations assuming a treatment intervention strategy of applying the treatment on trials 1-8 with the discrete system condition. Observations are separated by delay manipulations and the a) 25|0 contingency manipulation, b) 75|0 contingency manipulation, and c) 75|75 contingency manipulation. 140 and 120 represent an elevated and normal blood pressure levels respectively.

A) Delay by 25|0 Contingency Manipulation with the Discrete System

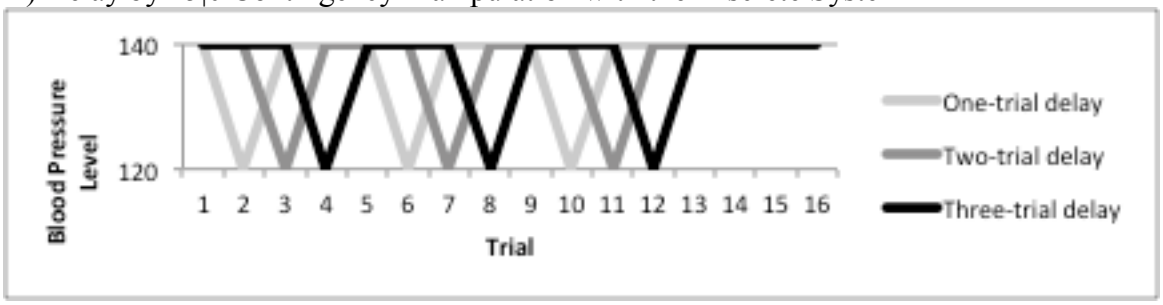
B) Delay by 75|0 Contingency Manipulation with the Discrete System



C) Delay by 75|75 Contingency Manipulation with the Discrete System

Figure F4. Illustration of blood pressure observations assuming a treatment intervention strategy that alternates applying two treatment interventions with withholding the treatment on two trials with the discrete system condition. Observations are separated by delay manipulations and the a) 25|0 contingency manipulation, b) 75|0 contingency manipulation, and c) 75|75 contingency manipulation. 140 and 120 represent an elevated and normal blood pressure levels respectively.

A) Delay by 25|0 Contingency Manipulation with the Discrete System



B) Delay by 75|0 Contingency Manipulation with the Discrete System

C) Delay by 75|75 Contingency Manipulation with the Discrete System

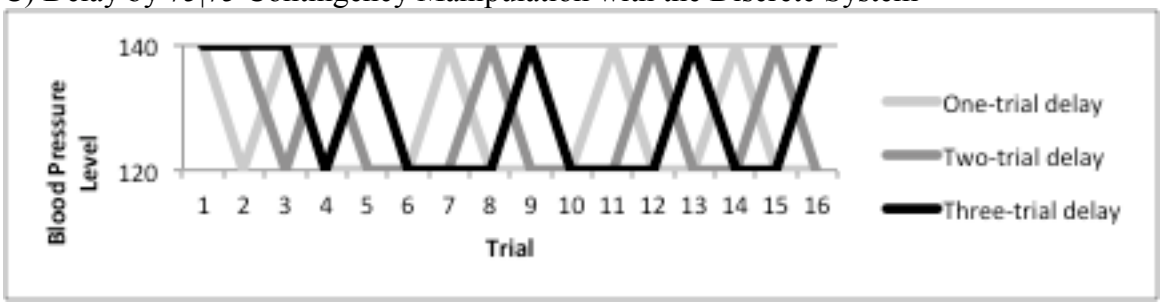
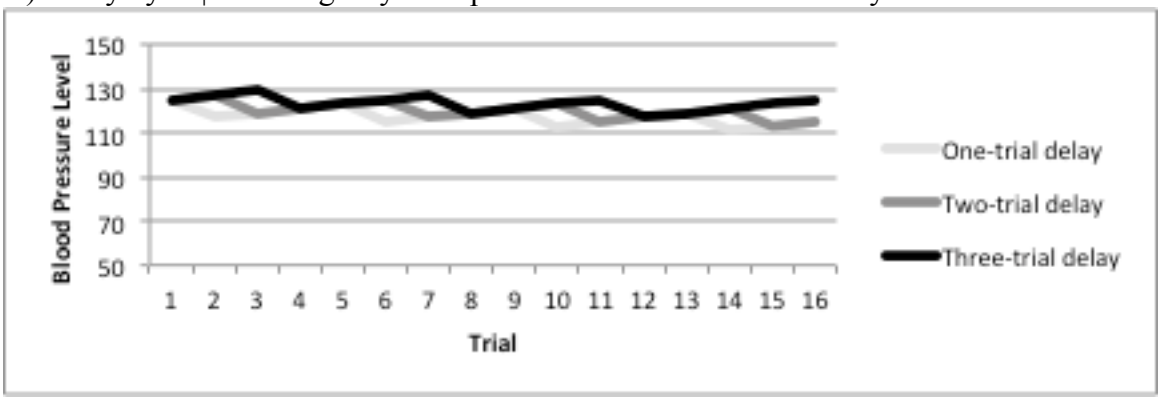


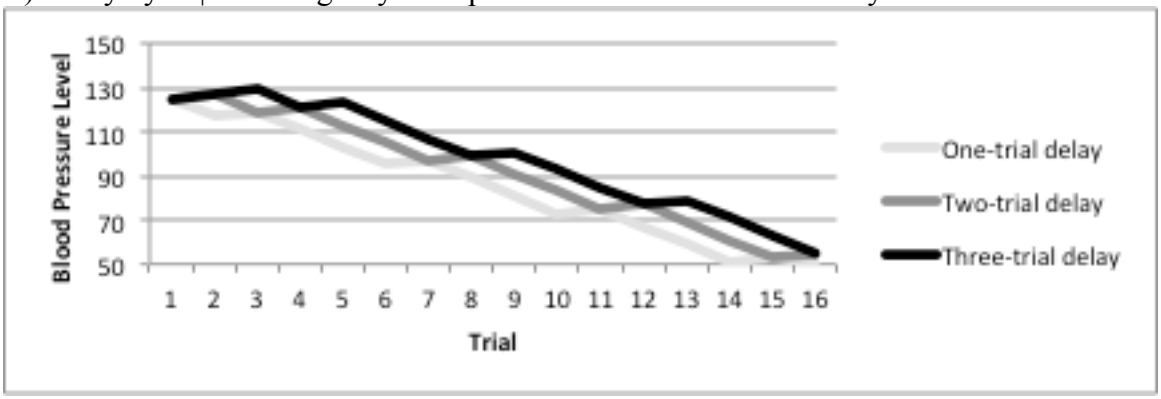
Figure F5. Illustration of blood pressure observations assuming a treatment-biased intervention strategy of applying the treatment on the first 11 trials and withholding the treatment for the remaining trials with the discrete system condition. Observations are separated by delay manipulations and the a) 25|0 contingency manipulation, b) 75|0 contingency manipulation, and c) 75|75 contingency manipulation. 140 and 120 represent an elevated and normal blood pressure levels respectively.

Continuous System

A) Delay by 25|0 Contingency Manipulation with the Continuous System



B) Delay by 75|0 Contingency Manipulation with the Continuous System



C) Delay by 75|75 Contingency Manipulation with the Continuous System

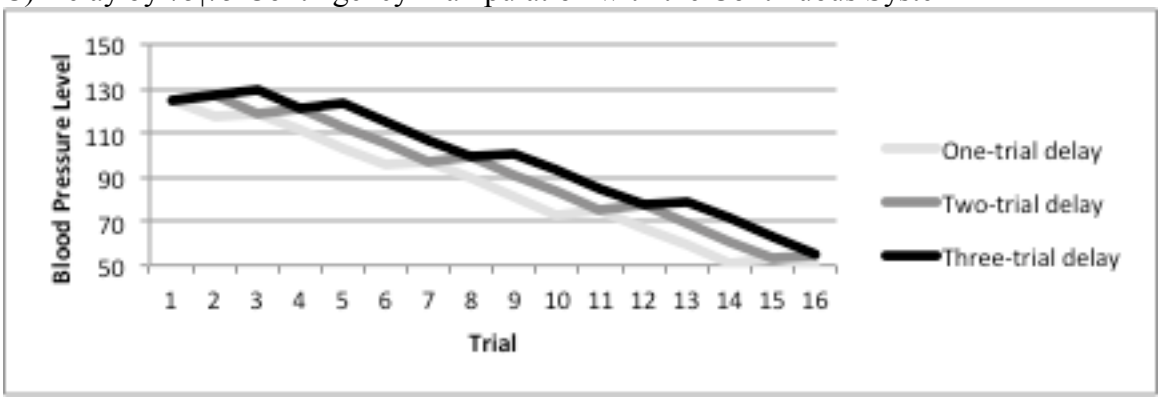
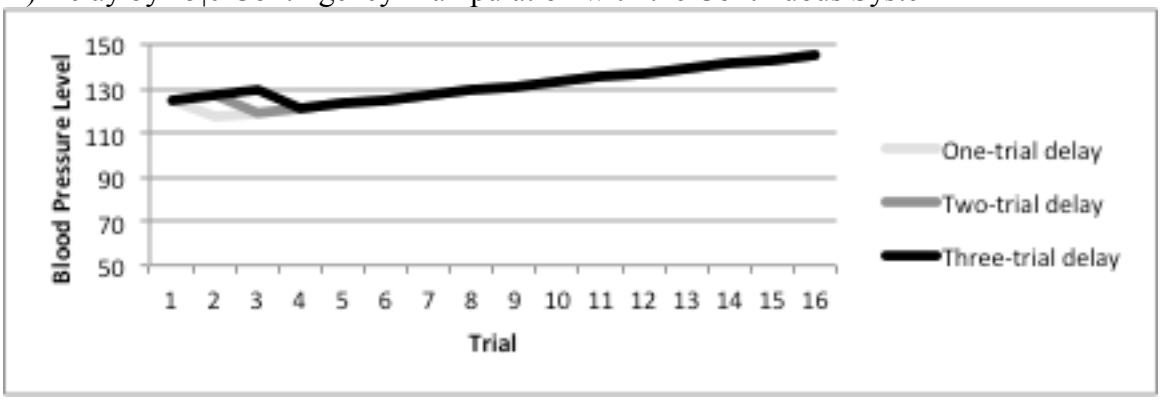
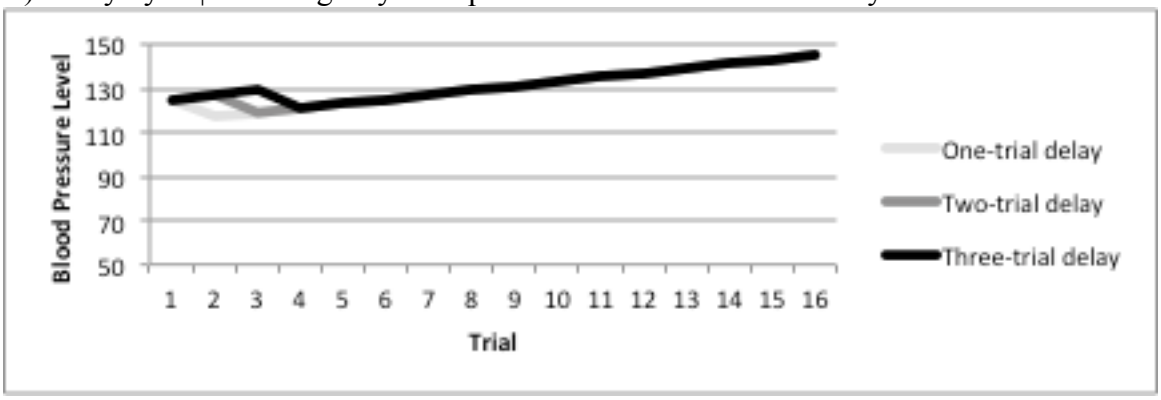


Figure F6. Illustration of blood pressure observations assuming a treatment intervention strategy of applying the treatment on all of the trials with the continuous system condition. Observations are separated by delay manipulations and the a) 25|0 contingency manipulation, b) 75|0 contingency manipulation, and c) 75|75 contingency manipulation. The observations are the same for the 75|0 and 75|75 contingency manipulations.

A) Delay by 25|0 Contingency Manipulation with the Continuous System



B) Delay by 75|0 Contingency Manipulation with the Continuous System



C) Delay by 75|75 Contingency Manipulation with the Continuous System

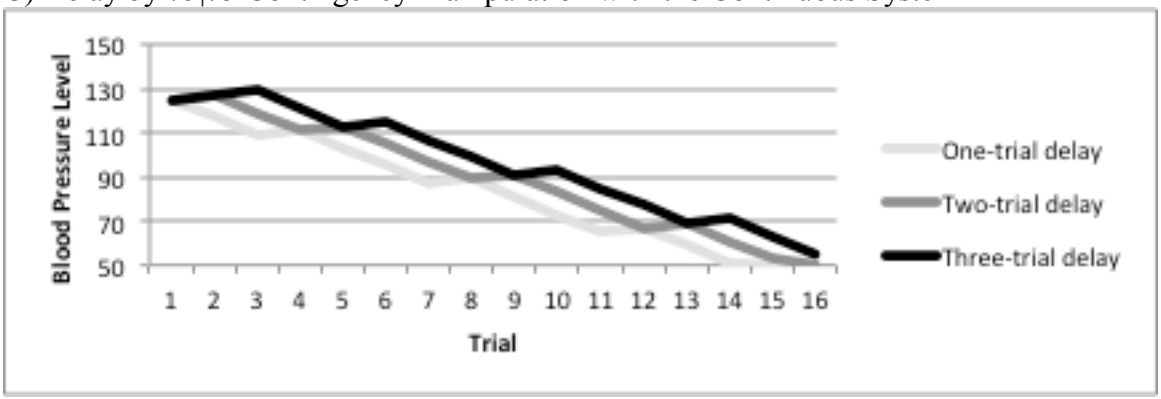
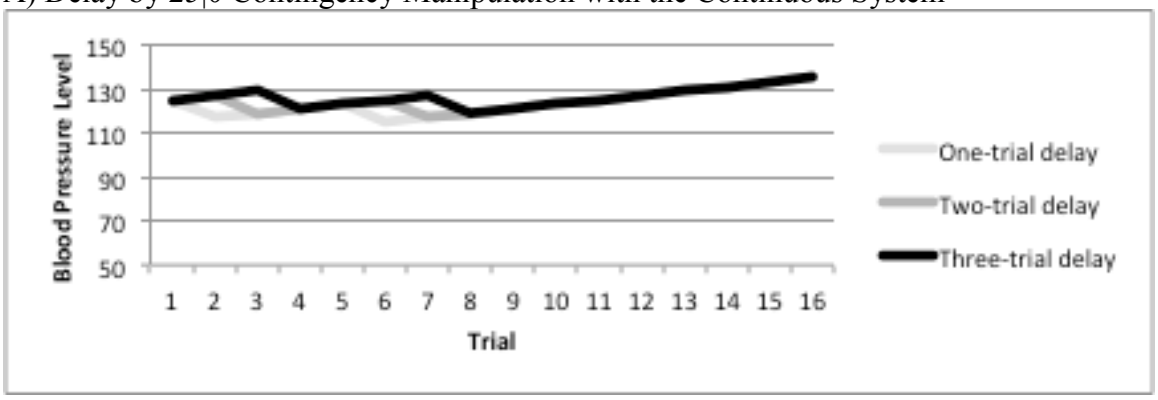
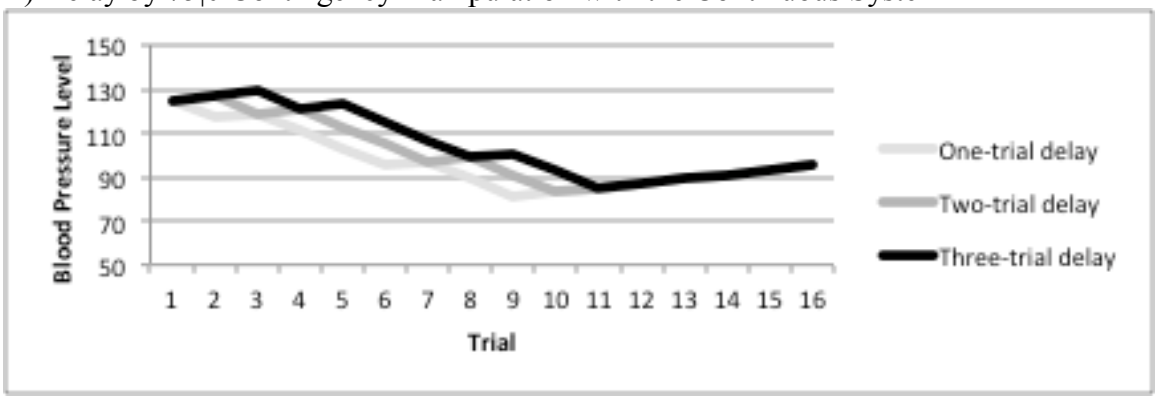


Figure F7. Illustration of blood pressure observations assuming a treatment intervention strategy of applying the treatment only on the first trial with the continuous system condition. Observations are separated by delay manipulations and the a) 25|0 contingency manipulation, b) 75|0 contingency manipulation, and c) 75|75 contingency manipulation. The observations are the same for the 25|0 and 75|0 contingency manipulations.

A) Delay by 25|0 Contingency Manipulation with the Continuous System



B) Delay by 75|0 Contingency Manipulation with the Continuous System



C) Delay by 75|75 Contingency Manipulation with the Continuous System

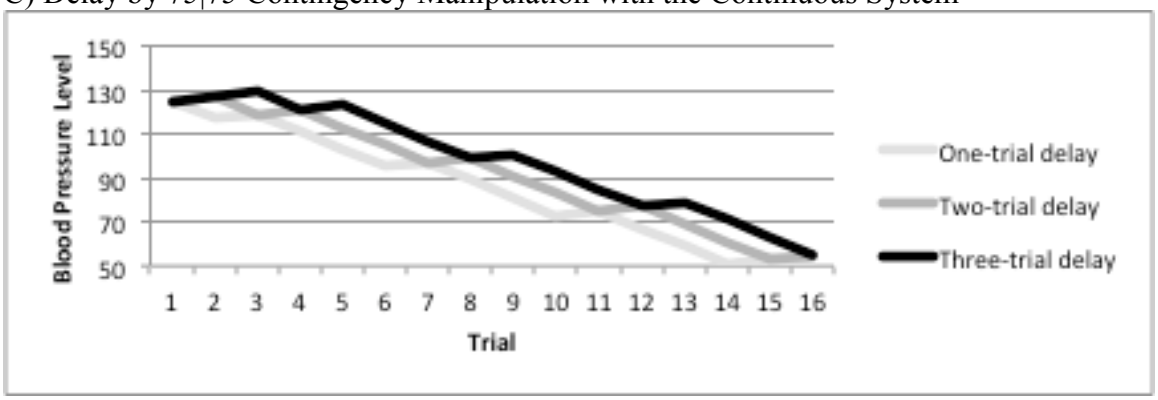
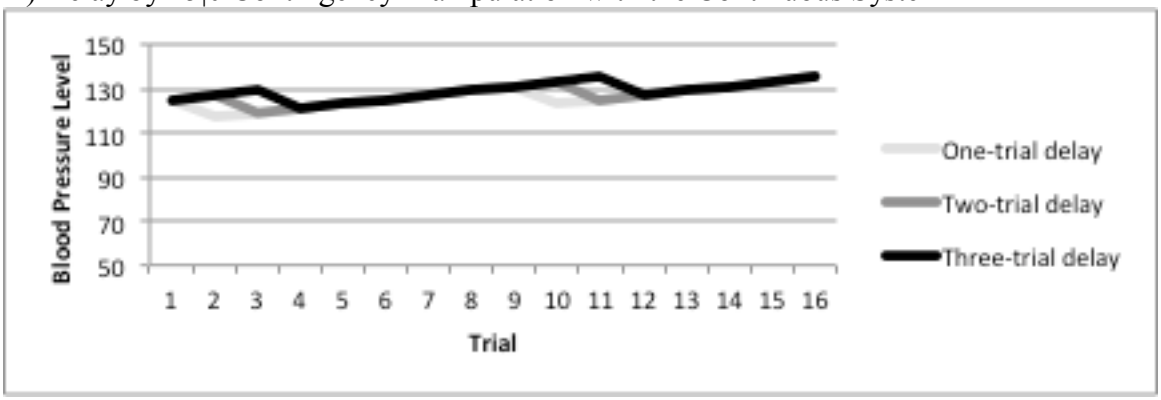
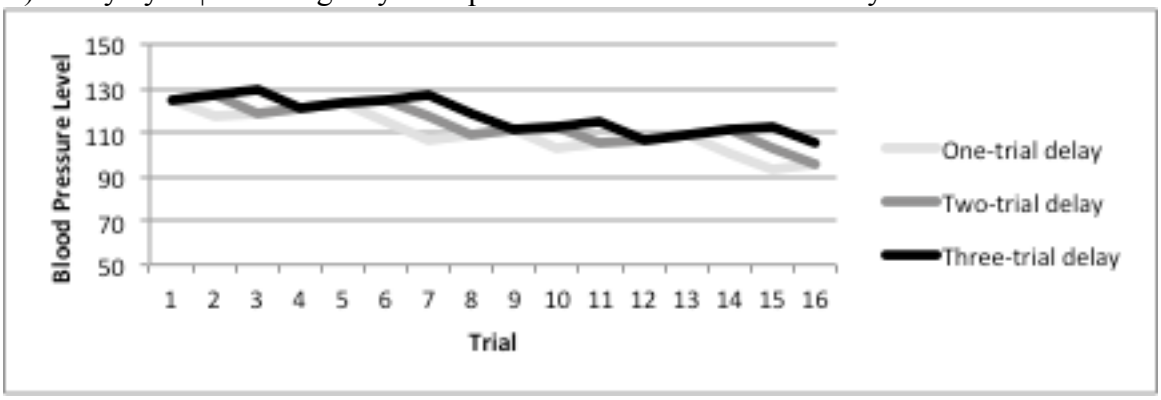


Figure F8. Illustration of blood pressure observations assuming a treatment intervention strategy of applying the treatment on trials 1-8 with the continuous system condition. Observations are separated by delay manipulations and the a) 25|0 contingency manipulation, b) 75|0 contingency manipulation, and c) 75|75 contingency manipulation.

A) Delay by 25|0 Contingency Manipulation with the Continuous System



B) Delay by 75|0 Contingency Manipulation with the Continuous System



C) Delay by 75|75 Contingency Manipulation with the Continuous System

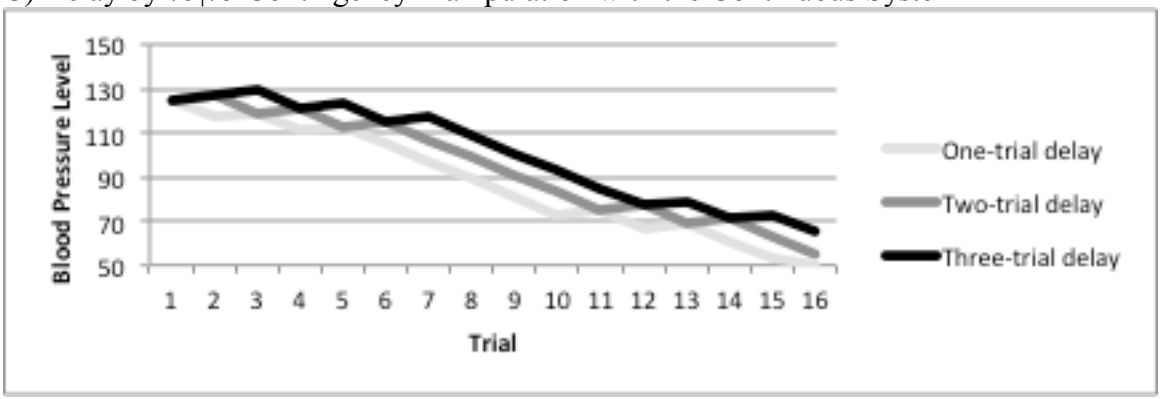
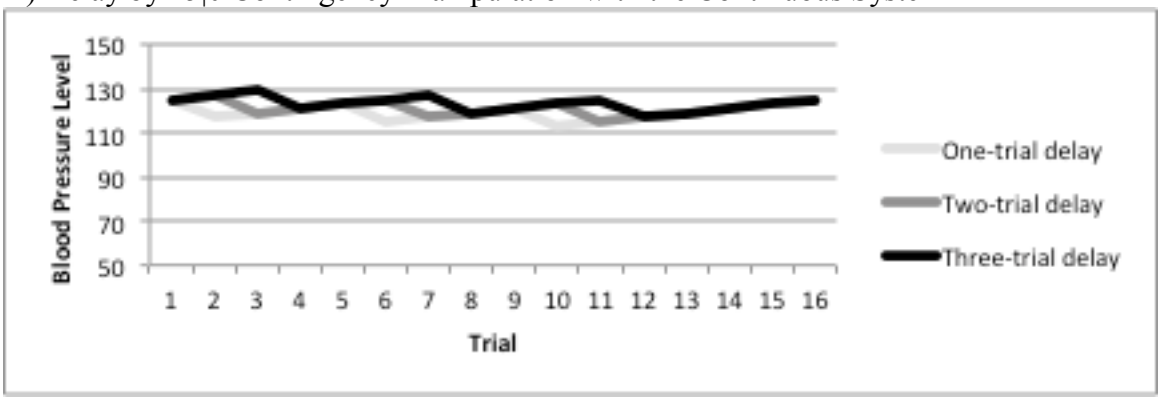
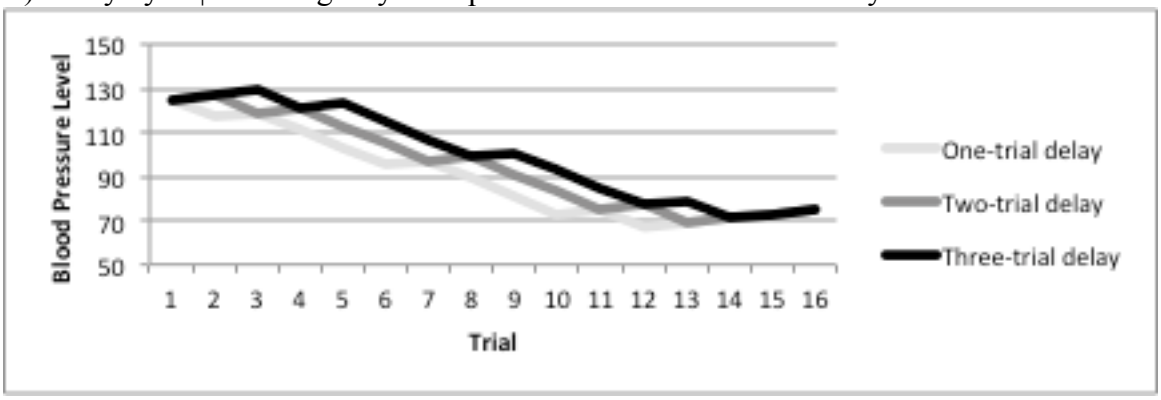


Figure F9. Illustration of blood pressure observations assuming a treatment intervention strategy that alternates applying two treatment interventions with withholding the treatment on two trials with the continuous system condition. Observations are separated by delay manipulations and the a) 25|0 contingency manipulation, b) 75|0 contingency manipulation, and c) 75|75 contingency manipulation.

A) Delay by 25|0 Contingency Manipulation with the Continuous System



B) Delay by 75|0 Contingency Manipulation with the Continuous System



C) Delay by 75|75 Contingency Manipulation with the Continuous System

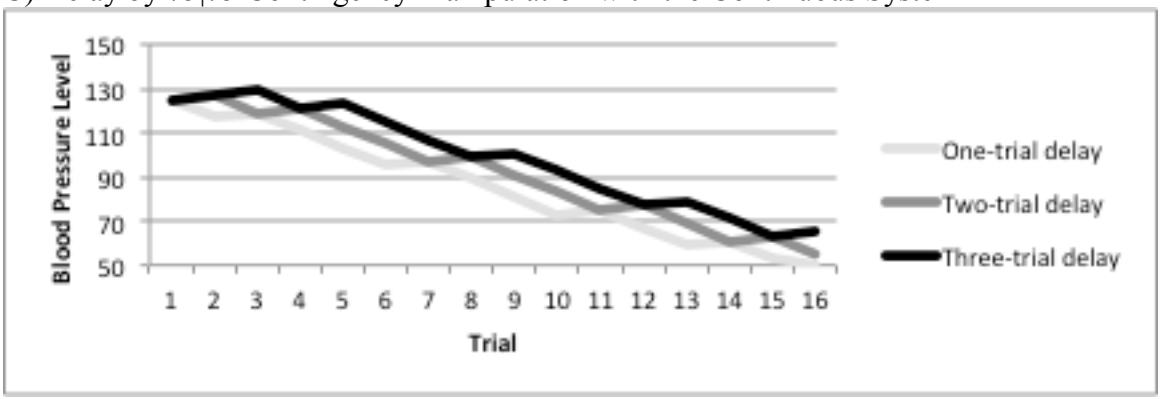


Figure F10. Illustration of blood pressure observations assuming a treatment-biased intervention strategy of applying the treatment on the first 11 trials and withholding the treatment for the remaining trials with the continuous system condition. Observations are separated by delay manipulations and the a) 25|0 contingency manipulation, b) 75|0 contingency manipulation, and c) 75|75 contingency manipulation.

APPENDIX G

Specification of the Effect to Null Effect Gain Ratio

The gain or weight associated with treatment interventions determines the speed at which a desired output (i.e., target blood pressure level) can be reached. More time and more apply treatment interventions are required to reach a target blood pressure level with lesser amounts of gain, but too much gain can result in less accuracy and oscillations around the target blood pressure threshold.

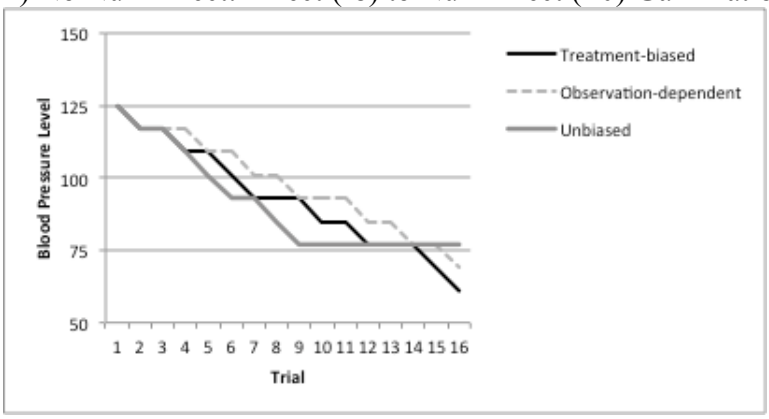
The influence of the gain parameter is a contextual factor that differs based on the nature of the system. With discrete system experiments, the system's output (e.g., blood pressure, object presence, happiness) is generally binary in nature (high or low, triangle or no triangle, smiley face or frown-y face) with the system returning to the baseline state after each trial (high blood pressure, no triangle, frown-y face). With such systems, a gain value is only associated with the presence of the effect (low blood pressure, triangle, smiley face) because the value associated with a null effect is the same as the baseline state (i.e., zero). Further, the specific value assigned to gain is meaningless given the desired output is either reached or not reached on each trial as a function of the binary outcome. In contrast, values associated with gain and the relationship between these values (in the case of more than intervention selection or type of output) is meaningful with continuous system experiments because a) the system's output is relational rather than binary, meaning the gain associated with interventions alters the system's output relative to the previous state rather than the baseline state, and b) the goal of reaching

and/or maintaining the target threshold extends across trials (absence of trial independence as a restriction).

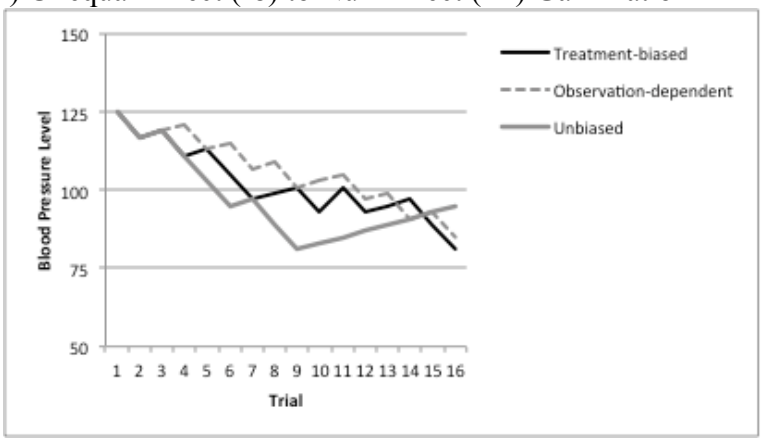
I examined a series of scenarios to appreciate how the specification of gain and the resulting ratios between the gain values assigned to the effect and null effect observations may affect on peoples' ability to control the system. I considered the same treatment intervention distributions that I used in Appendix C discussing observation dependence (that differ according to the strategic dimensions of observation dependence and intervention-bias) across three different effect to null effect ratios (no-null effect, unequal, and equal) that held the gain value associated with the effect constant. The no-null effect gain ratio (-8:0) is the most representative of the discrete system implementation, but is rather unrealistic relative to our scenario of hypertension management because it is characteristic of a temporary/curable condition (e.g., a headache), rather than a chronic management scenario. The unequal (-8: +2) and equal (-8: +8) gain ratios are more representative of a chronic management scenario because there is a cost associated with withholding the treatment. Figure G1 compares a treatment-biased and an unbiased, observation-independent intervention strategies, alongside an observation-dependent intervention strategy on blood pressure levels with a) the no null effect, b) the unequal, and c) the equal gain ratios with a 75|0 contingency by one-trial delay condition.[19]

[19] With the observation-dependent strategy, I applied and withheld the treatment when the blood pressure was elevated and normal respectively. This resulted in nine applications of the treatment. With the treatment-biased strategy, I applied 12 applications of the treatment with three withhold applications evenly interspersed.

A) No Null Effect: Effect (-8) to Null Effect (+0) Gain Ratio[20]



B) Unequal Effect (-8) to Null Effect (+2) Gain Ratio[21]



Finally, I continuously applied the treatment for eight trials with the unbiased strategy and withheld the treatment for the remaining trials. Across the scenarios, I applied a 75|0 contingency value (*P(effect|treatment)* = .75, *P(effect|~treatment)* = .00) and the outcome of each intervention occurred on the subsequent trial (one-trial delay manipulation).

[20] Blood pressure level on trial 16 for the no null- effect (-8) to null-effect (+0) gain ratio are 77, 69, and 61 for the unbiased, observation-dependent, and treatment-biased strategies respectively

[21] Blood pressure level on trial 16 for the unequal effect (-8) to null-effect (+2) gain ratio are 95, 85, and 81 for the unbiased, observation-dependent, and treatment-biased strategies respectively

C) Equal Effect (-8) to Null Effect (+8) Gain Ratio[22]

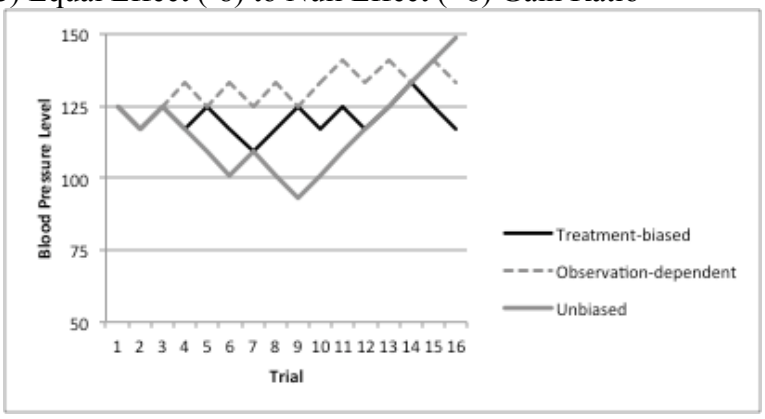


Figure G1. The relationship between the effect to null effect gain ratio and treatment intervention strategy on the ability to patients' blood pressure levels with a 75|0 contingency by one-trial delay condition

With all of the effect to null-effect gain ratios, the treatment-biased intervention strategy best supports the goal of lowering the patients' blood pressure followed by the observation-dependent and unbiased strategies due to the total number of treatments applied (12, 9, 8 respectively) across the 16 trials and the underlying contingency condition (75|0). This outcome corresponds to the % Normal values discussed relative to observation dependence in Appendix C. Also notable is the relationship between the effect to null-effect gain ratio and the decline in blood pressure level relative to the starting point, as well as differences attributable to intervention strategy. The equal gain ratio (+8:-8) generates the most variation in blood pressure observations between the three treatment intervention strategies. Also, the equal gain ratio may encourage the use of treatment-biased intervention strategies (at least with a contingency condition more

[22] Blood pressure level on trial 16 for the equal effect (-8) to null-effect (+8) gain ratio are 149, 133, and 117 for the unbiased, observation-dependent, and treatment-biased strategies respectively.

susceptible to strategic differences as with the 75|0 contingency condition), as this was the only treatment intervention strategy that resulted in a final system output on the 16th trial that was lower than the elevated starting point. In contrast, the no-null effect gain ratio (+8:0) supports an overall reduction in the blood pressure level relative to the elevated starting point across the treatment intervention strategies and there are minimal differences in blood pressure observations between the intervention strategies. However, it is not realistic that one application of a treatment will permanently affect a person's blood pressure given the context of hypertension management. The unequal gain ratio (+8:-2) offers a compromise in that there is a cost associated with withholding the treatment, yet it allows the blood pressure observation on the 16th trial to be reduced relative to the starting point and minimizes the impact of strategic differences.

APPENDIX H

Descriptions of Additional Dependent Measures

Table H1 contains descriptions of dependent measures not reviewed in the Method section. These include selection diversity, standard error, selections based on observations, selections based on changes in observations, selections based on unexpected changes, observation on the 16th trial, and response times. Results pertaining to these measures are included in subsequent appendices, but not reported in the body of the document.

Table H1
Descriptions of Additional Dependent Measures

Measure	Description
Selection diversity	I calculated participants' selections on a return matrix, which maps their selections at a particular trial (t) in relation to their selection on the subsequent ($t+1$). Given the two treatment interventions for participants to select within the learning task (apply and withhold), this matrix contained 4 cells. I used quantitative calculations that take into account the rate at which each cell was selected (e.g., Shannon-Weiner measure of information) to determine participants' overall selection diversity. • If participants selections are equally dispersed between the four possible combinations of t by $t+1$ options (apply apply, apply withhold, withhold apply, withhold withhold), then $H(t, t+1) = 2$. • If participants selections are less representative of equal selection dispersion across the four options, the value of $H(t, t+1)$ will decline. Finally, when participants only opt to apply the treatment (apply apply) or only opt to withhold the treatment (withhold withhold), then $H(t, t$

	+1) will be equal to zero.[23]
Standard error	Measure of the likelihood that participants will be able to identify the causality of the treatment options. Standard of error calculations take into account the number of trials that participants each of the treatment options relative to the actual contingency values associated with the task environment. Intervention selection strategies associated with lower standard of error values should facilitate participants' in correctly ascertaining the causality Table B1, in Appendix B depicts a set of selections strategies relative to the contingency conditions. o With the 75|0 and 25|0 contingency conditions, a strategy that is bias toward applying the treatment should assist participants in determining that the treatment variable lowers the patient's blood pressure level. However, with the 75|75 condition, a biased strategy in either direction leads to greater standard error values. The lowest standard error value, or the strategy with which the participant is most likely to realize the causality of (or lack of causality) the 75|75 condition is when an equal number of apply and do not apply treatment trials are considered (i.e., 8 trials).
Selections based on observations	Measure of the frequency that a participant selected a specific intervention (apply or withhold) relative the type of blood pressure observation (normal and elevated blood pressure). It is important to note that both participants' selections and the causal contingency manipulation influenced the probability of blood pressure observations. Specifically, the more participants chose to withhold the treatment relative to the 75|0, and 25|0 causal contingency conditions, the less likely they were to observe the patient's blood pressure level in the normal range. Also, with the continuous system condition, it was not feasible for participants to exit the elevated range with the 25|0 contingency condition regardless of intervention selection strategy.
Selections based on changes in observations	Measure considers intervention selections in relation to changes in blood pressure observations. A change in blood pressure was defined by comparing the observation on the current trial (*t*) relative to the previous trial (*t-1*). In addition to establishing whether a change occurred, I considered the direction of the change (increase or decrease in blood pressure observation). So, the measure considered the

[23] Participants only opting to withhold the treatment were eliminated from the final data set, so an $H(t, t+1) = 0$ relative to this condition will not occur.

	frequency of treatment interventions in relation to increases in blood pressure level ($t\text{-}(t\text{-}1) > 0$), decreases in blood pressure level (($t\text{-}(t\text{-}1) < 0$), and situations in which there was no change in blood pressure level ($t\text{-}(t\text{-}1) = 0$). There was always a change in blood pressure level with the continuous system, so neither the frequency calculation, nor analysis related to no change in blood pressure is applicable.
Change in treatment selection based on unexpected blood pressure observations	Measure of the frequency that participants switched their intervention selections when encountering an unexpected blood pressure observation. For this measure, I assumed that the expected observation was for the patient's blood pressure level to decrease and increase on the next trial with the application of and withholding of the treatment intervention respectively. With this assumption, the number of unexpected observations increased in combination with increases in the delay interval. The overall number of unexpected observations was also influenced by the causal contingency manipulation. Given the low *P(effect|treatment)* value, the 25|0 contingency manipulation produces more unexpected observations than the 75|0 contingency manipulation for expectations related to applying treatment option. Also, the 75|75 contingency manipulation produces more unexpected observations in relation to the withhold treatment option. To calculate the frequency value used for this measure, I summed the number of times participants changed their treatment selection (i.e., shifted from apply to withhold or from withhold to apply) that occurred in combination with an unexpected observation and divided this number by the total number of unexpected observations.
Value on the 16^{th} trial	The blood pressure observation shown to participants on the 16^{th} trial. For the discrete system condition, this value was either normal or elevated. For the continuous system condition, this value was the product of their cumulative set of treatment intervention selections.
Response times	The application recorded the amount of time (in seconds) between the 'next' button selections, which were used to advance through the learning and judgment tasks. • Learning task: measure considered total learning time and average trial time for each participant. • Judgment task: considered the amount of time participants spent determining if the treatment raised, lowered, or had no effect on the patient's condition.

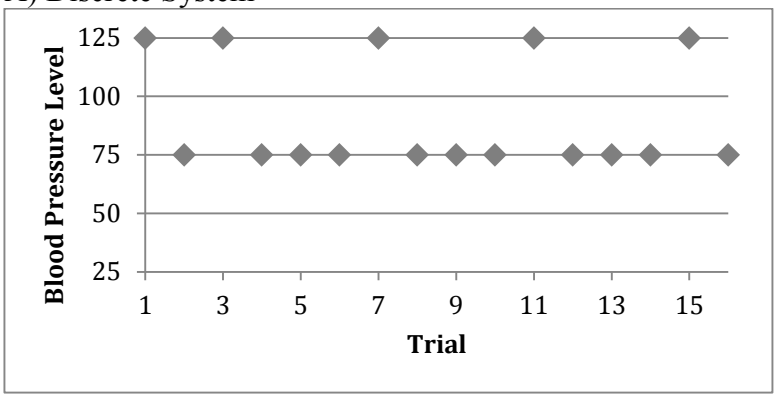
APPENDIX I

Assessment of System Control Options

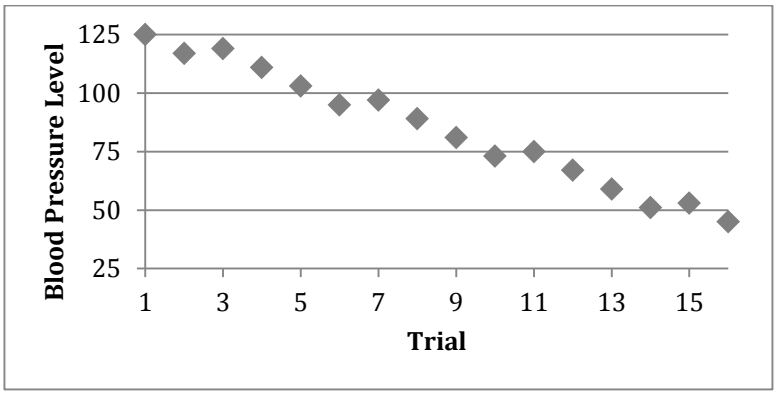
Two system control measures that are applicable to hypertension management are the frequency in a particular range (how often blood pressure is normal vs. elevated) or amount of directional change (how often does the blood pressure decline as opposed to elevate). To better understand these measures, I graphed the blood pressure observations for each system condition with the 75|0 contingency and one-trial delay manipulations with treatments applied on all trials (shown in Figure I1); the null-effect occurs on trials 3, 7, 11, and 15 with both system conditions.

With this example, values above and below 100 indicate elevated and normal blood pressure levels. The underlying contingency, and the effect to null effect gain ratio and gain specifications, affect the frequency that the blood pressure observation falls within the normal range with the continuous system. As a result, the use of this measurement may produce an inconsistent outcome between the system conditions. Figure I1 exemplifies this. Although 11 observations are in the normal range with both the discrete system (A) and continuous system with the -8:+2 effect to null effect gain ratio (B), a modification to this ratio results in only 4 observations in the normal range (C).

A) Discrete System



B) Continuous System – Effect to Null Effect Gain Ratio of -8:+2



C) Continuous System – Effect to Null Effect Gain Ratio of -5:+5

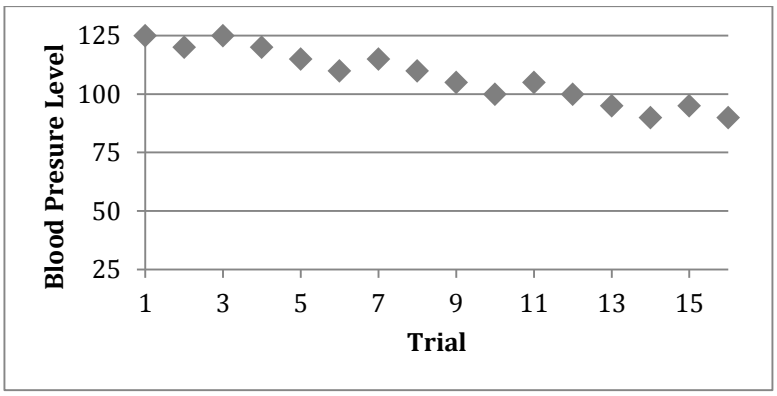


Figure 11. Blood pressure observations associated with the discrete and continuous systems with the 75|0 contingency manipulation and two gain ratios for the continuous system. The treatment is applied on every trial. Values above and below 100 indicate elevated and normal blood pressure levels.

Alternative system control measures are the frequency that blood pressure declines and the cumulative output of a series of selections. These measures are analogous; the only difference is whether the output value is presented as a percentage or in the units of the display scale. The outcome state of the display scale is representative of cumulative impact of the treatment interventions on every trial with the continuous system. This inherent counter is not true of the binomial output (elevated/normal) of the discrete system. However, participants' treatment interventions with the discrete system can be converted to the numeric scale of the continuous system to support system comparison.

APPENDIX J

Demographic and Debrief Questionnaire

Q1 Type in the code on written on your participant card

Q2 For your last patient, how difficult or easy was it to identify how the treatment influenced their blood pressure level?
- ❍ Very Difficult (1)
- ❍ Difficult (2)
- ❍ Somewhat Difficult (3)
- ❍ Neutral (4)
- ❍ Somewhat Easy (5)
- ❍ Easy (6)
- ❍ Very Easy (7)

Q3 For your last patient, what strategy did you use to determine how the treatment was influencing their blood pressure level?

Q4 Why did you use this strategy?

Q5 Did you use the same strategy for all of the patients?
- ❍ yes (1)
- ❍ no (2)

If Q5 is 'no,' then go to Q6, otherwise go to Q7

Q6 Explain how you modified your strategy across the three patients.

Q7 For your last patient, what was your expectation about how this treatment would work before you applied your treatment selections?

Q8 For your last patient, what did you think was happening?

Q9 If you thought the treatment influenced your last patient's blood pressure level, how much time did it take for the treatments to work (take effect)?

Q10 Generally, how much time do you think it takes for medical treatments to work (take effect)?

Q11 Generally, how much time do you think it takes for medical treatments to wear off?

Q12 For your last patient, how did you decide what assessment (or rating) to assign?

Q13 Generally, what percentage of treatment applications should influence a patient's condition for that treatment to be considered effective?

Q14 Doctors use too many medicines
- ❍ strongly disagree (1)
- ❍ disagree (2)
- ❍ uncertain (3)
- ❍ agree (4)
- ❍ strongly agree (5)

Q15 Natural remedies are safer than medicines
- ❍ strongly disagree (1)
- ❍ disagree (2)
- ❍ uncertain (3)
- ❍ agree (4)
- ❍ strongly agree (5)

Q16 Medicines do more harm than good
- ❍ strongly disagree (1)
- ❍ disagree (2)
- ❍ uncertain (3)
- ❍ agree (4)
- ❍ strongly agree (5)

Q17 People who take medicines should stop their treatment for a while every now and again
- ❍ strongly disagree (1)
- ❍ disagree (2)
- ❍ uncertain (3)
- ❍ agree (4)
- ❍ strongly agree (5)

Q18 Have you personally managed your own or someone else's medical condition?
- ❍ Yes (1)
- ❍ No (2)

If Q18 is 'Yes,' then complete Q19, Q20 and Q21, otherwise go to questions 24 to 29

Q19 Please describe the medical condition that you managed.

Q20 Where did you get information about this medical condition?

Q21 Did you take or give medication for this medical condition?
- ❍ Yes (1)
- ❍ No (2)

If Q21 is 'Yes,' then go to Q22, otherwise go to questions 24 to 29

Q22 Did you change the amount of medication given or taken by yourself?
- ❍ Yes (1)
- ❍ No (2)

If Q22 is 'Yes,' then go to Q23, otherwise go to questions 24 to 29

Q23 Please describe how you changed the amount of medication given or taken.

Q24 to Q29: The Qualtrics application randomly selected two of these six questions for each participant

Q24 the opposite of beautiful is the definition of which word?
- ❍ ugly (1)
- ❍ shy (2)
- ❍ blind (3)
- ❍ sensible (4)
- ❍ lazy (5)
- ❍ likely (6)

Q25 not able to see is the definition of which word?
- ❍ ugly (1)
- ❍ shy (2)
- ❍ blind (3)
- ❍ sensible (4)
- ❍ lazy (5)
- ❍ likely (6)

Q26 acting with good judgment is the definition of which word?
- ❍ ugly (1)
- ❍ shy (2)
- ❍ blind (3)
- ❍ sensible (4)
- ❍ lazy (5)
- ❍ likely (6)

Q27 probably going to happen is the definition of which word?
- ugly (1)
- shy (2)
- blind (3)
- sensible (4)
- lazy (5)
- likely (6)

Q28not wanting to work is the definition of which word?
- ugly (1)
- shy (2)
- blind (3)
- sensible (4)
- lazy (5)
- likely (6)

Q29 nervous of other people is the definition of which word?
- ugly (1)
- shy (2)
- blind (3)
- sensible (4)
- lazy (5)
- likely (6)

Q30 to Q35: The Qualtrics application randomly selected two of these six questions for each participant

Q30 someone who has been affected by something like a crime, accident or illness is the definition of which word?
- victim (1)
- litter (3)
- reward (4)
- courage (5)
- plug (6)
- currency (7)

Q31 rubbish is the definition of which word?
- victim (1)
- litter (3)
- reward (4)
- courage (5)
- plug (6)
- currency (7)

Q32 something good that you get or that happens because of something you have done is the definition of which word?
- ❍ victim (1)
- ❍ litter (3)
- ❍ reward (4)
- ❍ courage (5)
- ❍ plug (6)
- ❍ currency (7)

Q33 the ability to do something dangerous, frightening or difficult is the definition of which word?
- ❍ victim (1)
- ❍ litter (3)
- ❍ reward (4)
- ❍ courage (5)
- ❍ plug (6)
- ❍ currency (7)

Q34 an object used for connecting a machine to the electricity supply is the definition of which word?
- ❍ victim (1)
- ❍ litter (3)
- ❍ reward (4)
- ❍ courage (5)
- ❍ plug (6)
- ❍ currency (7)

Q35 the money used in a particular country is the definition of which word?
- ❍ victim (1)
- ❍ litter (3)
- ❍ reward (4)
- ❍ courage (5)
- ❍ plug (6)
- ❍ currency (7)

Q36 How old are you?

Q37 Are you a male or female?
- ❍ Male (1)
- ❍ Female (2)

Q38 What is your grade level?
- ❍ Freshman (1)
- ❍ Sophomore (2)
- ❍ Junior (3)
- ❍ Senior (4)

Q39 Is English your native language (the first language that you learned)?
- ❍ Yes (1)
- ❍ No (2)

If Q31 is 'No,' then complete Q40 and Q41, otherwise go to Q42

Q40 What is your native language?

Q41 Describe your ability to understand written content (i.e., read) in English.
- ❍ Poor (1)
- ❍ Fair (2)
- ❍ Good (3)
- ❍ Very Good (4)
- ❍ Excellent (5)

Q42 What college or school are you affiliated with?
- ❍ Raj Soin College of Business (1)
- ❍ College of Education and Human Services (2)
- ❍ College of Engineering and Computer Science (3)
- ❍ College of Liberal Arts (4)
- ❍ College of Nursing and Health (7)
- ❍ College of Science and Mathematics (9)
- ❍ University College (10)

If Q42 is 'College of Education and Human Services,' then go to Q43

Q43 What is your major/program?
- ❑ Athletic Training (1)
- ❑ Career, Technical, Adult Education (2)
- ❑ Community Health Education (3)
- ❑ Early Childhood Education (4)
- ❑ Middle Childhood Education (5)
- ❑ Middle Childhood Intervention Specialist (6)
- ❑ Organizational Leadership (7)
- ❑ Physical Education (8)
- ❑ Rehabilitation Services (9)
- ❑ Sign Language Interpreting (10)
- ❑ Sports Science (11)
- ❑ Wellness Studies (12)
- ❑ Undecided (13)

If Q42 is 'College of Engineering and Computer Science,' then go to Q44

Q44 What is your major/program?

- ❑ Biomedical Engineering (1)
- ❑ Computer Engineering (2)
- ❑ Computer Science (3)
- ❑ Electrical Engineering (4)
- ❑ Engineering Physics (5)
- ❑ Industrial and Systems Engineering (6)
- ❑ Material Science Engineering (7)
- ❑ Mechanical Engineering (8)
- ❑ Undecided (9)

If Q42 is 'College of Liberal Arts Is Selected,' then go to Q45

Q45 What is your major/program?

- ❑ Acting (1)
- ❑ African and African/American Studies (2)
- ❑ Anthropology (3)
- ❑ Art History (4)
- ❑ Classical Humanities (5)
- ❑ Communication Studies (6)
- ❑ Crime and Justice Studies (7)
- ❑ Dance (8)
- ❑ Design/Technology (9)
- ❑ Economics (10)
- ❑ English (11)
- ❑ French (12)
- ❑ Geography (13)
- ❑ German (14)
- ❑ Greek (15)
- ❑ History (16)
- ❑ International Studies (17)
- ❑ Latin (18)
- ❑ Liberal Studies (19)
- ❑ Mass Communication (20)
- ❑ Motion Pictures (21)
- ❑ Music (22)
- ❑ Music Education (23)
- ❑ Music History and Literature (24)
- ❑ Performance (25)
- ❑ Philosophy (26)
- ❑ Political Science (27)
- ❑ Religion (28)
- ❑ Selected Studies (29)
- ❑ Social Science Education (30)
- ❑ Social Work (31)

- ❑ Sociology (32)
- ❑ Spanish (33)
- ❑ Studio Art (34)
- ❑ Theater Studies (35)
- ❑ Urban Affairs (36)
- ❑ Women's Studies (37)
- ❑ Undecided (38)

If Q42 is 'College of Science and Mathematics,' then go to Q46

Q46 What is your major/program?
- ❑ Biological Sciences (1)
- ❑ Chemistry (2)
- ❑ Clinical Laboratory Science (3)
- ❑ Earth and Environmental Sciences (4)
- ❑ Integrated Science (5)
- ❑ Mathematics (6)
- ❑ Physics (7)
- ❑ Psychology (8)
- ❑ Statistics (9)
- ❑ Undecided (10)

If Q42 is 'Raj Soin College of Business,' then go to Q47

Q47 What is your major/program?
- ❑ Accountancy (1)
- ❑ Business Economics (2)
- ❑ Finance (3)
- ❑ Financial Services (4)
- ❑ Human Resource Management (5)
- ❑ International Business (6)
- ❑ Management (7)
- ❑ Management Info Systems (8)
- ❑ Marketing (9)
- ❑ Supply Chain Management (10)
- ❑ Undecided (11)

Thank you for participating in this study. Your input will help us to better understand if and when people can correctly identify effective causal variables when interacting with complex systems. As we are particularly interested in the medical domain, your questionnaire feedback will provide us with insight into how people think about treatments within the medical domain.

Q48 Do you have any additional questions, comments, or concerns about this study?

Please submit your responses and let the study facilitator know that you are complete.

APPENDIX K

Research Consent Form

CONSENT TO PARTICIPATE IN RESEARCH

Department of Psychology
Wright State University

Title of study	**Identifying variables that cause an effect**
Consent to participate	This signed consent is to certify my willingness to participate in this study, which assesses my ability to identify variables that cause an effect. I am free to refuse to participate in this study or to withdraw at any time. I understand that I will receive 1 research credit hour for every 30 minutes of my participation in this study. If I decide not to participate, or to withdraw before completing the study, I am free to so do without penalty and I will still receive 1 research credit hour for every 30 minutes.
Purpose of research	To gain a better understanding of how and when people can correctly identify effective causal variables after interacting with a complex system.
Procedures	For this study, I will work individually on a computer to apply pretend medical treatments to hypothetical patients and assess whether or not these treatments affect their medical condition. I will also provide non-identifying demographic information and answer questions about my experience with the study. I expect that participation will take about 1 hour. I understand that the research team will electronically record the information that I provide during the study session.
Risks	There are no recognized risks in participating.
Confidentiality	I understand that any information that is obtained from this study will be kept strictly confidential. I understand that all collected materials will be stored in password protected files and will only be available to the researchers. I understand that I will NOT be identified in any report or publication.
Availability of results	A summary of the study results may be requested by contacting the researchers listed below. The summary will show only aggregated (i.e., combined) data for the entire sample. No individual results will be available. The results of this study will be available after July, 2014.
Questions or complaints	If I have questions about this study or have research-related complaints I can contact the lead researcher Beth Bullemer at bullemer.2@wright.edu or her adviser, Valerie L. Shalin, PhD at valerie.shalin@wright.edu. If I have general questions about giving consent or my rights as a research participant in the study, I can call the Wright State University Institutional Review Board at (937) 775-4462.

My signature below means that I have freely agreed to participate in this study. I will be given a copy of this consent form.

_____

Participant Date

_____

Participant Name (Printed)

_____

Investigator

APPENDIX L

Study Introduction

The figures presented below are screenshots of the introduction shown to participants at the onset of the study. Figure L1 depicts the screen that presented participants with an overview of the study. Figure L2 presents the screen used to introduce the blood pressure widget. Figures L3, L4, and L5 present the screens used to introduce the learning task and interface. Figures L6 and L7 depict the screens that introduced the judgment task. Lastly, Figure L8 shows the screen that presented participants with the task instructions and goals.

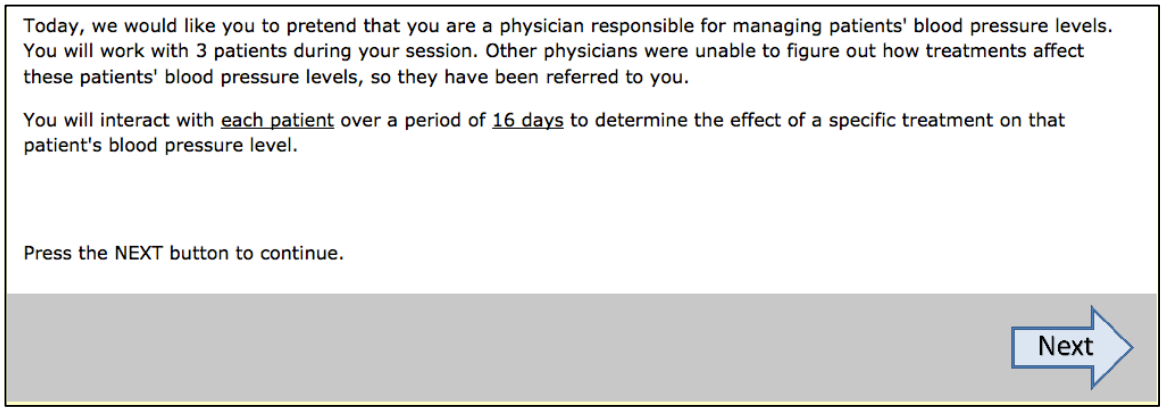


Figure L1. Study Overview

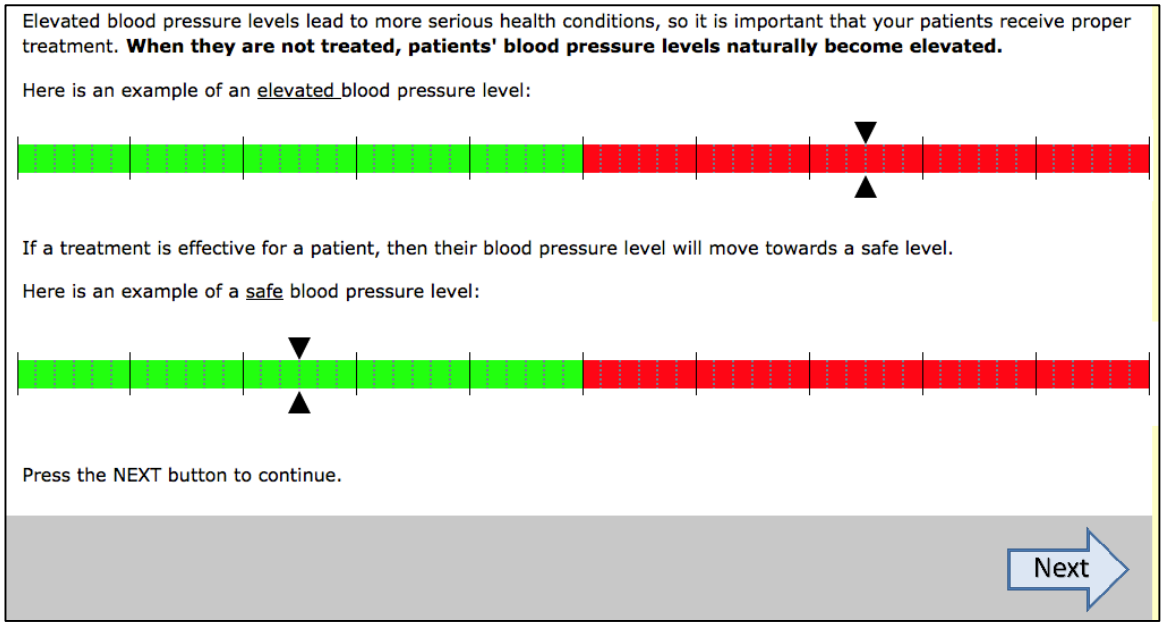


Figure L2. Description of the Blood Pressure (i.e., the effect) Widget

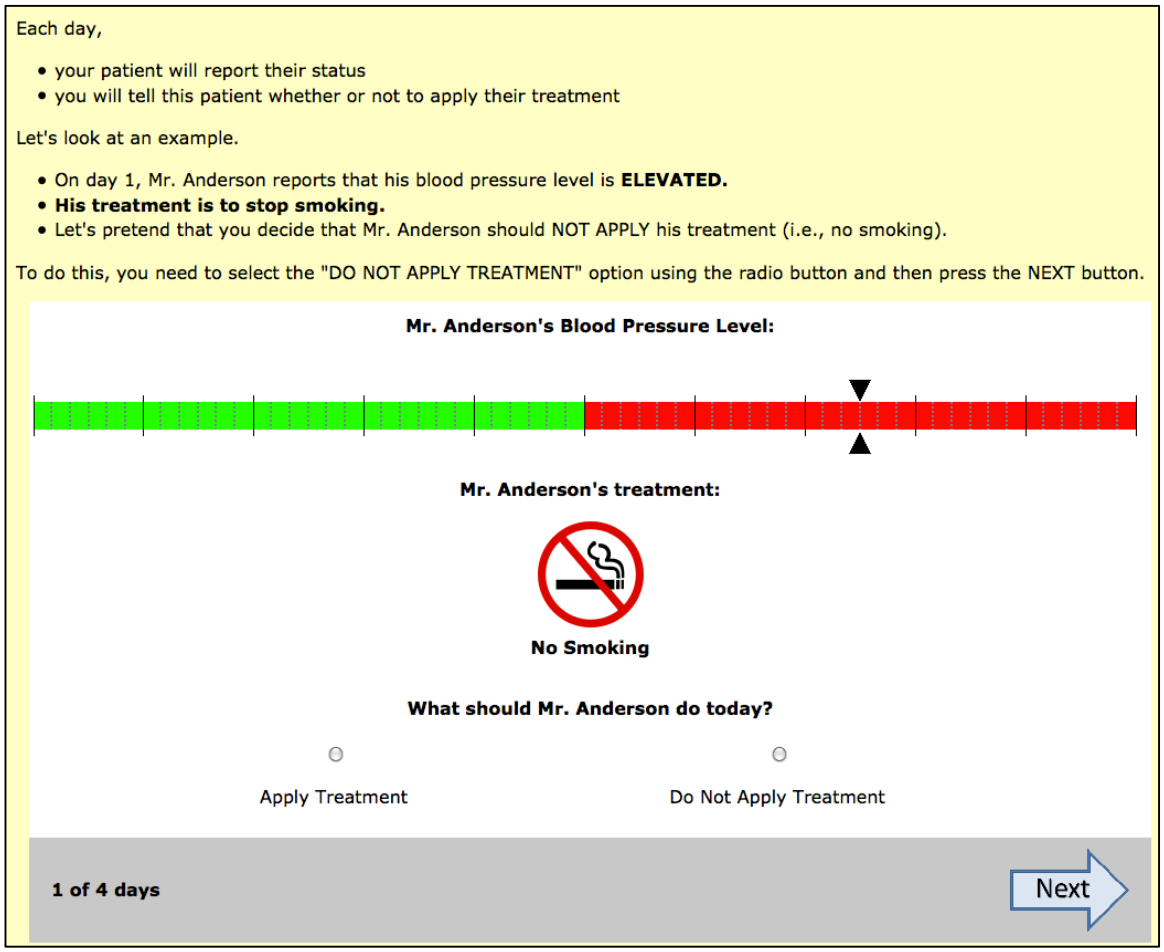


Figure L3. Introduction to the Learning Task and Interface

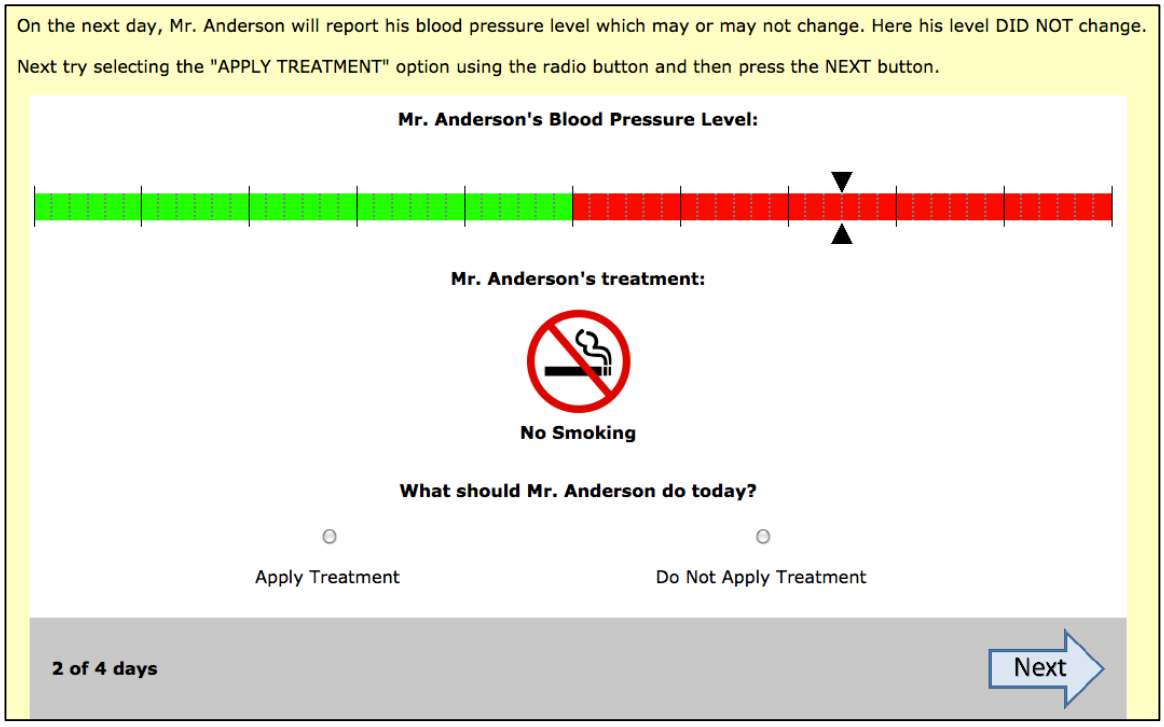


Figure L4. Introduction to the Learning Task and Interface continued

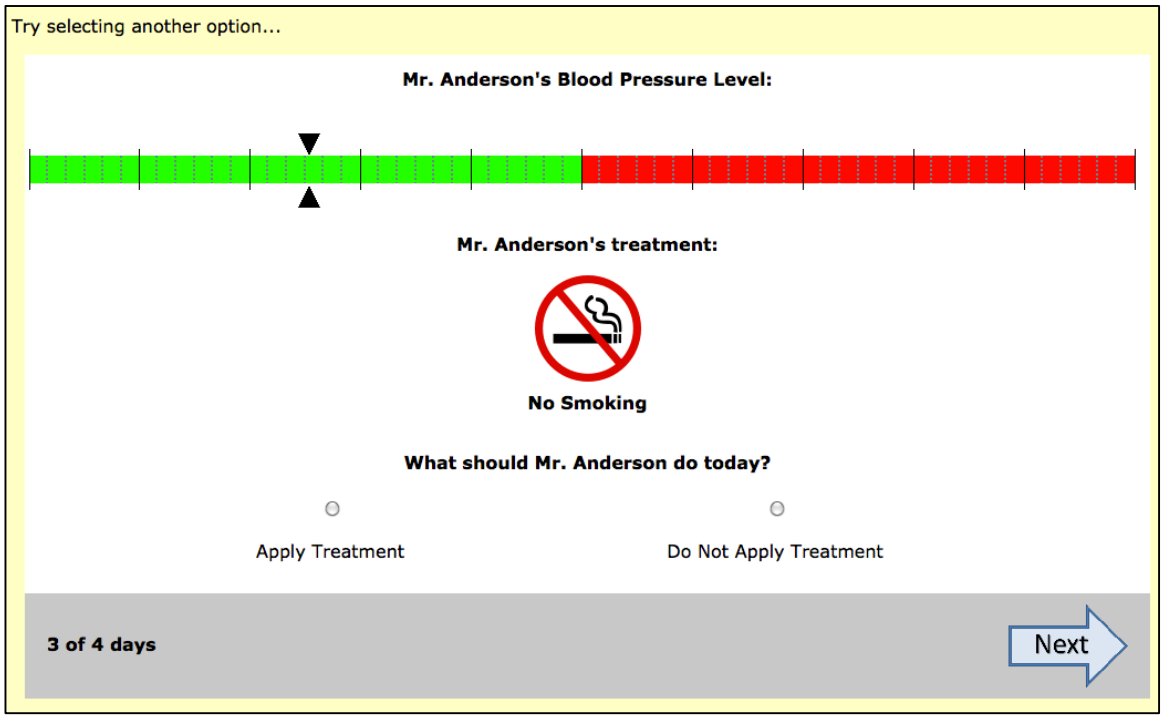


Figure L5. Introduction to the Learning Task and Interface continued

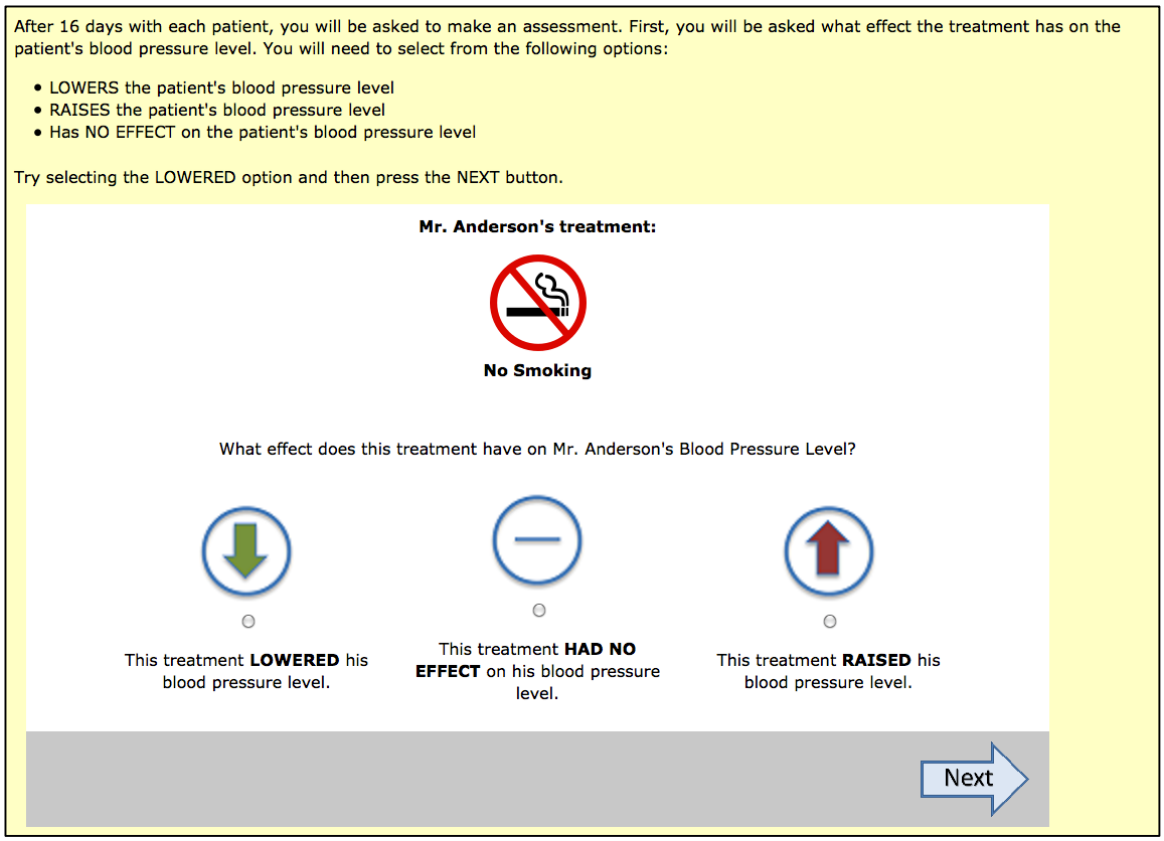


Figure L6. Introduction to the Judgment Task and Interface

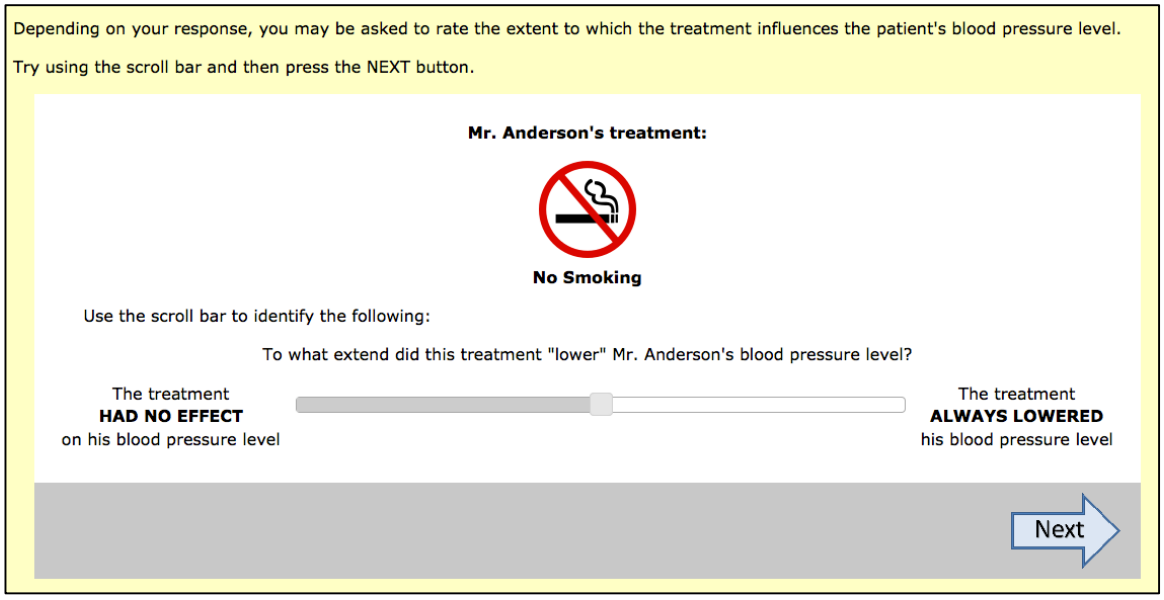


Figure L7. Introduction to the Judgment Task and Interface continued

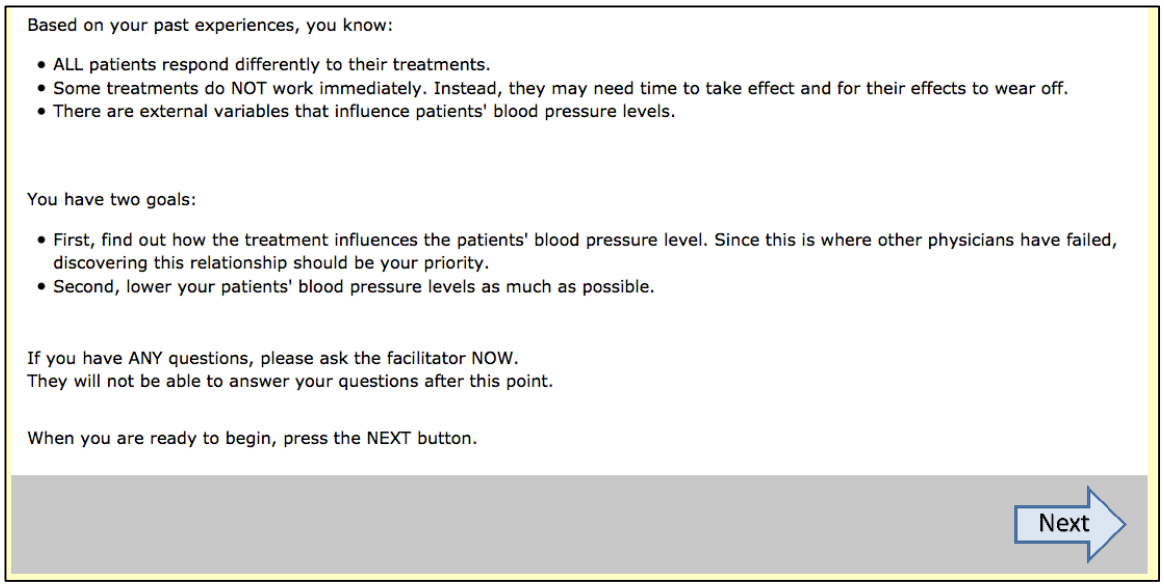


Figure L8. Task Instructions and Task Goals

APPENDIX M

Normality Assessments

Tables M1 and M5 contain the histograms and Q-Q plots of participants causal attributions separated by contingency. Tables M2 and M6 contain the histograms and Q-Q plots of the ability to control the system measure separated by contingency. Tables M3 and M7 contain the histograms and Q-Q plots related to treatment intervention selection measures. Tables M4 and M8 contain the histograms and Q-Q plots of response time data. Tables M1, M2, M3, and M4 graph data from the discrete system condition and Tables M5, M6, M7, and M8 graph data from the continuous system condition.

Discrete System

Table M1

Histograms and Q-Q Plots of Causal Attribution with the Discrete System Condition separated by Contingency

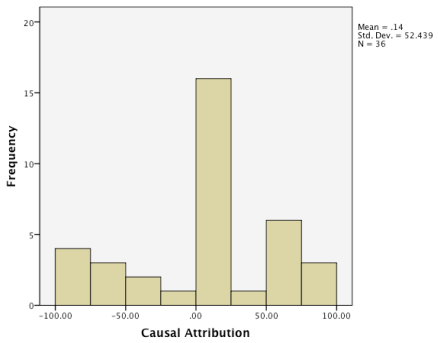
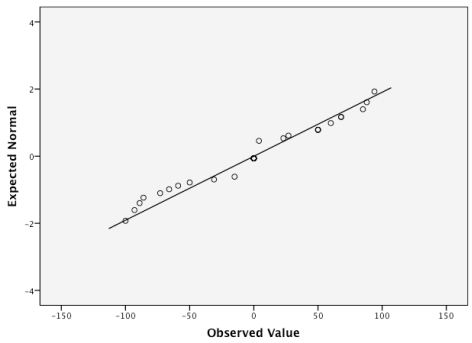
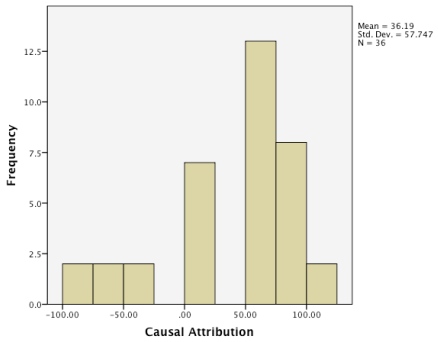
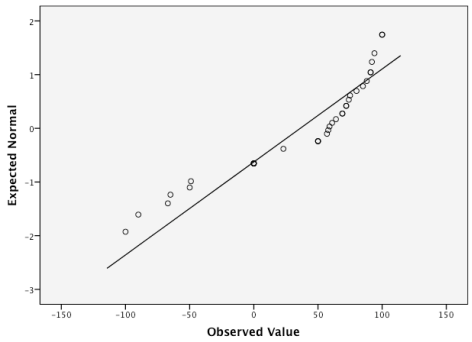
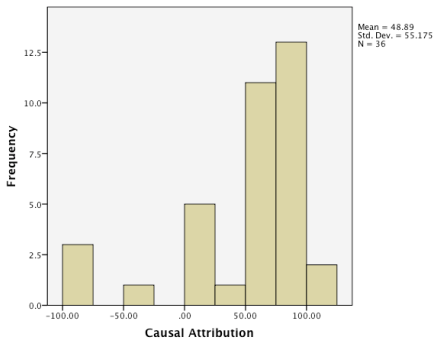
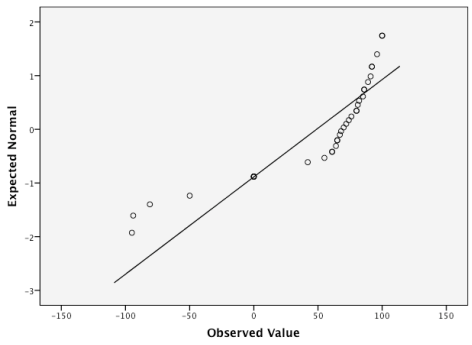
Histogram	Q-Q Plot
A) 25|0	
	
B) 75|0	
	
C) 75|75	
	

Table M2

Histograms and Q-Q Plots of Control of the System with the Discrete System Condition separated by Contingency

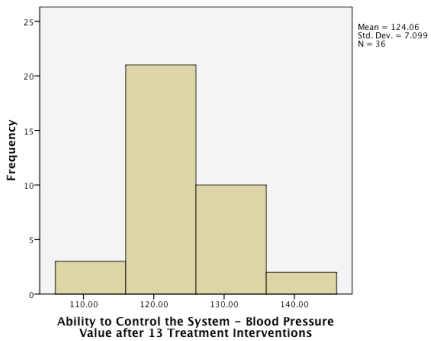
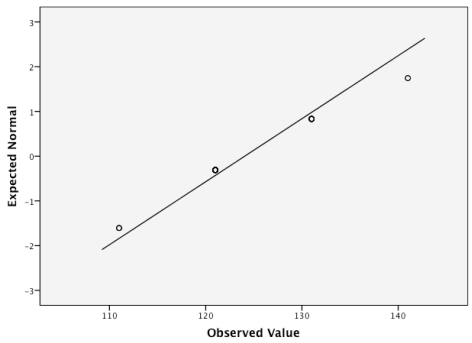
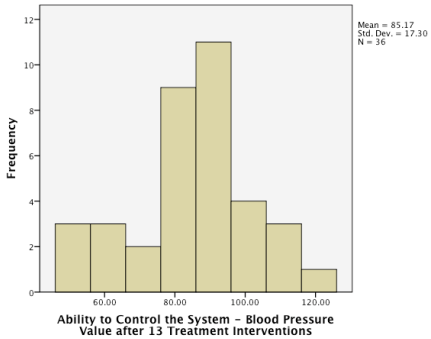
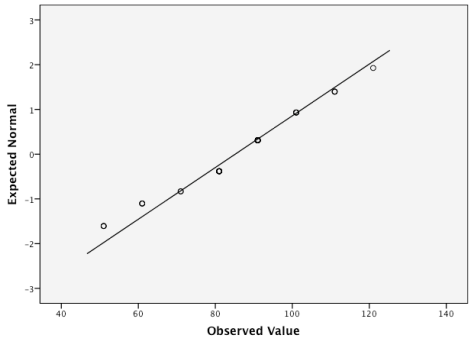
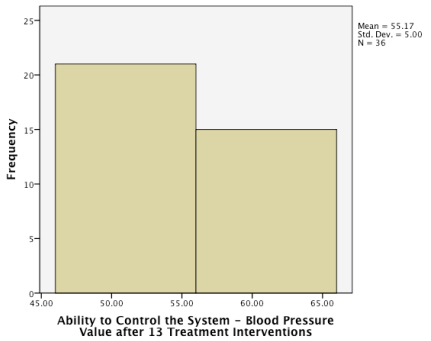
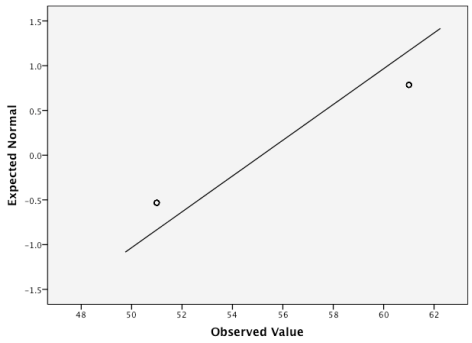
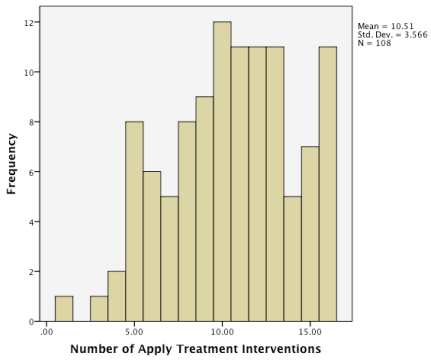
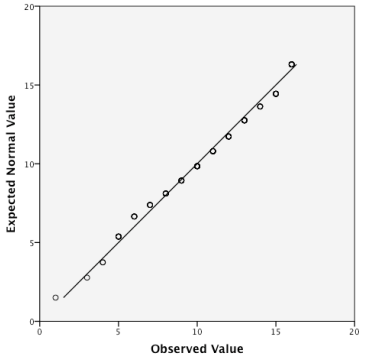
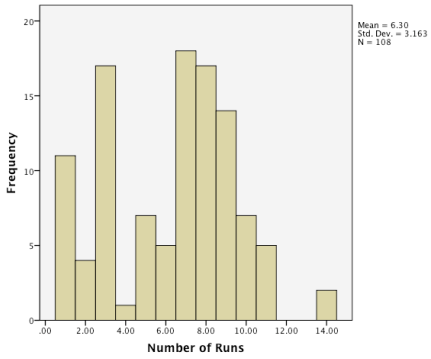
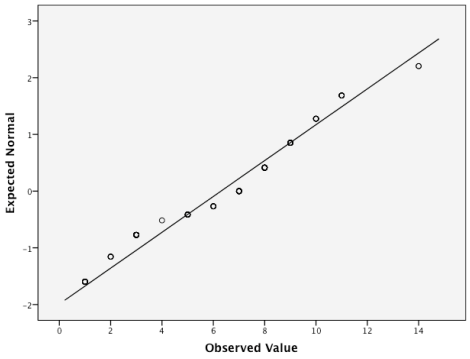
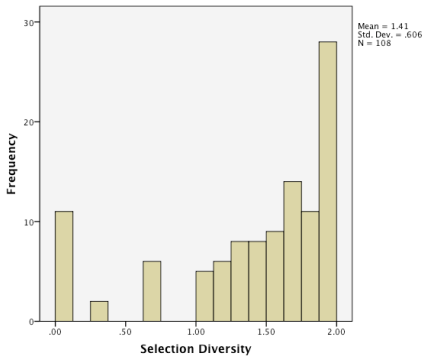
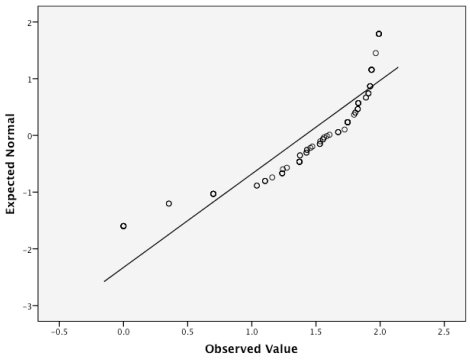
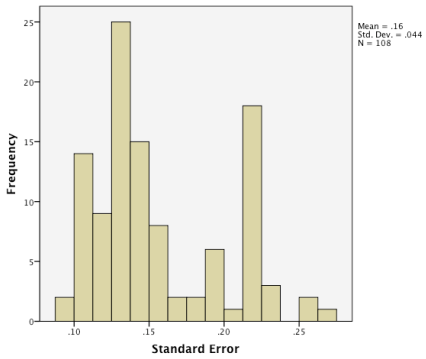
Histogram	Q-Q Plot
A) 25|0	
	
B) 75|0	
	
C) 75|75	
	

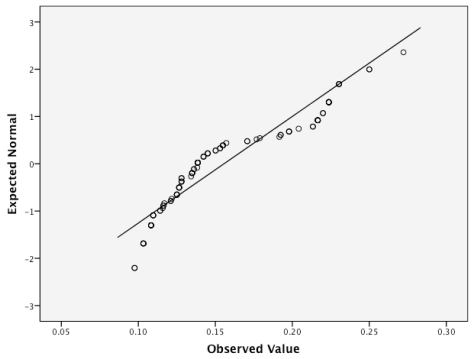
Table M3

Histograms and Q-Q Plots of Intervention Strategy Dependent Measures with the Discrete System Condition

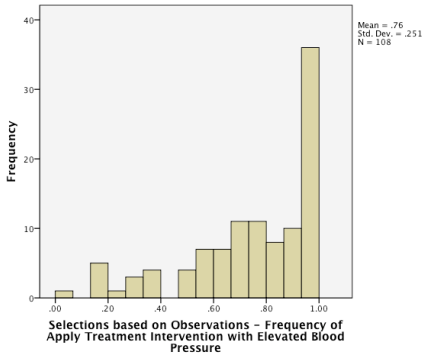
Histogram	Q-Q Plot
A) Number of Apply Treatment Interventions	
	
B) Number of Runs	
	
C) Selection Diversity	
	

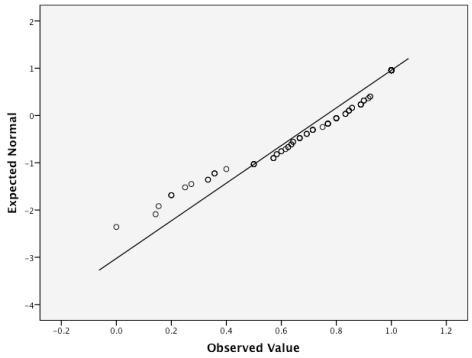
D) Standard Error



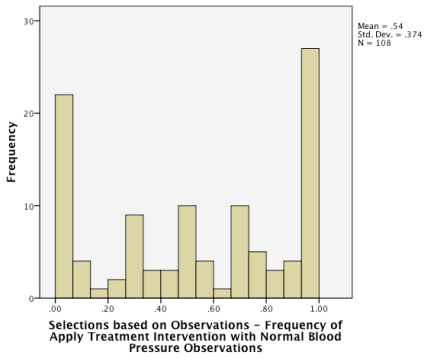


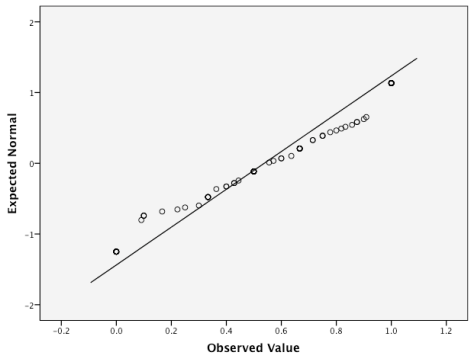
E) Selections based on Observations - Apply Treatment Intervention with Elevated Blood Pressure



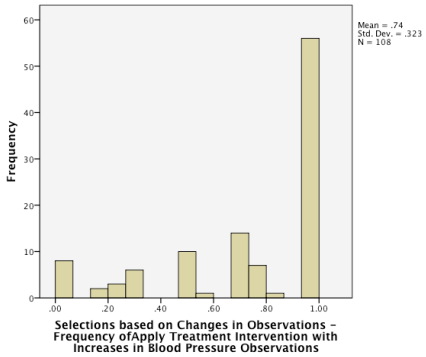


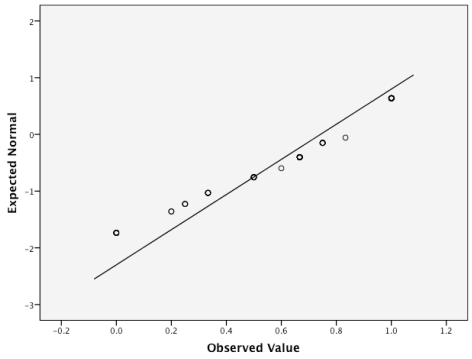
F) Selections based on Observations - Apply Treatment Intervention with Normal Blood Pressure Observations



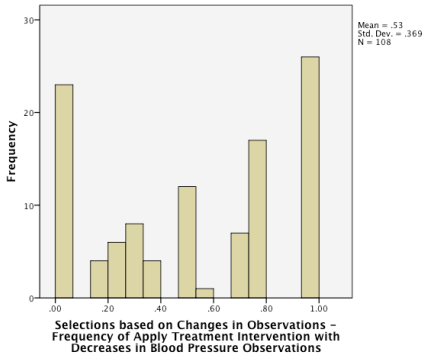


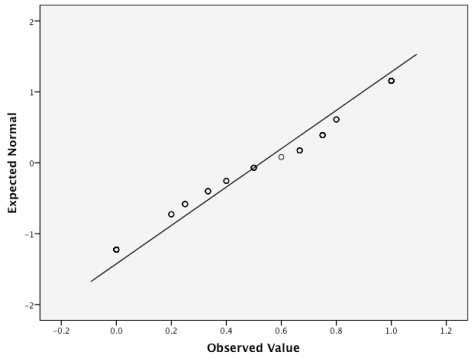
G) Selections based on Changes in Observations - Apply Treatment Intervention with Increases in Blood Pressure Observations



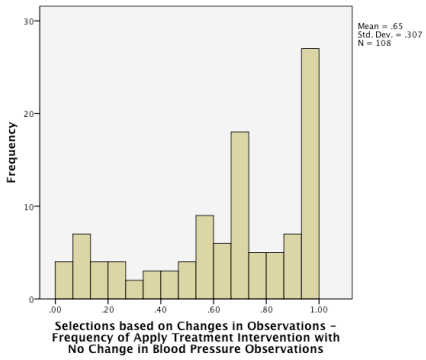


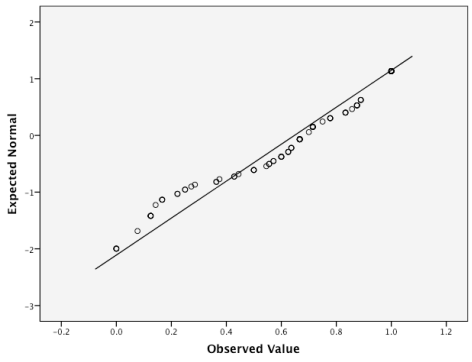
H) Selections based on Changes in Observations - Apply Treatment Intervention with Decreases in Blood Pressure Observations



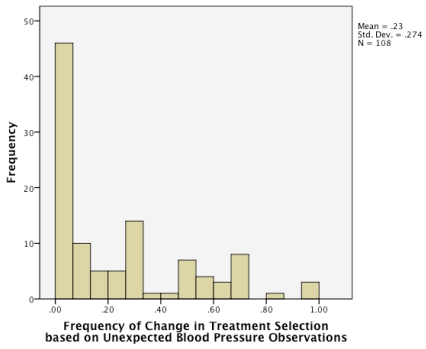


I) Selections based on Changes in Observations - Apply Treatment Intervention with No Change in Blood Pressure Observations





J) Change in Treatment Selection based on Unexpected Blood Pressure Observations



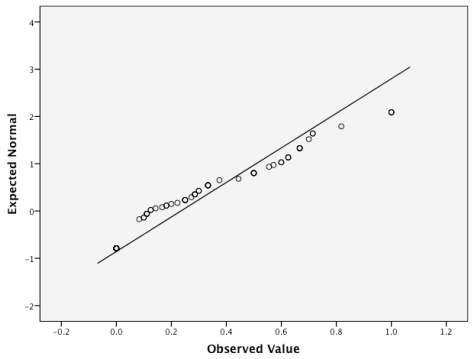
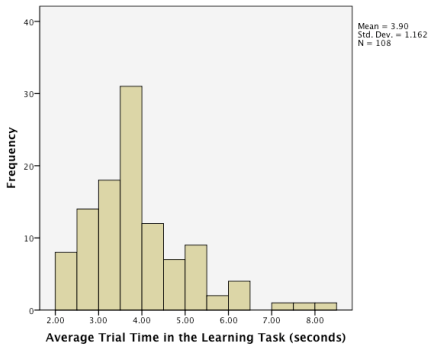


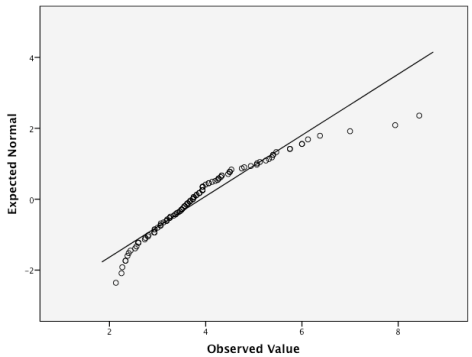
Table M4

Histograms and Q-Q Plots of Response Times with the Discrete System Condition

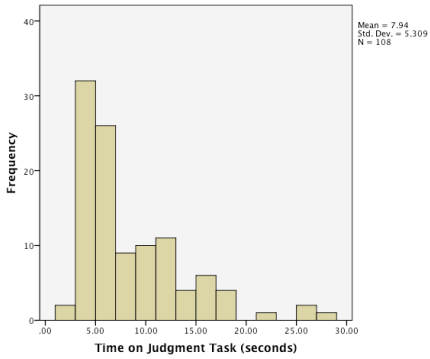
Histogram	Q-Q Plot

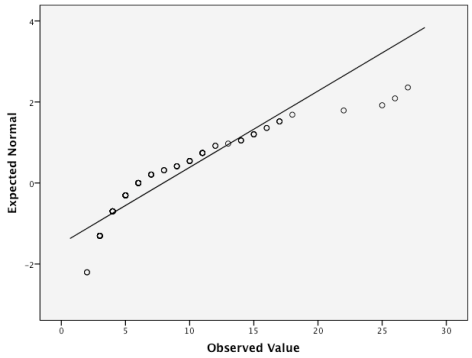
A) Average Trial Time in the Learning Task





B) Time on Judgment Task





Continuous System

Table M5

Histograms and Q-Q Plots of Causal Attribution with the Continuous System Condition separated by Contingency

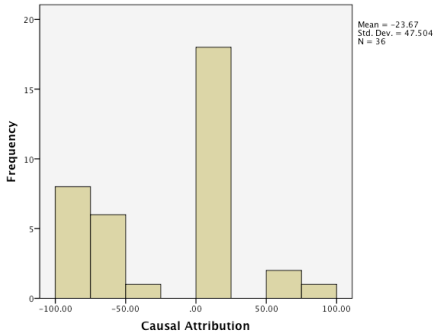
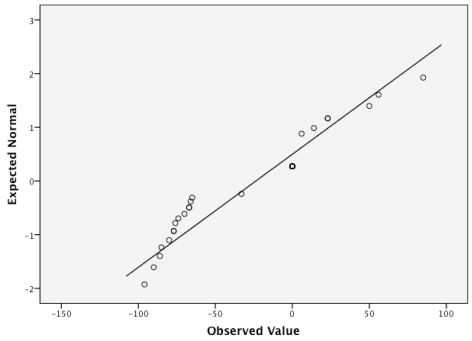
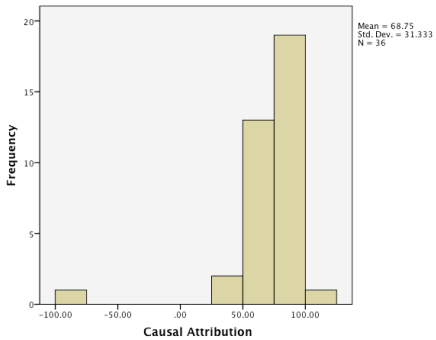
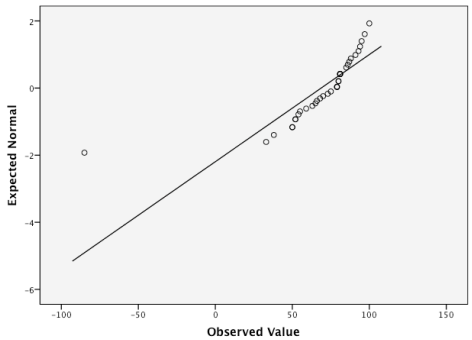
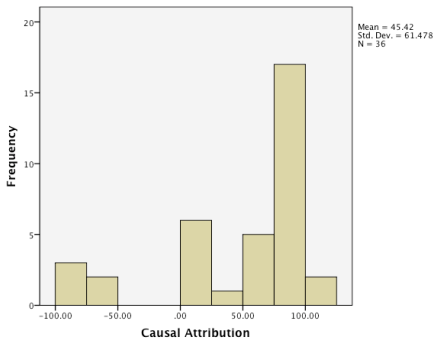
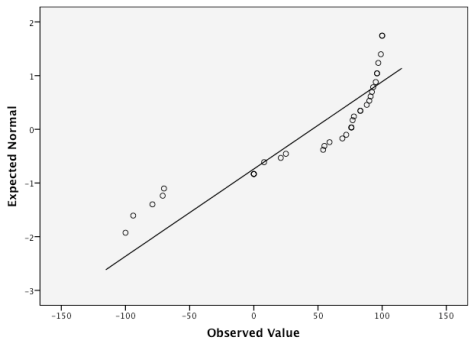
Histogram	Q-Q Plot
A) 25|0	
	
B) 75|0	
	
C) 75|75	
	

Table M6

Histograms and Q-Q Plots of Control of the System with the Continuous System Condition separated by Contingency

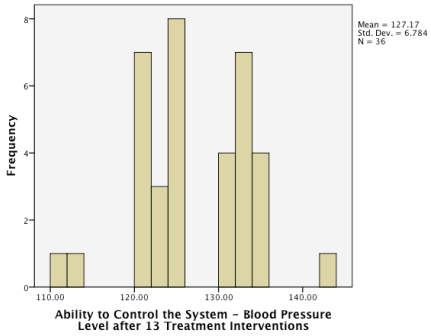
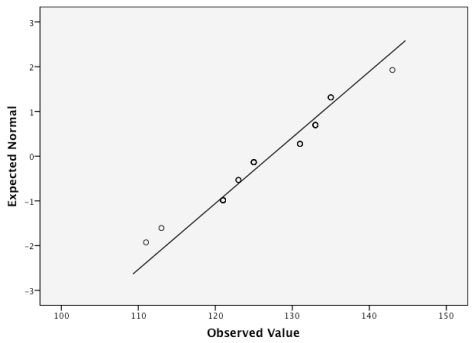
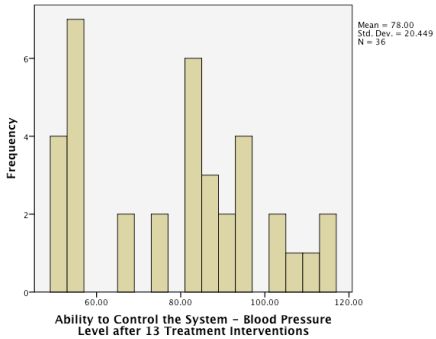
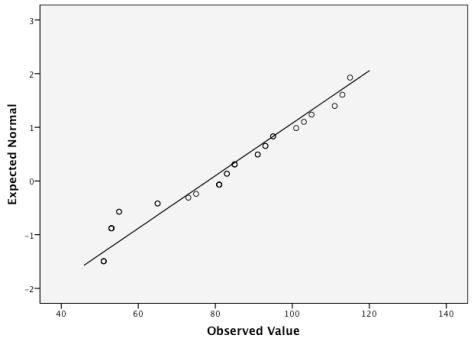
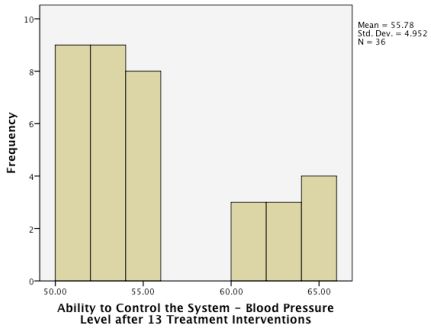
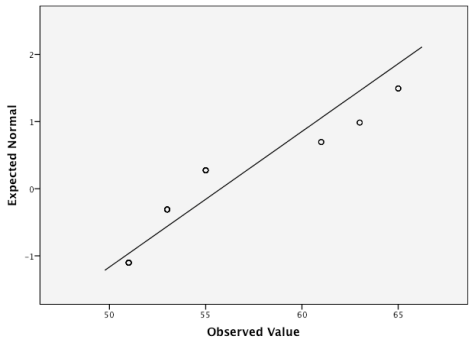
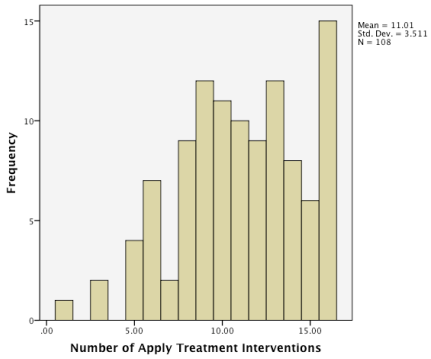
Histogram	Q-Q Plot
A) 25|0	
	
B) 75|0	
	
C) 75|75	
	

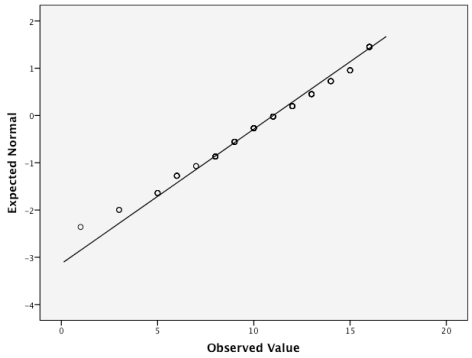
Table M7

Histograms and Q-Q Plots of Intervention Strategy Dependent Measures with the Continuous System Condition

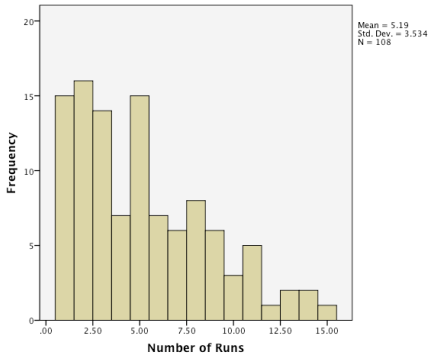
Histogram	Q-Q Plot

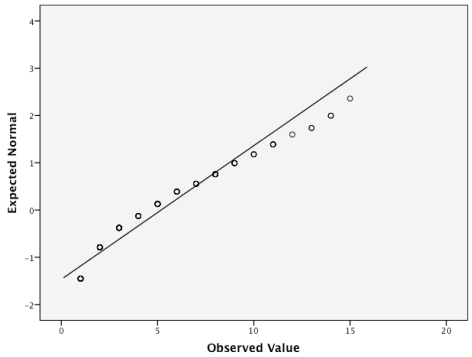
A) Number of Apply Treatment Interventions



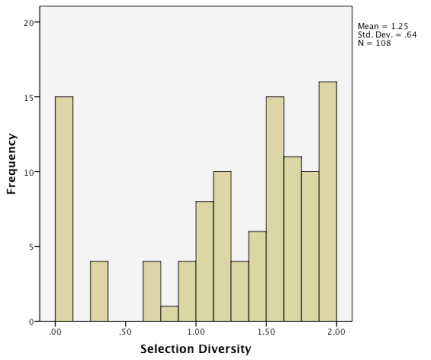


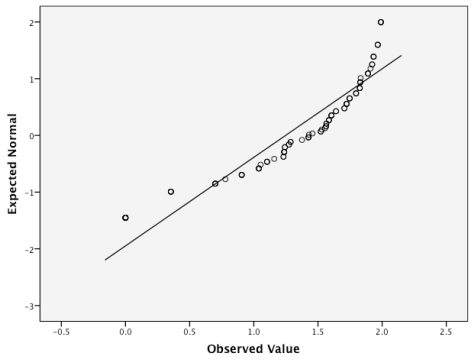
B) Number of Runs



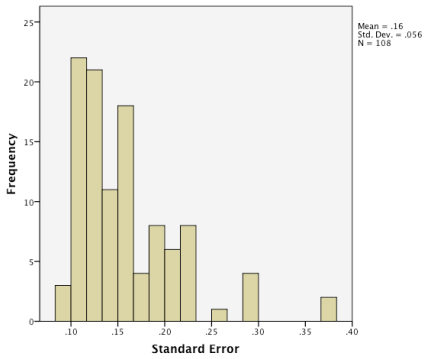


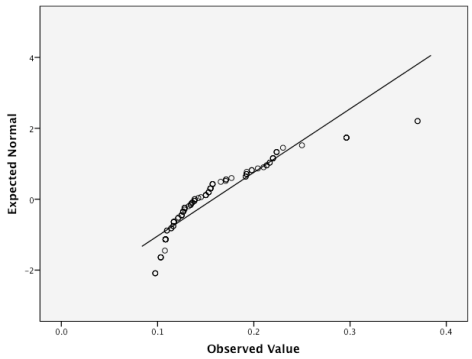
C) Selection Diversity



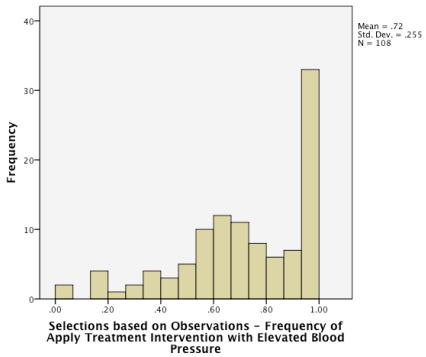


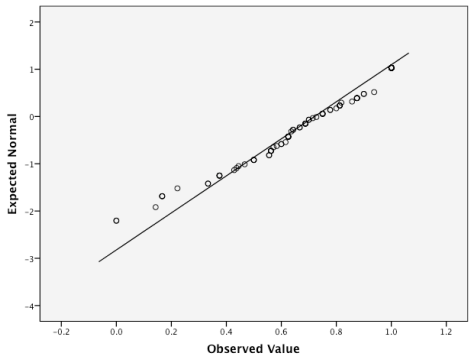
D) Standard Error



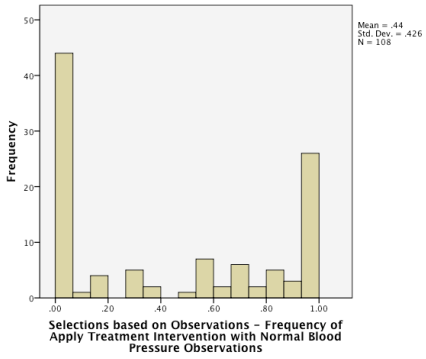


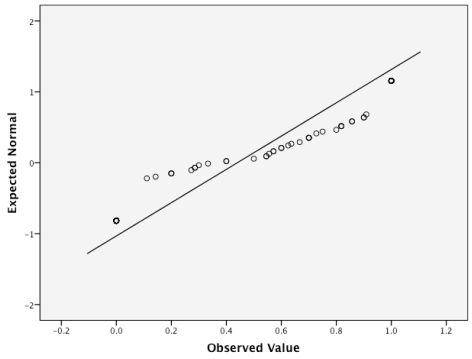
E) Selections based on Observations - Apply Treatment Intervention with Elevated Blood Pressure



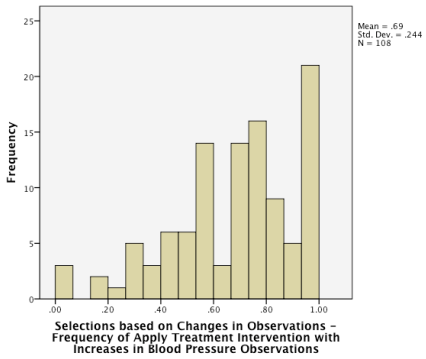


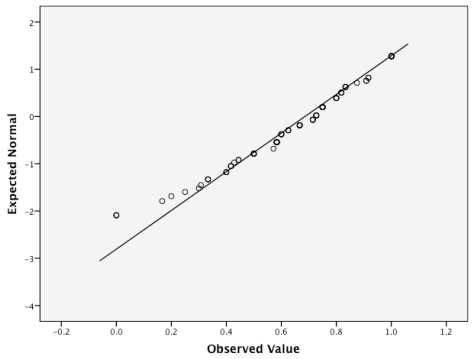
F) Selections based on Observations - Apply Treatment Intervention with Normal Blood Pressure Observations



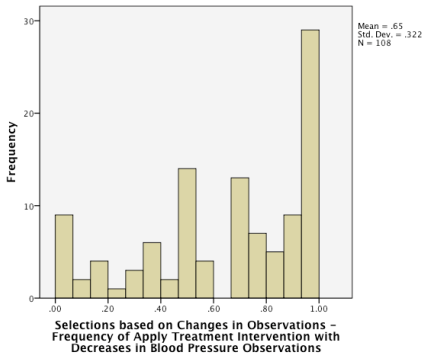


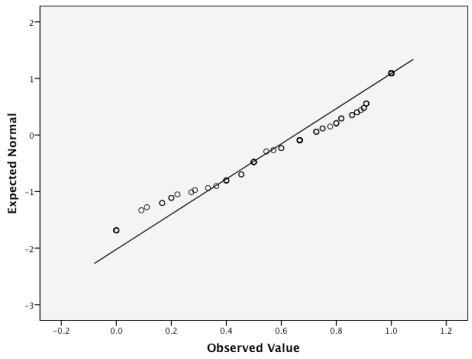
G) Selections based on Changes in Observations - Apply Treatment Intervention with Increases in Blood Pressure Observations



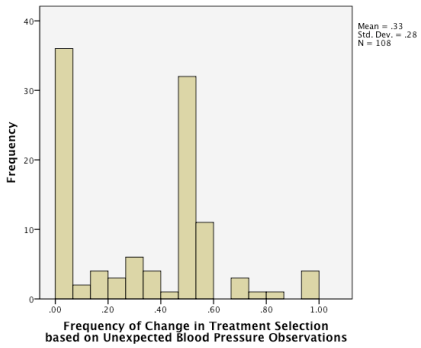


H) Selections based on Changes in Observations - Apply Treatment Intervention with Decreases in Blood Pressure Observations





I) Change in Treatment Selection based on Unexpected Blood Pressure Observations



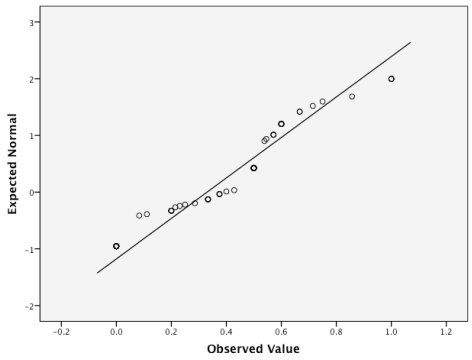
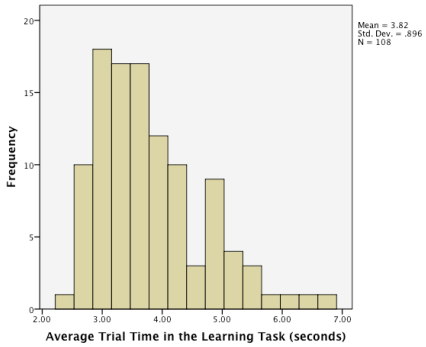
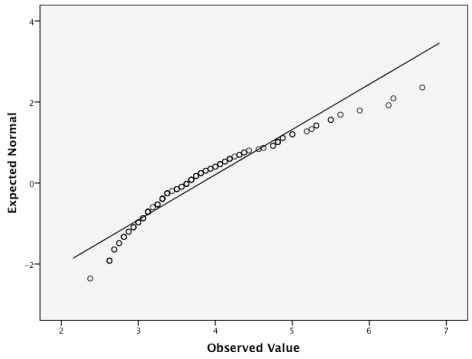
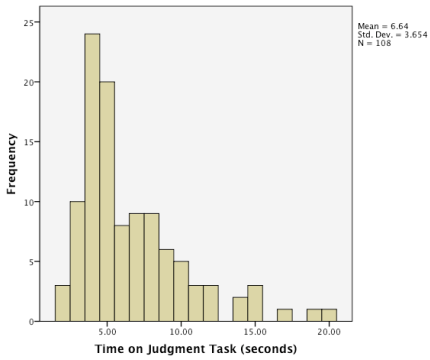
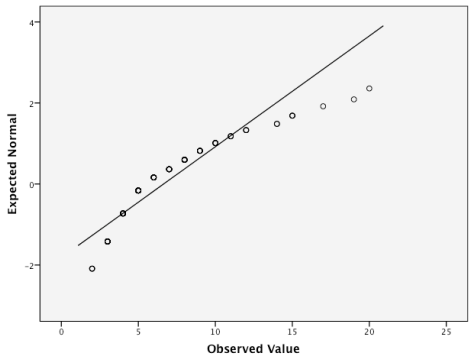


Table M8

Histograms and Q-Q Plots of Response Times with the Continuous System Condition

Histogram	Q-Q Plot
A) Average Trial Time in the Learning Task	
	
B) Time on Judgment Task	
	

APPENDIX N

Homogeneity of Variance Assessments

Table N1
*Tests of Homogeneity of Variance of Data from the Discrete System Condition separated by Contingency * Delay Experimental Manipulation Combinations*

Dependent Measure	*Levene Statistic*	*p*
Causal Attributions	1.583	.140
Ability to Control the System	5.022	.000
Number of Apply Treatment Interventions	2.818	.007
Number of Runs	2.681	.010
Selection Diversity	1.477	.175
Standard Error	7.464	.000
Selections based on Observations		
Apply Treatment Intervention with Elevated Blood Pressure Observations	2.646	.011
Apply Treatment Intervention with Normal Blood Pressure Observations	0.972	.462
Selections based on Changes in Observations		
Apply Treatment Intervention with Increases in Blood Pressure Observations	2.288	.027
Apply Treatment Intervention with Decreases in Blood Pressure Observations	0.980	.456
Apply Treatment Intervention with No Change in Blood Pressure Observations	4.866	.000
Change in Treatment Selection based on Unexpected Blood Pressure Observations	1.126	.353
Response Times		
Average Trial	0.969	.465
Judgment Task	2.591	.005

Note. *df* 1 = 8, *df* 2= 99

Table N2

*Tests of Homogeneity of Variance of Data from the Continuous System Condition separated by Contingency * Delay Experimental Manipulation Combinations*

Dependent Measure	*Levene Statistic*	*p*
Causal Attributions	2.805	.008
Ability to Control the System	12.340	.000
Number of Apply Treatment Interventions	2.886	.006
Number of Runs	3.563	.001
Selection Diversity	2.130	.040
Standard Error	5.599	.000
Selections based on Observations		
Apply Treatment Intervention with Elevated Blood Pressure Observations	5.753	.000
Apply Treatment Intervention with Normal Blood Pressure Observations	2.667	.030
Selections based on Changes in Observations		
Apply Treatment Intervention with Increases in Blood Pressure Observations	2.261	.029
Apply Treatment Intervention with Decreases in Blood Pressure Observations	1.985	.056
Change in Treatment Selection based on Unexpected Blood Pressure Observations	2.777	.008
Response Times		
Average Trial	0.875	.540
Judgment Task	0.831	.577

Note. $df\ 1 = 8$, $df\ 2 = 99$

APPENDIX O

Comprehensive Set of RBPF-3^2 and Paired Samples *t*-Tests

The RBPF-3^2 analyses review the data from each system condition separately. Table O1 highlights that there were no contingency by delay manipulations across the dependent measure with either system condition. Table O2 and O3 contain the contingency by system condition simple effects analyses. Tables O4 and O5 contain the delay by system condition simple effects analyses. Tables O6 and O7 contains the paired samples *t*-tests comparing the system conditions.

Table O1

RBPF-3^2 Analyses - Delay by Contingency Findings

	Discrete System		Continuous System	
	F(4, 64)	*p*	*F*(4, 64)	*p*
Causal attribution	1.52	.21	0.69	.60
Ability to control the system	0.66	.62	0.45	.77
Number of apply treatment interventions	1.28	.29	0.13	.97
Number of runs	1.34	.26	0.45	.77
Selection diversity	1.76	.15	0.75	.56
Standard error	0.59	.67	1.08	.37
Selections based on observations				
Elevated	2.09	.09	0.43	.79
Normal	0.36	.84	0.85	.50
Selections based on changes in observations				
Increases	1.25	.30	0.81	.52
Decreases	1.07	.38	0.77	.55
No change**	1.53	.21		
Selection changes based on unexpected observations	0.66	.62	0.80	.53
Response times				
Average trial in learning task	0.29	.88	0.22	.92
Judgment task	0.56	.69	0.86	.49

Note. *Analysis does not apply to the continuous system condition given there the observation always changed.

Table O2
RBPF-3[2] Analyses - Contingency by Discrete System Simple Effects

	M			
	25|0	75|0	75|75	*F*(2, 64)
Causal attributions	0.14 (52.44)	36.19 (57.75)	48.89 (55.18)	7.55*
Ability to control the system	124.06 (30.51)	85.17 (17.30)	55.17 (5.00)	NA[1]
Number of apply treatment interventions	11.33 (3.18)	10.94 (2.80)	9.25 (4.29)	4.12*[2]
Number of runs	5.81 (2.39)	7.06 (3.71)	6.03 (3.19)	2.29
Selection diversity	1.40 (0.51)	1.47 (0.61)	1.37 (0.69)	0.54
Standard error	0.15 (0.03)	0.13 (0.02)	0.19 (0.05)	22.99*[3]
Selections based on observations				
Elevated	.75 (.24)	.76 (.24)	.77 (.28)	0.17
Normal	.56 (.35)	.58 (.38)	.47 (.39)	1.51
Selections based on changes in observations				
Increases	.73 (.36)	.76 (.29)	.75 (.33)	0.14
Decreases	.57 (.36)	.57 (.37)	.45 (.38)	2.29
No change	.74 (.26)	.71 (.22)	.49 (.36)	7.67*[4]
Selection changes based on unexpected observations	.24 (.26)	.26 (.29)	.19 (.27)	0.97
Response times				
Average trial in learning task	61.31 (18.62)	61.86 (19.47)	59.31 (18.83)	0.14
Judgment task	8.19 (4.48)	7.56 (4.67)	8.08 (6.63)	0.31

Note. *$p < 0.05$. *SD* in parentheses.
[1] Differences are expected with the contingency manipulation.
[2] The number treatment selections applied to constructs containing the 75|75 contingency manipulations was statistically lower than those with the 25|0 manipulation, but neither differed significantly from the 75|0 contingency manipulation.
[3]Despite participants applying the treatment less with the 75|75 contingency manipulation, the standard error associated with this condition was still statistically greater than that of the 75|0 and 25|0 contingency manipulations.
[4]The frequency that participants applied treatment interventions when there was no change in blood pressure observation was significantly greater with the 25|0 and 75|0 contingency manipulations than the 75|75 contingency manipulation.

Table O3
RBPF-3[2] Analyses - Contingency by Continuous System Simple Effects

	M			
	25|0	75|0	75|75	*F*(2, 64)
Causal attributions	-23.67 (47.50)	68.75 (31.33)	45.42 (61.48)	38.72*
Ability to control the system	127.17 (6.78)	78.00 (20.45)	55.78 (4.95)	NA[1]
Number of apply treatment interventions	10.92 (2.47)	12.03 (3.02)	10.08 (4.53)	3.39*[2]
Number of runs	6.06 (2.83)	5.19 (4.08)	4.31 (3.45)	4.05*[3]
Selection diversity	1.53 (0.44)	1.12 (0.68)	1.09 (0.69)	9.23*[4]
Standard error	0.15 (0.02)	0.13 (0.02)	0.20 (0.07)	24.92*[5]
Selection based on observations				
Elevated	.68 (.15)	.81 (.20)	.67 (.35)	3.83*[6]
Normal	NA	.71 (.36)	.61 (.34)	NA*[7]
Selections based on changes in observations				
Increases	.66 (.16)	.73 (.20)	.66 (.33)	1.17
Decreases	.65 (.34)	.70 (.30)	.61 (.32)	0.72
Selection changes based on unexpected observations	.29 (.23)	.27 (.29)	.20 (.29)	1.55
Response times				
Average trial in learning task	60.86 (14.85)	61.81 (12.43)	60.50 (15.89)	0.10
Judgment task	7.81 (4.46)	6.03 (3.10)	6.08 (3.04)	2.89

Note. *$p < 0.05$. *SD* in parentheses.
[1]Differences are expected with the contingency manipulation.
[2]The number treatment selections applied to constructs containing the 75|75 contingency manipulations was statistically lower than those with the 75|0 manipulation.
[3]Participants switched between the treatment options significantly more with the 25|0 contingency manipulation than with the 75|75 contingency manipulation.
[4]Participants' intervention selections were more variable with the 25|0 contingency manipulation than with the 75|0 and 75|75 contingency manipulations.
[5]The standard error associated with the 75|75 contingency manipulation was still statistically greater than that of the 75|0 and 25|0 contingency manipulations.
[6]Participants applied the treatment more frequently when the blood pressure observation was elevated with the 75|0 contingency manipulation than with the 75|75 contingency manipulation.
[7]Blood pressure observations did not enter the normal range with the 25|0 contingency manipulation.

Table O4

RBPF-3[2] Analyses - Delay by Discrete System Simple Effects

	M			
	One-Trial Delay	Two-Trial Delay	Three-Trial Delay	*F*(2, 64)
Causal attributions	34.36 (57.10)	30.06 (53.62)	20.81 (64.89)	0.57
Ability to control the system	86.00 (30.57)	88.22 (29.43)	90.17 (31.93)	1.26
Number of apply treatment interventions	10.92 (3.89)	10.67 (3.54)	9.94 (3.28)	0.86
Number of runs	5.92 (3.42)	6.11 (2.86)	6.86 (3.20)	1.28
Selection diversity	1.28 (0.62)	1.43 (0.65)	1.52 (0.53)	3.46*[1]
Standard error	.16 (.05)	.15 (.04)	.16 (.04)	0.60
Selections based on observations				
Elevated	.80 (.23)	.78 (.26)	.70 (.26)	2.52
Normal	.60 (.39)	.49 (.38)	.52 (.36)	1.49
Selections based on changes in observations				
Increases	.82 (.25)	.77 (.34)	.64 (.35)	4.13*[2]
Decreases	.61 (.38)	.44 (.36)	.54 (.37)	3.27*[3]
No change	.65 (.34)	.67 (.31)	.63 (.28)	0.15
Selection changes based on unexpected observations	.22 (.26)	.17 (.24)	.31 (.30)	4.02*[4]
Response times				
Average trial in learning task	61.11 (16.90)	62.06 (22.15)	59.31 (17.44)	0.34
Judgment task	8.36 (6.23)	8.50 (5.80)	6.97 (3.49)	0.87

Note. *$p < 0.05$. *SD* in parentheses.

[1] Participants' intervention selections were more diverse or variable with the three-trial delay manipulation than with the one-trial delay manipulation.

[2] Participants applied the treatment more frequently in combinations with increases in blood pressure with the one-trial delay manipulation than with the two-trial and three-trial delay manipulations.

[3] Participants applied the treatment more frequently after observing a decrease in blood pressure with the one-trial delay manipulation than the two-trial delay manipulation.

[4]Participants switched their treatment selection when viewing unexpected observations more often with the three-trial delay manipulation than with the two-trial delay manipulation.

Table O5
RBPF-3[2] Analyses - Delay by Continuous System Simple Effects

	M			
	One-Trial Delay	Two-Trial Delay	Three-Trial Delay	*F*(2, 64)
Causal attributions	38.39 (57.70)	30.83 (68.03)	21.28 (60.50)	1.23
Ability to control the system	84.61(32.09)	86.89 (34.50)	89.44 (31.66)	1.33
Number of apply treatment interventions	11.58 (3.17)	11.14 (3.62)	10.31 (3.69)	1.50
Number of runs	5.39 (3.84)	4.39 (2.92)	5.78 (3.72)	2.71
Selection diversity	1.21 (0.61)	1.21 (0.70)	1.33 (0.62)	0.70
Standard error	0.16 (0.06)	0.15 (0.05)	0.16 (0.06)	0.46
Selections based observations				
Elevated	.78 (.24)	.72 (.29)	.67 (.22)	1.92
Normal	.47 (.41)	.47 (.43)	.38 (.44)	0.96
Selections based on changes in observations				
Increases	.76 (.21)	.66 (.26)	.64 (.25)	2.93
Decreases	.69 (.30)	.69 (.28)	.57 (.37)	1.79
Selection changes based on unexpected observations	.25 (.28)	.22 (.22)	.30 (.31)	1.10
Response times				
Average trial in learning task	61.31 (15.87)	61.19 (14.14)	60.67 (13.28)	0.02
Judgment task	7.00 (4.15)	6.39 (3.21)	6.53 (3.61)	0.29

Note. $*p < 0.05$. *SD* in parentheses.

Table O6
Paired Samples t-tests - Effect of System on Causal Attributions

	M				*95% CI*		
	Discrete	Continuous	*t(35)*	*p*	*LL*	*UL*	*d*
Contingency							
25|0	0.14 (52.44)	-23.67 (47.50)	2.27	0.03	2.47	45.14	0.38
75|0	36.19 (57.75)	68.75 (31.33)	-3.47	< .01	-51.61	-13.51	-0.58
75|75	48.89 (55.18)	45.42 (61.48)	0.24	0.81	-25.50	32.44	0.04
Delay							
One-trial	34.36 (57.10)	38.39 (57.70)	-0.36	0.72	-26.80	18.74	-0.06
Two-trial	30.06 (53.62)	30.83 (68.03)	-0.06	0.95	-26.77	25.23	-0.01
Three-trial	20.81 (64.89)	21.28 (60.50)	-0.04	0.97	-26.04	25.10	-0.01

Note. *SD* in parentheses. Discrete = Discrete System Condition. Continuous = Continuous System Condition. CI = Confidence Interval; LL= lower limit; UL=upper limit. I conducted paired *t*-tests for separating the data by contingency and delay with data from participants in the continuous system condition subtracted from data from participants in the discrete system condition. Positive outcomes suggest higher attributions with the discrete system condition and negative outcomes suggest higher attributions with the continuous system condition. All tests two-tailed.

Table O7
Paired Samples t-tests - Effect of System on System Control

	M				*95% CI*		
	Discrete	Continuous	*t(35)*	*p*	*LL*	*UL*	*d*
Contingency							
25|0	124.06 (7.10)	127.17 (6.78)	-1.99	.06	-6.29	0.07	-.33
75|0	85.17 (17.30)	78.00 (20.45)	1.80	.08	-0.90	15.23	.30
75|75	55.17 (5.00)	55.78 (4.95)	-0.56	.58	-2.82	1.59	-.09
Delay							
One-trial	86.00 (30.57)	84.61 (32.09)	0.61	.55	-3.27	6.05	.10
Two-trial	88.22 (29.43)	86.89 (34.50)	0.43	.67	-4.93	7.60	.07
Three-trial	90.17 (31.93)	89.44 (31.66)	0.28	.78	-4.36	5.81	.05

Note. *SD* in parentheses Discrete = Discrete System Condition. Continuous = Continuous System Condition. CI = Confidence Interval; LL= lower limit; UL=upper limit. I conducted paired *t*-tests for separating the data by contingency and delay with data from participants in the continuous system condition subtracted from data from participants in the discrete system condition. Positive outcomes suggest higher attributions with the discrete system condition and negative outcomes suggest higher attributions with the continuous system condition. All tests two-tailed.

APPENDIX P

Analyses considering Alternative Conversion Techniques, Causal Attribution Scales, and Sequence Effects

I reviewed and conducted a subset of analyses using subsets of and transformations of the causal attribution dependent measure. These alternative approaches to the -100 to 100 rating scale considered the distributions of causal ratings data with outliers removed, with abbreviated scales (0 to 100), with transformed data to adjust for non-normality, and with the data separated by sequence order.

Figure P1 through Figure P6 depict the distribution of data by casual contingency manipulation and system condition for each of these techniques. In all of these figures, the unaltered data set is depicted in the upper left corner and is labeled *-100 to 100 scale, raw data.* Across these distributions, the y-axis represents the frequency that the causal attribution was selected on a scale ranging from zero to 30. The x-axis represents causal attribution scores from -100 to 100 separated by 5-point intervals with the exceptions of the square root and logarithmic transformations. With the square root transformation (Sqrt(101-X)), the x-axis ranges from zero to 20 and the data is charted in 0.5 point intervals. With the logarithmic transformation (log10(101-X)), the x-axis ranges from zero to three and is charted in 0.1 point intervals. In the following sections, I will describe the conversion techniques and discuss their influences on analyses using causal attributions as the dependent measure in more detail.

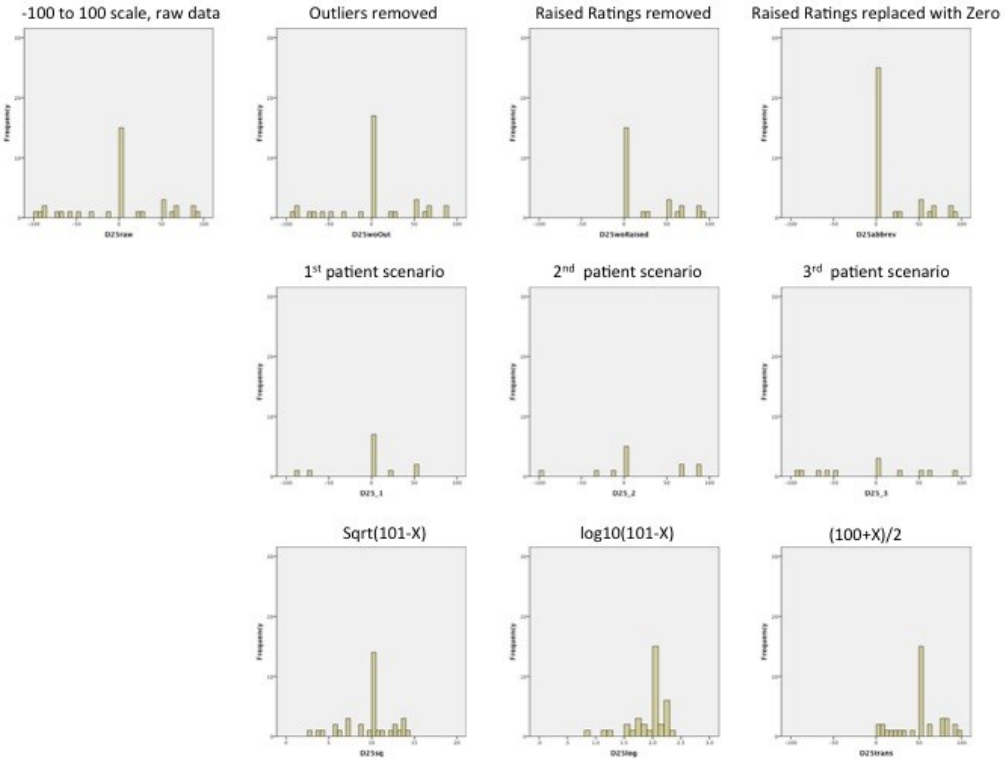


Figure P1. Distributions for the 25|0 Contingency Manipulations with the Discrete System Condition.

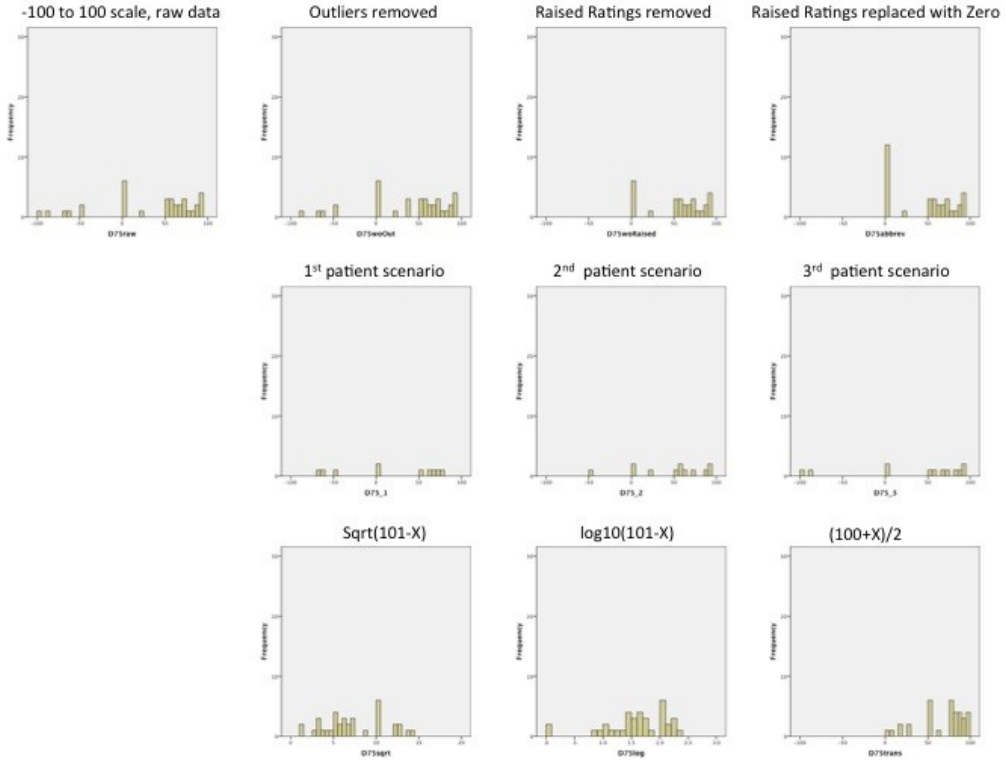


Figure P2. Distributions for the 75|0 Contingency Manipulations with the Discrete System Condition.

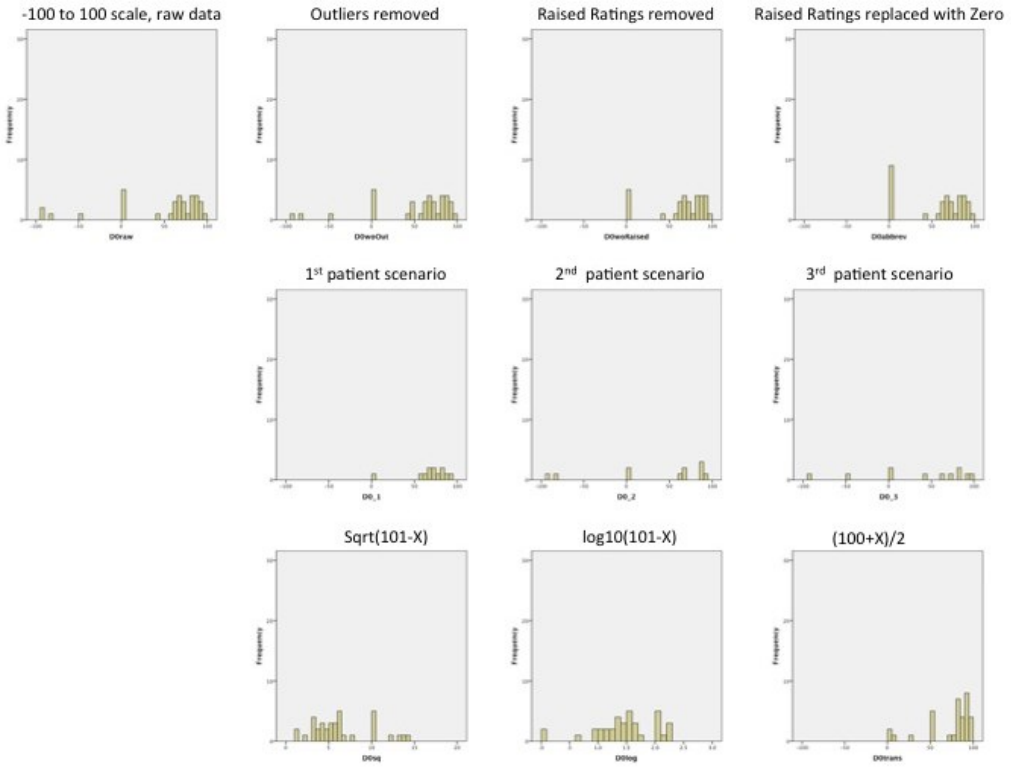


Figure P3. Distributions for the 75|75 Contingency Manipulations with the Discrete System Condition.

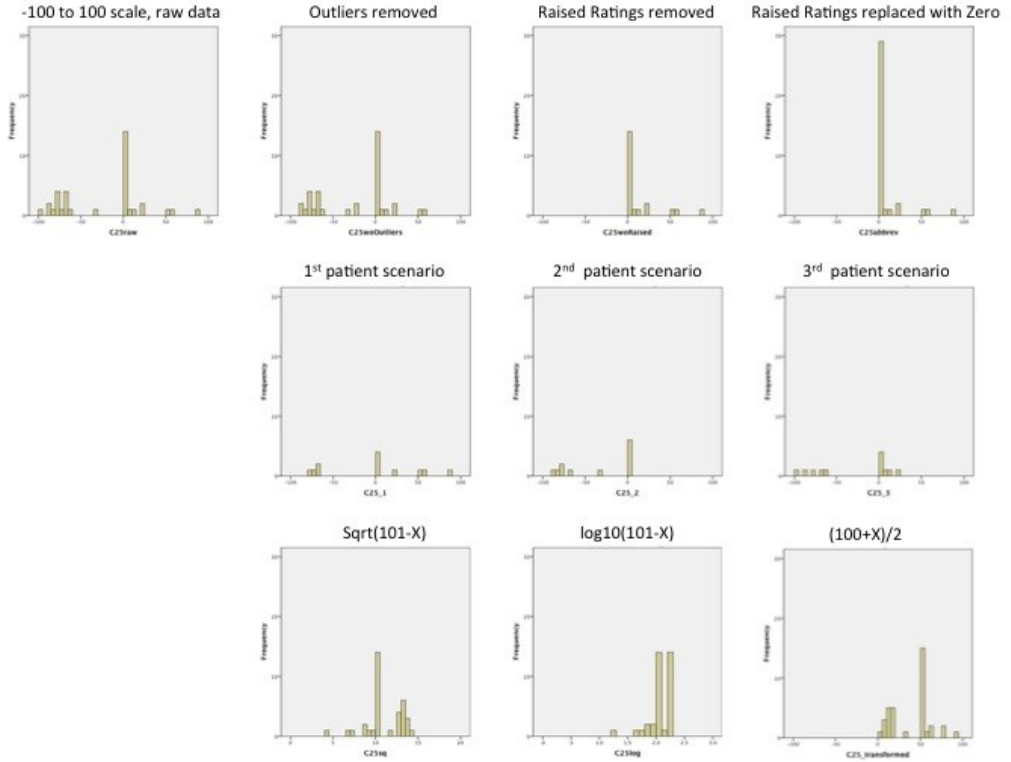


Figure P4. Distributions for the 25|0 Contingency Manipulations with the Continuous System Condition.

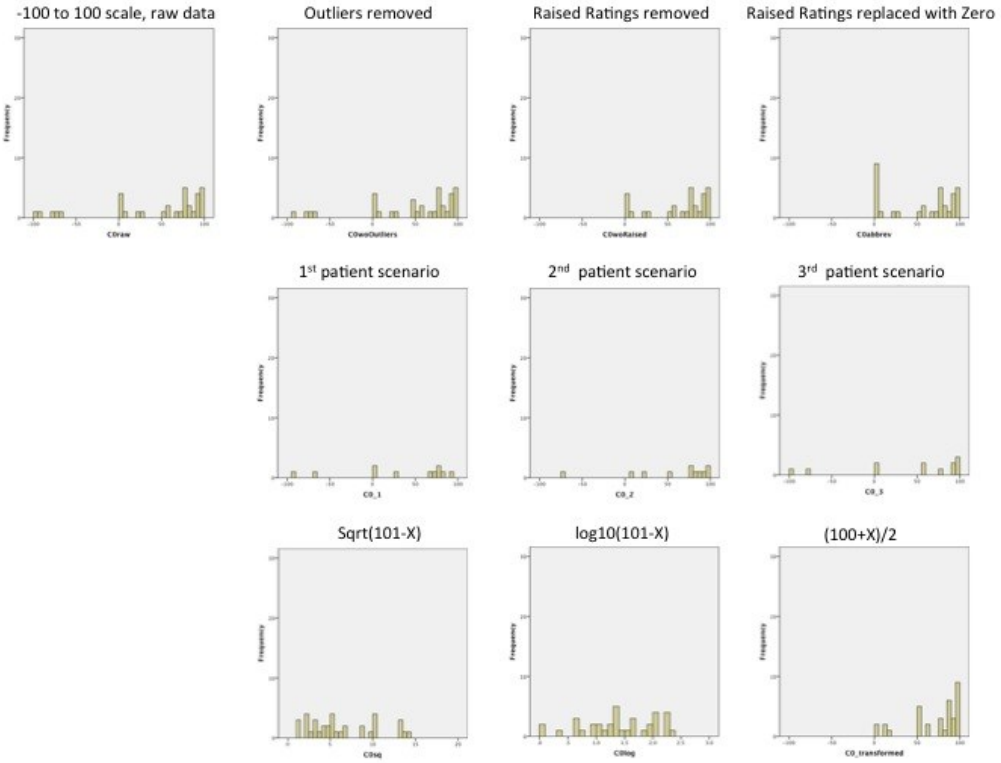


Figure P5. Distributions for the 75|0 Contingency Manipulations with the Continuous System Condition.

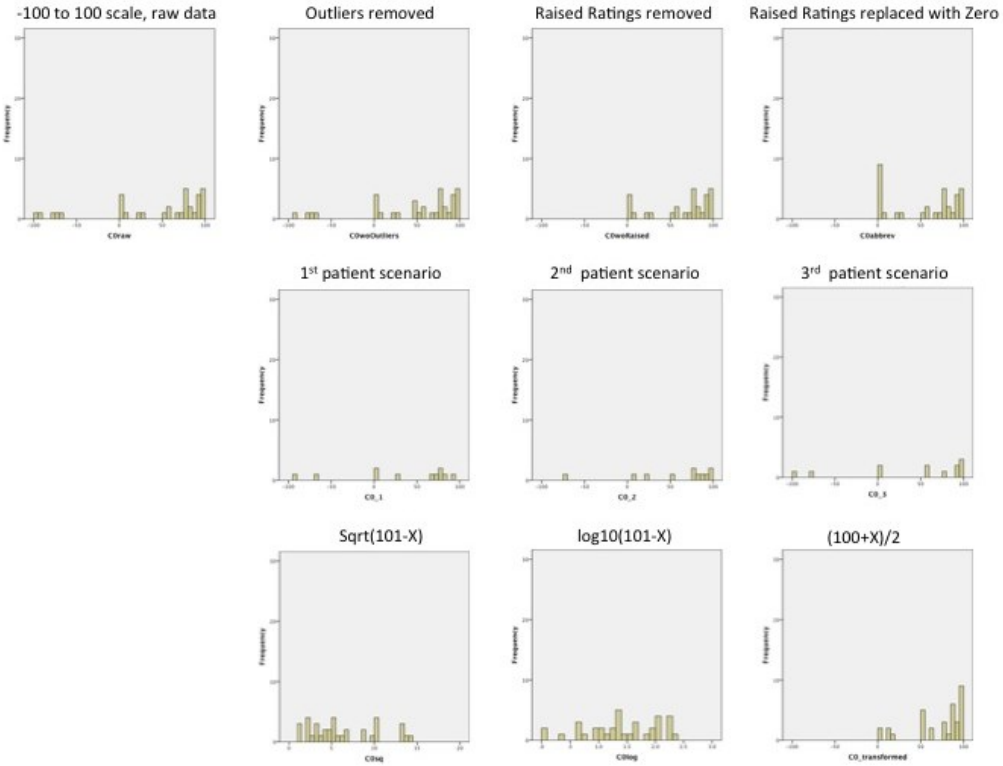


Figure P6. Distributions for the 75|75 Contingency Manipulations with the Continuous System Condition.

Outliers Removed

I reviewed the data set with outliers beyond two standard deviations of the mean of each contingency manipulation removed and replaced with the mean value of that contingency manipulation. This method resulted in the replacement of five of the 108 ratings from discrete system condition's data set and four from the continuous system condition's data set.

Discrete system condition. With the discrete system, the RBPF-3^2 ANOVA findings with the outliers removed were the same as the raw data analysis, which found a significant contingency main effect ($F(2,64) = 10.35$, $p < 0.01$). The relationship between the contingency manipulations were the same with participants assigning less causality to the 25|0 manipulation ($M = 0.31$, $SD = 47.03$) than the 75|0 ($M = 36.43$, $SD = 50.72$) and 75|75 ($M = -50.05$, $SD = 47.98$) contingency manipulations.

Continuous system condition. As with the raw data analysis, there were no significant interaction or a delay main effect with the RBPF-3^2 ANOVA using the continuous system data set. However, the 75|0 contingency manipulation ($M = 72.15$, $SD = 16.33$) was greater than 75|75 ($M = 46.42$, $SD = 54.81$), which was greater than 25|0 ($M = -24.68$, $SD = 42.06$): $F(2,64) = 48.63$, $p < 0.01$. This is a departure from the raw data analysis that found that the 75|0 and 75|75 contingency manipulations were not significantly different.

Raised Ratings Removed

I removed the raised ratings (-100 to 0) from the data set for a subset of analyses that compared the results from the present study to previous findings. Tables P1 and P2

present the findings from independent samples *t*-tests that compare the data set with the raised ratings removed to the raw data set.

Table P1
Comparison between the raw data and data set with the raised ratings removed with the discrete system condition

	M						
	Raw data	Raised removed	*n*	*t*	*df*	*p*	*d*
One-trial delay*	34.36 (57.10)	53.30 (40.04)	30	-1.58	62.31	.12	-0.40
Two-trial delay*	30.06 (53.62)	46.45 (36.09)	31	-1.49	61.62	.14	-0.38
Three-trial delay*	20.80 (64.89)	53.04 (33.76)	27	-2.56	55.15	.01	-0.69
25|0	0.14 (52.44)	25.65 (33.65)	26	-2.18	60.00	.03	-0.56
75|0*	36.19 (57.75)	57.47 (33.78)	30	-1.86	57.87	.07	-0.49
75|75*	48.89 (55.18)	65.00 (31.35)	32	-1.50	56.61	.14	-0.40

Note. $n_{raw\ data}$=36. *n* = number of raised removed data points. *SD* in parentheses. *Variances unequal and t-test performed assuming inequality. All tests two-tailed.

Table P2
Comparison between the raw data and data set with the raised ratings removed with the continuous system condition

	M						
	Raw data	Raised removed	*n*	*t*	*df*	*p*	*d*
One-trial delay*	38.39 (57.70)	54.47 (36.72)	32	-1.39	60.08	.17	-0.36
Two-trial delay*	30.83 (68.03)	65.81 (33.72)	27	-2.68	53.90	.01	-0.73
Three-trial delay*	21.28 (60.50)	48.07 (36.01)	28	-2.20	58.43	.03	-0.58
25|0*	-23.67 (47.50)	12.24 (23.46)	21	-3.81	53.90	.00	-1.04
75|0	68.75 (31.33)	73.14 (17.19)	35	-0.73	69.00	.47	-0.18
75|75*	45.42 (61.48)	66.10 (34.68)	31	-1.73	56.63	.09	-0.46

Note. $n_{raw\ data}$=36. *n* = number of raised removed data points. *SD* in parentheses. *Variances unequal and t-test performed assuming inequality. All tests two-tailed.

Raised Ratings Replaced with Zero

If participants felt that a treatment raised the patients' blood pressure levels, it is arguable that they would have determined that the treatment had no effect (i.e., did not influence the patient's blood pressure) on a 0 (no effect) to 100 (always lowered) scale. So, I explored replacing the raised ratings with zeros to construct an abbreviated rating

scale analogous to the scale used within the causal reasoning literature. With the discrete and continuous conditions, the result of the RBPF-3^2 ANOVAs were the same as with the raw data set.

Scale reduction

The scale reduction technique (i.e., (100+X)/2) was another method of reducing the range of the scale to be analogous with the 100-point scales used in previous research. As with the raw data set, the RBPF-3^2 ANOVAs resulted in contingency main effects with both the discrete ($F(2, 64) = 7.54$, $p < .01$) and continuous ($F(2, 64) = 38.71$, $p < .01$) system conditions. Also, the relationships between the contingency manipulations remained the same as the causal attributions associated with the 75|0 and 75|75 contingency manipulations being significantly greater than the causal attributions associated with the 25|0 contingency manipulation.

Separated by Sequence

I reviewed the separated causal attributions relative to sequence order to explore if there were significant sequencing effects. With these analyses, the number of replications of each experimental manipulation combination (contingency * delay) was reduced from 12 to four. Table P3 presents the findings from the RBPF-3^2 ANOVAs related to the contingency main effects for the discrete and continuous system conditions separated by sequence order.

Table P3
RBPF-3^2 Contingency Main Effect Findings relative to Sequence Order

	Discrete system		Continuous system	
	F(2, 16)	*p*	*F(2, 16)*	*p*
1st position	9.55	< .01	4.03	.03
2nd position	1.16	.34	27.43	< .01
3rd position	2.88	.09	9.55	< .01

Discrete system condition. None of the findings associated with the data separated by sequence order were the same as with the combined analysis. Although, there was a significant contingency main effect with the 1st position in the sequence, only 75|75 contingency manipulation was greater than the 25|0 manipulation rather than both the 75|0 and 75|0 contingency manipulations. Further, there was not a significant contingency main effects related to the 2nd and 3rd positions.

Continuous system condition. Unlike the discrete system condition, the contingency main effect persisted across the three positions in the sequence with the continuous system condition. However, only the 75|0 contingency manipulation was greater than the 25|0 contingency manipulation with the first position in the sequence order. The analyses considering the second and third positions replicated the findings of the combined analysis with the ratings associated with the 75|0 and 75|75 contingency manipulations being greater than those related to the 25|0 contingency manipulation.

Data Transformations

Given the raw data set is not normally distributed (see Appendix M), I attempted to transform the data to reduce the amount of skew. I explored transforming the data by taking the square root of the causal attribution added to 101 (to account for the negative values associated with the raised ratings) and by taking the log of the causal attribution added to 101. Neither of these transformations altered the RBPF-3^2 ANOVAs findings that there was a significant contingency main effect with both the discrete or continuous system conditions (shown in Table P4). Also, the finding that the 25|0 contingency manipulation was significantly less than the 75|0 and 75|75 contingency manipulations persisted with both transformation techniques.

Table P4

RBPF-3[2] Contingency Main Effect Findings with Transformed Data Sets

	Discrete System		Continuous System	
	F(2, 64)	*p*	*F(2, 64)*	*p*
Log10	8.33	< 0.01	32.39	< 0.01
Sqrt	7.54	< 0.01	42.79	< 0.01

APPENDIX Q

Possible System Control Differences given Intervention Strategy Extremes

Large differences in system control performance were likely only with the 75|0 contingency manipulation given the limited number of trials. This is especially true given system control measure only considered 13 treatment interventions, rather than 16, in order to assess the same number of treatment interventions across the delay conditions. Table Q1 presents range of system control outcomes with the intervention distribution extremes (minimum number of treatment interventions (1 apply treatment: 12 withhold treatments) and maximum number of treatment interventions (13 apply treatments: 0 withhold treatments)) relative to each of the contingency manipulations. It is not feasible for system control to differ as a function of delay with the 75|75 contingency manipulation, because applying and withholding the treatment have the same outcome. Although there was less opportunity for variation given the limited number of trials with the 25|0 contingency manipulation than with the 75|0 contingency manipulation, system control outcomes were at neither extreme of the possible outcomes.

Table Q1

Range System Control Outcomes with Extreme Intervention Distributions relative to System Control with each of the Contingency Manipulation in the Discrete and Continuous System Conditions

	1 apply: 12 withhold	13 apply: 0 withhold	Discrete		Continuous	
			M	*SD*	*M*	*SD*
25|0	141	111	124.06	7.10	127.17	6.78
75|0	141	50	85.17	17.30	78.00	20.45
75|75	50	50	55.17	5.00	55.78	4.95

Note. Discrete = Discrete System Condition, Continuous = Continuous System condition. Assumes a starting blood pressure level of 125, with a maximum level of 150 and minimum level of 50. Applying treatment applications reduced blood pressure level by 8 and withholding the treatment increased the blood pressure level by 2.

APPENDIX R

Correlations between Dependent Measures

To justify further exploration into the implications of intervention strategy on reasoning outcomes, I performed a series of Pearson product-moment correlational analyses assessing the relationships between dependent measures describing participants' intervention selections and reasoning outcomes. I considered the reasoning outcomes associated with both the causal reasoning (i.e., causal attributions) and system dynamics (i.e., ability to control the system) literatures. The results of these analyses suggest that intervention strategy may explain causal attribution better than the ability to control the system, and further highlights the differences between the system conditions.

Table R1 and R2 contains the correlations between the dependent measures for the discrete system and continuous system conditions respectively. Both of these analyses violate the assumption of data independence given participants interacted with multiple experimental manipulation combinations.

Tables R3, R4, and R5 contain the correlations with the data separated by contingency manipulations for the discrete system. Tables R6, R7, and R8 contain the correlations with the data separated by delay manipulations for the discrete system. Tables R9, R10, and R11 contain the correlations with the data separated by contingency manipulations for the continuous system. Tables R12, R13, and R14 review the correlations with the data separated by delay manipulations for the continuous system.

The apply treatment intervention with no change in blood pressure observations dependent measure is not applicable with the continuous system condition because the blood pressure observation either increase or decreases on every trial. Although I removed this item (number 11) from the series of correlations with the continuous system condition data, I did not change the numbering scheme for remaining items to retain consistency between the discrete and continuous system analyses.

Table R1

Correlations between Measures with Data from the Discrete System Condition

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Causal Attribution														
2. System Control	-.431**													
3. Number of Apply Treatments	.259**	.021												
4. Number of Runs	-.184	.145	-.577**											
5. Selection Diversity	-.067	-.050	.123	-.234*										
6. Standard Error	-.085	-.170	-.735**	.375**	-.155									
7. Treatment|Elevated	.338**	-.220*	.500**	-.179	.131	-.348**								
8. Treatment|Normal	.133	.014	.566**	-.613**	.059	-.375**	-.235*							
9. Treatment | Increase	.245*	-.151	.329**	-.130	.090	-.232*	.705**	-.113						
10. Treatment | Decrease	.089	.039	.550**	-.585**	.060	-.348**	-.175	.920**	-.180					
11. Treatment | No Change	.091	.142	.833**	-.384**	.057	-.641**	.321**	.411**	.205*	.386**				
12. Switch given Unexpected Observation	-.059	-.066	-.118	.135	.156	-.041	-.069	-.137	-.074	-.149	-.007			
13. Value on 16th Trial	.011	.006	-.042	.033	.078	.107	.038	-.048	-.005	-.010	-.029	-.075		
14. Average Trial Time	-.084	.054	-.037	.217*	-.019	.035	.061	-.142	.022	-.107	.032	.145	.036	
15. Judgment Time	-.147	.056	-.362**	.240*	.025	.228*	-.106	-.271**	-.031	-.286**	-.275**	.068	.092	.435**

Note. $N = 108$. * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Table R2

Correlations between Measures with Data from the Continuous System Condition

	1	2	3	4	5	6	7	8	9	10	12	13	14
1. Causal Attribution													
2. System Control	-.532**												
3. Number of Apply	.489**	-.039											
4. Number of Runs	-.277**	.270**	-.497**										
5. Selection Diversity	-.413**	.379**	-.683**	.689**									
6. Standard Error	-.163	-.237*	-.501**	.178	.367**								
7. Treatment|Elevated	.389**	-.093	.777**	-.488**	-.557**	-.408**							
8. Treatment|Normal	.727**	-.673**	.524**	-.306**	-.564**	-.156	.311**						
9. Treatment | Increase	.455**	-.096	.803**	-.333**	-.547**	-.367**	.633**	.423**					
10. Treatment | Decrease	.380**	-.046	.777**	-.495**	-.533**	-.366**	.602**	.431**	.405**				
12. Switch given Unexpected	-.209*	.231*	-.097	.111	.106	-.075	-.140	-.218*	-.124	-.045			
13. Value on 16th Trial	-.580**	.918**	-.155	.357**	.489**	-.158	-.196*	-.738**	-.196*	-.131	.237*		
14. Average Trial Time	.003	.026	.054	.131	-.031	-.146	.006	.052	.051	.097	-.041	-.017	
15. Judgment Time	-.109	.228*	-.029	.081	.099	-.070	-.054	-.206*	-.044	.066	.107	.182	.421**

Note. $N = 108$. * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Table R4

Correlations between Measures with 75|0 Contingency Manipulation Data from the Discrete System Condition

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Causal Attribution														
2. System Control	-.509**													
3. Number of Apply Treatments	.530**	-.934**												
4. Number of Runs	-.450**	.717**	-.765**											
5. Selection Diversity	.028	-.270	.208	-.261										
6. Standard Error	-.511**	.840**	-.910**	.573**	-.164									
7. Treatment|Elevated	.259	-.635**	.581**	-.171	.169	-.712**								
8. Treatment|Normal	.283	-.387*	.491**	-.679**	.032	-.226	-.393*							
9. Treatment | Increase	.358*	-.434**	.399*	-.170	.183	-.436**	.746**	-.357*						
10. Treatment | Decrease	.165	-.412*	.520**	-.644**	.031	-.266	-.303	.945**	-.409*					
11. Treatment | No Change	.559**	-.832**	.859**	-.537**	.162	-.907**	.685**	.199	.404*	.183				
12. Switch given Unexpected Observation	-.282	-.035	-.032	.040	.149	.049	.121	-.136	.081	-.142	-.077			
13. Value on 16th Trial	-.071	-.041	-.085	.157	.221	.187	.016	-.067	.095	-.021	-.200	-.041		
14. Average Trial Time	-.314	.179	-.145	.252	-.212	.039	.113	-.325	.124	-.287	-.090	-.108	.058	
15. Judgment Time	-.329	.472**	-.563**	.377*	-.003	.526**	-.294	-.345*	-.094	-.414*	-.388*	.071	.179	.235

Note. N= 36 * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Table R5

Correlations between Measures with 75|75 Contingency Manipulation Data from the Discrete System Condition

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Causal Attribution														
2. System Control	-.023													
3. Number of Apply Treatments	.372*	.043												
4. Number of Runs	-.036	.100	-.643**											
5. Selection Diversity	-.173	.126	.036	-.332*										
6. Standard Error	-.167	.352*	-.670**	.733**	-.111									
7. Treatment|Elevated	.556**	-.084	.236	-.144	.117	-.152								
8. Treatment|Normal	.182	.037	.940**	-.622**	-.014	-.648**	-.099							
9. Treatment | Increase	.536**	-.220	.225	-.052	.059	-.139	.912**	-.079						
10. Treatment | Decrease	.091	.210	.821**	-.550**	-.009	-.506**	-.097	.875**	-.114					
11. Treatment | No Change	.200	.111	.841**	-.553**	-.044	-.494**	-.068	.879**	-.092	.885**				
12. Switch given Unexpected Observation	-.019	-.113	-.098	-.118	.326	-.039	-.147	-.050	-.184	-.022	.015			
13. Value on 16th Trial	.186	-.250	.182	-.083	-.031	-.160	.269	.101	.165	.173	.242	-.112		
14. Average Trial Time	.071	-.189	-.121	.247	.093	.188	-.003	-.098	-.003	-.019	-.039	.475**	.164	
15. Judgment Time	-.066	-.243	-.514**	.337*	.002	.305	-.076	-.476**	-.074	-.475**	-.411*	.191	.123	.683**

Note. N= 36 * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Table R6

Correlations between Measures with One-Trial Delay Manipulation Data from the Discrete System Condition

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Causal Attribution														
2. System Control	-.412*													
3. Number of Apply Treatments	.213	.136												
4. Number of Runs	-.342*	.067	-.577**											
5. Selection Diversity	-.204	-.130	.280	-.274										
6. Standard Error	-.204	-.252	-.878**	.368*	-.287									
7. Treatment|Elevated	.273	-.360*	.279	-.039	.129	-.277								
8. Treatment|Normal	.230	.175	.752**	-.706**	.200	-.606**	-.256							
9. Treatment | Increase	.332*	-.293	.142	-.012	.264	-.264	.742**	-.271						
10. Treatment | Decrease	.162	.140	.740**	-.676**	.225	-.558**	-.206	.975**	-.283					
11. Treatment | No Change	.105	.221	.946**	-.429**	.162	-.839**	.210	.640**	-.005	.607**				
12. Switch given Unexpected Observation	.163	-.246	-.071	-.092	.148	-.023	-.076	.014	-.021	-.021	-.054			
13. Value on 16th Trial	.061	-.095	-.166	.048	-.036	.150	.226	-.250	.165	-.251	-.141	.118		
14. Average Trial Time	-.063	-.104	-.178	.153	.002	.263	.103	-.180	-.040	-.124	-.185	.052	.149	
15. Judgment Time	-.094	-.109	-.492**	.313	-.056	.417*	.014	-.442**	.067	-.466**	-.453**	-.045	.325	.519**

Note. N= 36 * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Table R7

Correlations between Measures with Two-Trial Delay Manipulation Data from the Discrete System Condition

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Causal Attribution														
2. System Control	-.359*													
3. Number of Apply Treatments	.105	.224												
4. Number of Runs	.094	.021	-.596**											
5. Selection Diversity	.070	-.035	-.038	-.145										
6. Standard Error	.326	-.294	-.534**	.546**	-.056									
7. Treatment|Elevated	.116	.112	.508**	-.183	.122	-.285								
8. Treatment|Normal	.176	-.014	.642**	-.662**	.100	-.251	-.140							
9. Treatment | Increase	.326	.086	.444**	-.080	-.013	-.194	.825**	-.144						
10. Treatment | Decrease	-.045	.107	.557**	-.602**	.094	-.179	-.121	.860**	-.208					
11. Treatment | No Change	-.078	.365*	.737**	-.314	-.158	-.430**	.222	.418*	.155	.508**				
12. Switch given Unexpected Observation	-.130	-.171	-.099	.333*	.167	-.188	-.038	-.211	-.122	-.157	.132			
13. Value on 16th Trial	.214	.146	.214	-.203	.041	-.031	.010	.226	.032	.270	.167	-.108		
14. Average Trial Time	-.016	.072	-.084	.355*	-.091	.127	.039	-.216	.074	-.154	.078	.364*	.192	
15. Judgment Time	-.130	.033	-.287	.204	.272	.115	-.122	-.256	-.093	-.244	-.113	.147	.235	.491**

Note. N= 36 * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Table R8

Correlations between Measures with Three-Trial Delay Manipulation Data from the Discrete System Condition

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Causal Attribution														
2. System Control	-.497**													
3. Number of Apply Treatments	.426**	-.290												
4. Number of Runs	-.215	.312	-.543**											
5. Selection Diversity	-.025	-.011	.194	-.370*										
6. Standard Error	-.315	.032	-.770**	.211	-.136									
7. Treatment|Elevated	.550**	-.396*	.706**	-.261	.239	-.480**								
8. Treatment|Normal	-.010	-.109	.246	-.472**	-.109	-.245	-.360*							
9. Treatment | Increase	.093	-.249	.341*	-.201	.196	-.237	.543**	-.019						
10. Treatment | Decrease	.127	-.107	.341*	-.507**	-.114	-.325	-.223	.923**	-.128					
11. Treatment | No Change	.227	-.172	.795**	-.397*	.218	-.607**	.556**	.118	.450**	.015				
12. Switch given Unexpected Observation	-.210	.189	-.227	.239	.183	.117	-.138	-.208	-.123	-.235	-.119			
13. Value on 16th Trial	-.104	-.036	.018	.059	.299	.053	.141	-.226	.086	-.280	.048	-.052		
14. Average Trial Time	-.193	.188	.164	.173	.087	-.326	.030	-.012	-.032	-.028	.215	-.036	-.140	
15. Judgment Time	-.337*	.407*	-.347*	.281	-.226	.119	-.394*	-.023	-.193	-.065	-.254	.087	-.213	.211

Note. N= 36 * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Table R9

Correlations between Measures with 25|0 Contingency Manipulation Data from the Continuous System Condition

	1	2	3	4	5	6	7	8	9	10	12	13	14
1. Causal Attribution													
2. System Control	-.296												
3. Number of Apply	.373*	-.327											
4. Number of Runs	-.447**	.220	-.618**										
5. Selection Diversity	-.460**	.358*	-.798**	.599**									
6. Standard Error	-.294	.363*	-.939**	.564**	.660**								
7. Treatment|Elevated	.373*	-.327	1.000**	-.618**	-.798**	-.939**							
8. Treatment|Normal	.c	.c	.c	.c	.c	.c	.c						
9. Treatment | Increase	.349*	-.298	.905**	-.493**	-.686**	-.867**	.905**	.c					
10. Treatment | Decrease	.184	-.293	.528**	-.414*	-.462**	-.596**	.528**	.c	.228				
12. Switch given Unexpected	-.023	-.112	-.020	-.084	-.091	.095	-.020	.c	-.128	.052			
13. Value on 16th Trial	-.427**	.333*	-.822**	.538**	.708**	.796**	-.822**	.c	-.803**	-.371*	.042		
14. Average Trial Time	-.107	.033	.108	.192	-.023	-.112	.108	.c	.024	.259	-.211	.104	
15. Judgment Time	.291	-.133	.204	-.069	-.121	-.239	.204	.c	.092	.333*	.104	-.284	.269

Note. N= 36 * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed). .c Condition does not apply to the 25|0 contingency manipulation in the continuous system condition because the blood pressure value never entered the normal range

Table R10

Correlations between Measures with 75|0 Contingency Manipulation Data from the Continuous System Condition

	1	2	3	4	5	6	7	8	9	10	12	13	14
1. Causal Attribution													
2. System Control	-.109												
3. Number of Apply	.521**	-.371*											
4. Number of Runs	-.315	.328	-.677**										
5. Selection Diversity	-.403*	.381*	-.865**	.659**									
6. Standard Error	-.474**	.277	-.890**	.534**	.651**								
7. Treatment|Elevated	.203	-.302	.718**	-.720**	-.650**	-.617**							
8. Treatment|Normal	.561**	-.289	.747**	-.314	-.590**	-.761**	.314						
9. Treatment | Increase	.244	-.299	.660**	-.153	-.641**	-.578**	.314	.457**					
10. Treatment | Decrease	.528**	-.234	.812**	-.778**	-.623**	-.721**	.718**	.647**	.153				
12. Switch given Unexpected	.038	.106	-.172	.098	.191	.203	-.126	-.167	.012	-.124			
13. Value on 16th Trial	-.448**	.446**	-.929**	.625**	.850**	.809**	-.773**	-.656**	-.696**	-.702**	.153		
14. Average Trial Time	.050	.152	.138	.107	-.129	-.220	.016	.213	.167	.065	-.045	-.097	
15. Judgment Time	-.182	.206	.012	.067	.003	-.098	-.018	.105	.055	-.014	-.036	.023	.501**

Note. N= 36 * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Table R11

Correlations between Measures with 75|75 Contingency Manipulation Data from the Continuous System Condition

	1	2	3	4	5	6	7	8	9	10	12	13	14
1. Causal Attribution													
2. System Control	-.157												
3. Number of Apply	.724**	-.077											
4. Number of Runs	-.112	-.054	-.443**										
5. Selection Diversity	-.215	.004	-.653**	.757**									
6. Standard Error	-.213	-.207	-.351*	.278	.566**								
7. Treatment|Elevated	.461**	-.053	.733**	-.413*	-.502**	-.266							
8. Treatment|Normal	.690**	-.136	.904**	-.336*	-.571**	-.285	.390*						
9. Treatment | Increase	.678**	-.045	.835**	-.454**	-.501**	-.295	.670**	.703**					
10. Treatment | Decrease	.691**	-.087	.958**	-.354*	-.618**	-.313	.641**	.911**	.664**				
12. Switch given Unexpected	-.184	.264	-.081	.133	-.039	-.164	-.153	-.013	-.172	-.045			
13. Value on 16th Trial	-.182	.163	-.173	.143	.050	.231	-.257	-.111	-.352*	-.048	.075		
14. Average Trial Time	.033	-.097	-.030	.118	.040	-.180	-.060	-.037	.004	-.047	.060	-.199	
15. Judgment Time	-.116	.144	-.237	.161	.207	-.004	-.199	-.290	-.173	-.245	.109	.022	.636**

Note. N= 36 * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Table R12

Correlations between Measures with One-Trial Delay Manipulation Data from the Continuous System Condition

	1	2	3	4	5	6	7	8	9	10	12	13	14
1. Causal Attribution													
2. System Control	-.470**												
3. Number of Apply	.374*	.049											
4. Number of Runs	-.148	.148	-.509**										
5. Selection Diversity	-.344*	.300	-.738**	.645**									
6. Standard Error	-.025	-.421*	-.326	-.018	.299								
7. Treatment|Elevated	.275	-.104	.673**	-.514**	-.550**	-.076							
8. Treatment|Normal	.711**	-.721**	.414*	-.160	-.529**	.035	.229						
9. Treatment | Increase	.365*	-.222	.639**	-.119	-.586**	-.083	.276	.488**					
10. Treatment | Decrease	.152	.253	.699**	-.598**	-.417*	-.238	.486**	.127	.096				
12. Switch given Unexpected	-.291	.402*	.144	.001	-.014	-.336*	-.136	-.209	.017	.201			
13. Value on 16th Trial	-.546**	.896**	-.143	.347*	.468**	-.341*	-.281	-.770**	-.331*	.085	.397*		
14. Average Trial Time	.078	-.056	.045	.196	.010	-.064	-.116	.067	.149	.103	-.105	-.028	
15. Judgment Time	-.081	.484**	.002	.125	.247	-.153	-.284	-.305	-.063	.177	.216	.434**	.269

Note. N= 36 * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Table R13

Correlations between Measures with Two-Trial Delay Manipulation Data from the Continuous System Condition

	1	2	3	4	5	6	7	8	9	10	12	13	14
1. Causal Attribution													
2. System Control	-.675**												
3. Number of Apply	.387*	-.121											
4. Number of Runs	-.541**	.526**	-.641**										
5. Selection Diversity	-.497**	.433**	-.763**	.818**									
6. Standard Error	-.336*	-.034	-.888**	.441**	.524**								
7. Treatment|Elevated	.357*	-.099	.783**	-.563**	-.625**	-.834**							
8. Treatment|Normal	.758**	-.734**	.509**	-.569**	-.557**	-.415*	.262						
9. Treatment | Increase	.482**	-.196	.864**	-.581**	-.650**	-.737**	.741**	.418*					
10. Treatment | Decrease	.292	-.112	.833**	-.635**	-.629**	-.784**	.635**	.472**	.597**				
12. Switch given Unexpected	-.265	.271	-.178	.192	.273	.018	-.009	-.330*	-.151	-.123			
13. Value on 16th Trial	-.707**	.946**	-.184	.576**	.510**	-.003	-.156	-.764**	-.278	-.133	.272		
14. Average Trial Time	-.193	.083	-.044	.217	.047	-.029	-.054	-.037	-.064	-.063	-.004	.030	
15. Judgment Time	-.258	.063	-.066	.124	-.015	.141	-.029	-.205	-.006	-.129	.232	.004	.573**

Note. N= 36 * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Table R14

Correlations between Measures with Three-Trial Delay Manipulation Data from the Continuous System Condition

	1	2	3	4	5	6	7	8	9	10	12	13	14
1. Causal Attribution													
2. System Control	-.412*												
3. Number of Apply	.673**	-.002											
4. Number of Runs	-.184	.191	-.396*										
5. Selection Diversity	-.357*	.384*	-.545**	.660**									
6. Standard Error	-.171	-.220	-.409*	.213	.323								
7. Treatment|Elevated	.518**	-.041	.874**	-.455**	-.457**	-.434**							
8. Treatment|Normal	.708**	-.564**	.612**	-.234	-.597**	-.151	.444**						
9. Treatment | Increase	.470**	.153	.857**	-.372*	-.409*	-.395*	.773**	.381*					
10. Treatment | Decrease	.619**	-.214	.792**	-.317	-.553**	-.248	.702**	.614**	.444**				
12. Switch given Unexpected	-.020	.030	-.084	.143	-.045	.078	-.201	-.084	-.042	-.053			
13. Value on 16th Trial	-.447**	.909**	-.102	.198	.477**	-.088	-.113	-.680**	.049	-.281	.028		
14. Average Trial Time	.145	.063	.160	-.005	-.169	-.355*	.242	.122	.079	.219	-.017	-.052	
15. Judgment Time	-.016	.109	-.049	-.012	.059	-.128	.160	-.117	-.111	.089	-.044	.076	.492**

Note. N= 36 * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

APPENDIX S

Correlations between Outcome Dependent Measures, and Intervention Strategy and Outcome Dependent Measures

I performed a series of Pearson product-moment correlational analyses to better understand the interdependencies between the dependent measures. Here, I focus on relationships between dependent measures describing participants' reasoning outcomes (i.e., causal attributions and ability to control the system), as well as their relationship to intervention strategy to justify further exploration into differences in reasoning process and the implication of reasoning process on reasoning outcome (see Appendix R for the comprehensive set of correlational analyses reviewing the relationships between all of the dependent measures). I performed these analyses across experimental manipulations, as well as separated by contingency and delay, which ensures data independence.

Causal Attributions and System Control

Table S1 contains the results of correlational analyses between causal attributions and system control. With both the discrete and continuous system conditions, causal attributions are negatively correlated with the across experimental manipulation analysis that violated the assumption of data independence. This relationship only persists with the 75|0 by three-trial delay when the data is separated by experimental combination.

Table S1
Correlations between Number of Causal Attribution and System Control Measures

	Discrete System		Continuous System	
	r	*p*	*r*	*p*
Across Experimental Manipulations	-.431	< .001	-.532	< .001
25|0 and				
One-trial delay	-.141	.663	-.487	.108
Two-trial delay	.033	.919	-.131	.686
Three-trial delay	-.305	.335	-.569	.054
75|0 and				
One-trial delay	-.315	.318	-.022	.947
Two-trial delay	-.090	.782	-.088	.785
Three-trial delay	-.815	.001	-.125	.698
75|75 and				
One-trial delay	-.379	.225	-.125	.698
Two-trial delay	.306	.334	.183	.568
Three-trial delay	.095	.769	-.487	.108

Note. $N = 108$ for the across experimental manipulations analysis; $n = 12$ for each analyses. All tests two-tailed.

Causal Attributions

Tables S2 and S3 contains the results of correlational analyses between the number of apply treatment selections and the number of runs relative to participants' causal attribution ratings (i.e., the reasoning outcome dependent measure used within the causal reasoning literature). With both the discrete and continuous system conditions, the number of apply treatment interventions appears to be positively correlated with causal attributions. In contrast, the number of runs appears to negatively relate, or have no relationship to participants' causal attributions.

Table S2
Correlations between Number of Apply Treatment Interventions and Causal Attribution Measures

	Discrete System		Continuous System	
	r	*p*	*r*	*p*
Across Experimental Manipulations	.259	.007	.489	.000
25|0	.228	.182	.373	.025
75|0	.530	.001	.521	.001
75|75	.372	.025	.724	.000
One-trial delay	.213	.213	.374	.025
Two-trial delay	.105	.542	.387	.020
Three-trial delay	.426	.010	.673	.000

Note. $N = 108$ for the across experimental manipulations analysis; $n = 36$ for the remaining analyses. All tests two-tailed.

Table S3
Correlations between Number of Runs and Causal Attribution Measures

	Discrete System		Continuous System	
	r	*p*	*r*	*p*
Across Experimental Manipulations	-.184	.057	-.277	.004
25|0	-.128	.456	-.447	.006
75|0	-.450	.006	-.315	.062
75|75	-.036	.836	-.112	.514
One-trial delay	-.342	.041	-.148	.391
Two-trial delay	.094	.584	-.541	.001
Three-trial delay	-.215	.209	-.184	.283

Note. $N = 108$ for the across experimental manipulations analysis; $n = 36$ for the remaining analyses. All tests two-tailed.

Control of the System

Table S4 and S5 contains the results of correlational analyses between the number of apply treatment selections and the number of runs relative to participants' blood pressure level after 13 treatment selections. Participants' ability to control the system

was negatively correlated to the number of apply treatment applications with the 25|0 and 75|0 contingency manipulations with the discrete system condition, but was not significantly correlated across the experimental manipulations. This is a rather unsurprising finding as the only manner in which the blood pressure observation can lower is by applying the treatment with these two contingency manipulations. A similar pattern of findings was true relative to the continuous system, however the relationship between participants treatment applications and their ability to control the system was only significant with the 75|0 contingency manipulation.

The relationship between the number of runs and the system control measure differed between the system conditions. With the discrete system condition, the number of runs was positively correlated with their ability to control the system, meaning the less participants switched between the treatment options resulted in lower blood pressure observations after 13 treatment interventions, only with the 25|0 and 75|0 contingency manipulations and not across the experimental manipulations. Conversely, the number of runs was positively correlated with system control across experimental manipulations with the continuous system, but not significantly correlated when the data was separated by contingency. With the continuous system, the only significant correlation with the data separated by experimental manipulations was with the two-trial delay manipulation.

Table S4
Correlations between Number of Apply Treatment Interventions and System Control Measures

	Discrete System		Continuous System	
	r	*p*	*r*	*p*
Across Experimental Manipulations	.021	.826	-.039	.689
25|0	-.907	.000	-.327	.051
75|0	-.934	.000	-.371	.026
75|75	.043	.802	-.077	.656
One-trial delay	.136	.429	.049	.777
Two-trial delay	.224	.189	-.121	.481
Three-trial delay	-.290	.086	-.002	.990

Note. $N = 108$ for the across experimental manipulations analysis. $N = 36$ for the remaining analyses. All tests two-tailed.

Table S5
Correlations between Number of Runs and System Control Measures

	Discrete System		Continuous System	
	r	*p*	*r*	*p*
Across Experimental Manipulations	.145	.135	.270	.005
25|0	.373	.025	.220	.197
75|0	.717	.000	.328	.051
75|75	.100	.562	-.054	.753
One-trial delay	.067	.698	.148	.388
Two-trial delay	.021	.904	.526	.001
Three-trial delay	.312	.064	.191	.266

Note. $N = 108$ for the across experimental manipulations analysis. $N = 36$ for the remaining analyses. All tests two-tailed.

APPENDIX T

Multiple Regressions Assessing the Influence of System Control on Causal Attribution

I performed a series of multiple regressions to assess the influence of system control on causal attribution, while controlling for delay, with the data separated by system condition and levels of contingency. The outcomes of the multiple regressions are the same as the correlations reported in the main body of the document. Table T1 and Table T2 contain the outcomes pertaining to the discrete and continuous system conditions, respectively.

Table T1

Regression Analyses between System Control and Causal Attribution with the Discrete System Condition

						95% CI	
	b	*SE*	β	*t*	*p*	LL	UL
25|0	-1.27	1.30	-0.17	-0.97	0.34	-3.92	1.39
75|0	-1.63	0.50	-0.49	-3.25	< 0.01	-2.65	-0.61
75|75	-0.23	1.94	-0.02	-0.12	0.91	-4.17	3.72

Note. CI = Confidence Interval; LL= lower limit; UL=upper limit.

Table T2

Regression Analyses between System Control and Causal Attribution with the Continuous System Condition

						95% CI	
	b	*SE*	β	*t*	*p*	LL	UL
25|0	-2.11	1.22	-0.30	-1.74	0.09	-4.59	0.36
75|0	-0.13	0.26	-0.09	-0.51	0.62	-0.66	0.39
75|75	-1.35	2.32	-0.11	-0.58	0.56	-6.07	3.37

Note. CI = Confidence Interval; LL= lower limit; UL=upper limit.

APPENDIX U

Strategy Shift within Learning Tasks

To assess whether participants shifted their intervention strategy in the midst of individual learning tasks (or patient scenarios), I compared their intervention selections on trials 1 through 8 with their intervention selections on trials 9 through 16 using Fisher's Exact Tests. I was unable to calculate a *p* value using the Fisher's Exact Test with several of the observations because zero instances of one or more intervention types within were observed. To account for this limitation, I performed additional tests using adjusted data, specifically I added to one to every cell. Table U1 summarizes the findings, which suggest that relatively few participants switched their treatment intervention strategy half way through the patient scenarios. Tables U2, U3, and U4 contain the analyses of the discrete system condition. Tables U5, U6, and U7 present the analyses of the continuous system condition data.

Table U1
Summary of the Fisher Exact Test Assessing Strategy Shifts within Leaning Tasks

System Condition	Same	Different	NA
Discrete	76 (107)	1 (1)	31
Continuous	55 (98)	2 (6)	51

Note. Same = Same intervention strategy used on trials 1-8 than on trials 9-16 given $p > .05$. Different = different intervention strategy used on trials 1-8 than on trials 9-16 given $p < .05$. NA = not able to perform test because of a zero in one or more of the cells. Parentheses summarize Fisher's Exact test calculated with adjusted data (i.e., +1 was added to every cell).

Table U2

Strategy Shift within a Patient Scenario - Fisher's Exact Tests related to the 25|0 Contingency Manipulation in the Discrete System Condition separated by Delay

	Trials 1-8		Trials 9-16			
	Treatment	~Treatment	Treatment	~Treatment	*p*	*p***
One-trial delay	2	6	1	7	1.00	1.00
	7	1	7	1	1.00	1.00
	7	1	4	4	0.28	0.35
	6	2	6	2	1.00	1.00
	5	3	8	0	NA	0.30
	7	1	6	2	1.00	1.00
	4	4	7	1	0.28	0.35
	8	0	7	1	NA	1.00
	8	0	5	3	NA	0.30
	8	0	8	0	NA	1.00
	6	2	7	1	1.00	1.00
	5	3	6	2	1.00	1.00
Two-trial delay	7	1	5	3	0.57	0.63
	7	1	5	3	0.57	0.63
	6	2	7	1	1.00	1.00
	4	4	6	2	0.61	0.65
	6	2	7	1	1.00	1.00
	6	2	8	0	NA	0.58
	4	4	8	0	NA	0.14
	6	2	4	4	0.61	0.65
	8	0	8	0	NA	1.00
	8	0	8	0	NA	1.00
	5	3	4	4	1.00	1.00
	7	1	4	4	0.28	0.35
Three-trial delay	4	4	4	4	1.00	1.00
	3	5	7	1	0.12	0.17
	4	4	0	8	NA	0.14
	6	2	5	3	1.00	1.00
	4	4	2	6	0.61	0.65
	6	2	6	2	1.00	1.00
	8	0	5	3	NA	0.30
	7	1	6	2	1.00	1.00
	7	1	6	2	1.00	1.00
	4	4	2	6	0.61	0.65
	4	4	2	6	0.61	0.65
	7	1	6	2	1.00	1.00

Note. NA = not able to perform test because of a zero in one or more of the cells. ***p* = Fisher's Exact test calculated with adjusted data (i.e., +1 was added to every cell).

Table U3

Strategy Shift within a Patient Scenario - Fisher's Exact Tests related to the 75|0 Contingency Manipulation in the Discrete System Condition separated by Delay

	Trials 1-8		Trials 9-16			
	Treatment	~Treatment	Treatment	~Treatment	*p*	***p*
One-trial delay	3	5	6	2	0.31	0.37
	6	2	6	2	1.00	1.00
	5	3	7	1	0.57	0.63
	7	1	8	0	NA	1.00
	7	1	8	0	NA	1.00
	5	3	4	4	1.00	1.00
	8	0	8	0	NA	1.00
	6	2	8	0	NA	0.58
	5	3	4	4	1.00	1.00
	5	3	6	2	1.00	1.00
	7	1	0	8	NA	1.00
	7	1	5	3	0.57	0.63
Two-trial delay	5	3	5	3	1.00	1.00
	4	4	5	3	1.00	1.00
	5	3	5	3	1.00	1.00
	8	0	6	2	NA	0.58
	5	3	5	3	1.00	1.00
	4	4	7	1	0.28	0.35
	8	0	8	0	NA	1.00
	5	3	5	3	1.00	1.00
	4	4	8	0	NA	0.14
	5	3	4	4	1.00	1.00
	4	4	4	4	1.00	1.00
	5	3	5	3	1.00	1.00
Three-trial delay	2	6	3	5	1.00	1.00
	4	4	6	2	0.61	0.65
	6	2	3	5	0.31	0.37
	6	2	6	2	1.00	1.00
	7	1	4	4	0.28	0.35
	5	3	4	4	1.00	1.00
	5	3	3	5	0.62	0.69
	8	0	8	0	NA	1.00
	5	3	4	4	1.00	1.00
	4	4	3	5	1.00	1.00
	8	0	5	3	NA	0.30
	7	1	8	0	NA	1.00

Note. NA = not able to perform test because of a zero in one or more of the cells. ***p* = Fisher's Exact test calculated with adjusted data (i.e., +1 was added to every cell).

Table U4

Strategy Shift within a Patient Scenario - Fisher's Exact Tests related to the 75|75 Contingency Manipulation in the Discrete System Condition separated by Delay

	Trials 1-8		Trials 9-16			
	Treatment	~Treatment	Treatment	~Treatment	*p*	***p*
One-trial	3	5	2	6	1.00	1.00
delay	3	5	2	6	1.00	1.00
	5	3	6	2	1.00	1.00
	4	4	4	4	1.00	1.00
	8	0	8	0	NA	1.00
	5	3	5	3	1.00	1.00
	8	0	8	0	NA	1.00
	3	5	2	6	1.00	1.00
	3	5	2	6	1.00	1.00
	8	0	8	0	NA	1.00
	3	5	2	6	1.00	1.00
	3	5	2	6	1.00	1.00
Two-	4	4	7	1	0.28	0.35
trial	7	1	8	0	NA	1.00
delay	5	3	3	5	0.62	0.69
	1	7	0	8	NA	1.00
	4	4	1	7	0.28	0.35
	8	0	7	1	NA	1.00
	8	0	8	0	NA	1.00
	5	3	7	1	0.57	0.63
	1	7	3	5	0.57	0.63
	5	3	2	6	0.31	0.37
	4	4	3	5	1.00	1.00
	3	5	3	5	1.00	1.00
Three-	5	3	5	3	1.00	1.00
trial	4	4	3	5	1.00	1.00
delay	5	3	3	5	0.62	0.69
	4	4	4	4	1.00	1.00
	7	1	7	1	1.00	1.00
	8	0	8	0	NA	1.00
	8	0	3	5	NA	0.06
	4	4	2	6	0.61	0.65
	7	1	8	0	NA	1.00
	7	1	3	5	0.12	0.17
	4	4	2	6	0.61	0.65
	7	1	1	7	0.01	0.02

Note. NA = not able to perform test because of a zero in one or more of the cells. ***p* = Fisher's Exact test calculated with adjusted data (i.e., +1 was added to every cell).

Table U5

Strategy Shift within a Patient Scenario - Fisher's Exact Tests related to the 25|0 Contingency Manipulation in the Continuous System Condition separated by Delay

	Trials 1-8		Trials 9-16			
	Treatment	~Treatment	Treatment	~Treatment	*p*	***p*
One-trial	4	4	6	2	0.61	0.65
delay	7	1	5	3	0.57	0.63
	6	2	6	2	1.00	1.00
	8	0	6	2	NA	0.58
	5	3	4	4	1.00	1.00
	3	5	4	4	1.00	1.00
	2	6	8	0	NA	0.02
	8	0	7	1	NA	1.00
	6	2	8	0	NA	0.58
	6	2	6	2	1.00	1.00
	7	1	5	3	0.57	0.63
	6	2	7	1	1.00	1.00
Two-	2	6	4	4	0.61	0.65
trial	5	3	5	3	1.00	1.00
delay	8	0	8	0	NA	1.00
	4	4	7	1	0.28	0.35
	4	4	6	2	0.61	0.65
	6	2	5	3	1.00	1.00
	4	4	8	0	NA	0.14
	4	4	5	3	1.00	1.00
	7	1	6	2	1.00	1.00
	8	0	6	2	NA	0.58
	5	3	5	3	1.00	1.00
	4	4	4	4	1.00	1.00
Three-	6	2	5	3	1.00	1.00
trial	6	2	8	0	NA	0.58
delay	5	3	6	2	1.00	1.00
	5	3	8	0	NA	0.30
	8	0	5	3	NA	0.30
	4	4	4	4	1.00	1.00
	3	5	3	5	1.00	1.00
	6	2	3	5	0.31	0.37
	5	3	4	4	1.00	1.00
	6	2	4	4	0.61	0.65
	5	3	3	5	0.62	0.67
	7	1	4	4	0.28	0.35

Note. NA = not able to perform test because of a zero in one or more of the cells. ***p* = Fisher's Exact test calculated with adjusted data (i.e., +1 was added to every cell).

Table U6

Strategy Shift within a Patient Scenario - Fisher's Exact Tests related to the 75|0 Contingency Manipulation in the Continuous System Condition separated by Delay

	Trials 1-8		Trials 9-16			
	Treatment	~Treatment	Treatment	~Treatment	*p*	***p*
One-trial	5	3	8	0	NA	0.30
delay	5	3	4	4	1.00	1.00
	8	0	8	0	NA	1.00
	8	0	6	2	NA	0.58
	4	4	6	2	0.61	0.65
	5	3	4	4	1.00	1.00
	7	1	4	4	0.28	0.35
	4	4	4	4	1.00	1.00
	5	3	6	2	1.00	1.00
	6	2	6	2	1.00	1.00
	8	0	8	0	NA	1.00
	8	0	8	0	NA	1.00
Two-	2	6	8	0	NA	0.02
trial	8	0	7	1	NA	1.00
delay	8	0	0	8	NA	< .01
	8	0	8	0	NA	1.00
	5	3	6	2	1.00	1.00
	8	0	5	3	NA	0.30
	7	1	5	3	0.57	0.63
	8	0	5	3	NA	0.30
	5	3	8	0	NA	0.30
	4	4	2	6	0.61	0.65
	6	2	7	1	1.00	1.00
	8	0	8	0	NA	1.00
Three-	8	0	3	5	NA	0.06
trial	7	1	8	0	NA	1.00
delay	5	3	5	3	1.00	1.00
	8	0	8	0	NA	1.00
	7	1	8	0	NA	1.00
	4	4	8	0	NA	0.14
	3	5	7	1	0.12	0.17
	4	4	2	6	0.61	0.65
	8	0	8	0	NA	1.00
	7	1	1	7	0.01	0.02
	5	3	4	4	1.00	1.00
	6	2	8	0	NA	0.58

Note. NA = not able to perform test because of a zero in one or more of the cells. ***p* = Fisher's Exact test calculated with adjusted data (i.e., +1 was added to every cell).

Table U7

Strategy Shift within a Patient Scenario - Fisher's Exact Tests related to the 75|75 Contingency Manipulation in the Continuous System Condition separated by Delay

	Trials 1-8		Trials 9-16			
	Treatment	~Treatment	Treatment	~Treatment	*p*	***p*
One-trial	3	5	6	2	0.31	0.37
delay	8	0	5	3	NA	0.30
	0	8	6	2	NA	0.02
	8	0	8	0	NA	1.00
	7	1	1	7	0.01	0.02
	8	0	4	4	NA	0.14
	8	0	8	0	NA	1.00
	5	3	0	8	NA	0.30
	8	0	6	2	NA	0.58
	7	1	6	2	1.00	1.00
	3	5	2	6	1.00	1.00
	8	0	7	1	NA	1.00
Two-	4	4	1	7	0.28	0.35
trial	8	0	8	0	NA	1.00
delay	8	0	8	0	NA	1.00
	8	0	8	0	NA	1.00
	1	7	8	0	NA	< .01
	8	0	8	0	NA	1.00
	1	7	8	0	NA	< .01
	3	5	5	3	0.62	0.67
	8	0	1	7	NA	< .01
	8	0	5	3	NA	0.30
	1	7	2	6	1.00	1.00
	1	7	4	4	0.28	0.35
Three-	3	5	4	4	1.00	1.00
trial	6	2	8	0	NA	0.58
delay	6	2	0	8	NA	0.02
	5	3	3	5	0.62	0.67
	4	4	7	1	0.28	0.35
	4	4	2	6	0.61	0.65
	4	4	5	3	1.00	1.00
	7	1	8	0	NA	1.00
	2	6	1	7	1.00	1.00
	4	4	6	2	0.61	0.65
	8	0	8	0	NA	1.00
	1	7	0	8	NA	1.00

Note. NA = not able to perform test because of a zero in one or more of the cells. ***p* = Fisher's Exact test calculated with adjusted data (i.e., +1 was added to every cell).

APPENDIX V

Strategy Shift between Learning Tasks

Tables V1 and V2 present the chi-square goodness-of-fit analyses that I used to assess if intervention strategy was consistent between the three learning task/sequence orders with both the discrete and continuous system conditions. Indeed, it was consistent as $p > .05$ with all of the analyses.

Table V1
Distribution of Treatment Selection Processes across Learning Tasks with the Discrete System Condition

	1st	*2nd*	*3rd*	X^2	*p*	*V*
Treatment-biased	12	15	18	1.20	.55	.12
Unbiased	24	19	16	1.66	.44	.12
Withhold-bias	0	2	2	2.00	.37	.50
Observation-independent	12	17	16	0.93	.63	.10
Observation-dependent	24	19	20	0.67	.72	.07

Note. 1st, 2nd, and 3rd = order within the 3 learning tasks. *V = Cramer's V.*

Table V2
Distribution of Treatment Selection Processes across Learning Tasks with the Continuous System Condition

	1st	*2nd*	*3rd*	X^2	*p*	*V*
Treatment-biased	16	18	16	0.16	.92	.04
Unbiased	19	17	19	0.15	.93	.04
Withhold-bias	1	1	1	0.00	1.00	.00
Observation-independent	22	28	24	0.76	.68	.07
Observation-dependent	14	8	12	1.65	.44	.16

Note. 1st, 2nd, and 3rd = order within the 3 learning tasks. *V = Cramer's V.*

APPENDIX W

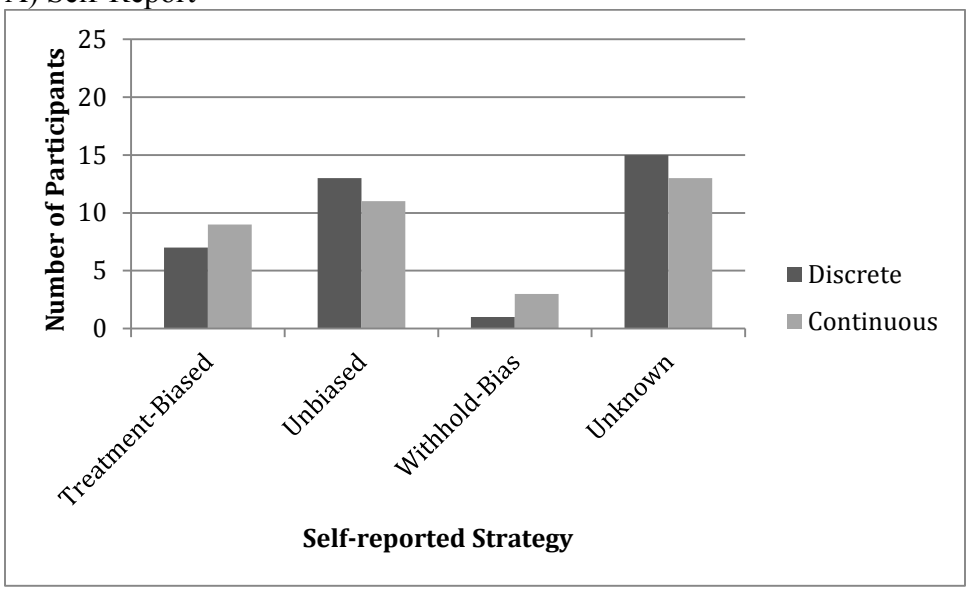
Self-report of Intervention Strategy

Categorization of Responses

I asked participants to describe the intervention strategy that they used in combination with their third patient scenario. My intent was to use these self-reports to validate the findings of my strategy analysis. With each participant's response, I looked for words and phases to categorize their intervention strategies using the same classification schemes that I used for the strategy analysis (i.e., intervention bias and observation dependence). When participants did not refer to either observation dependence or treatment biases,[24] I categorized the self-reported strategy as *unknown.* Figures W1 and W2 depict the distribution of intervention strategies with the third patient scenario based on participants' self-reported and observationally derived data.

[24] I coded observation dependence and intervention bias separately using the same response. Given this, a participant's self-reported intervention strategy may have been categorized as unknown relative to observation dependence, but categorized as treatment-biased/unbiased/withhold-biased relative to intervention biases and vice versa.

A) Self-Report



B) Observational Strategy Analysis

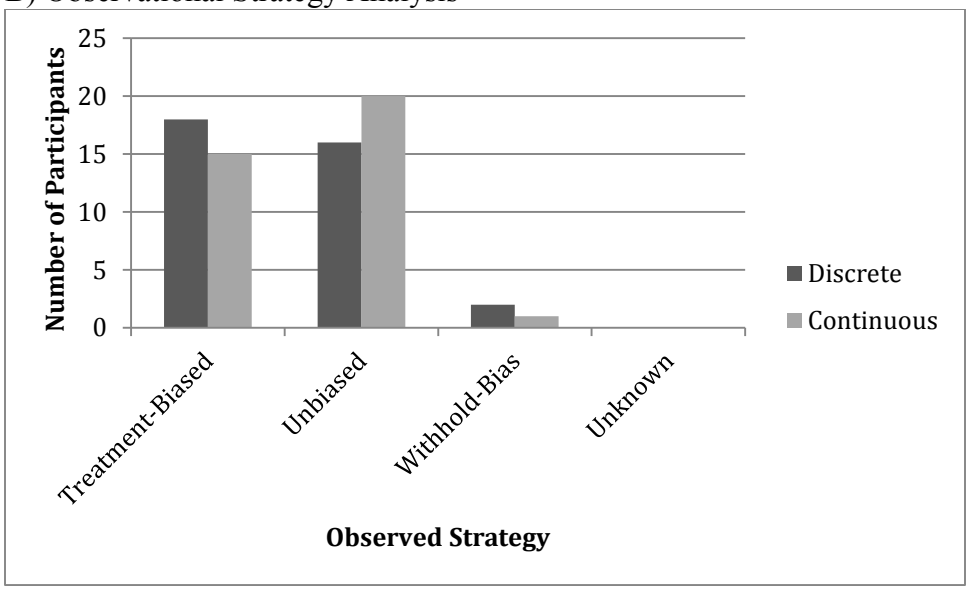
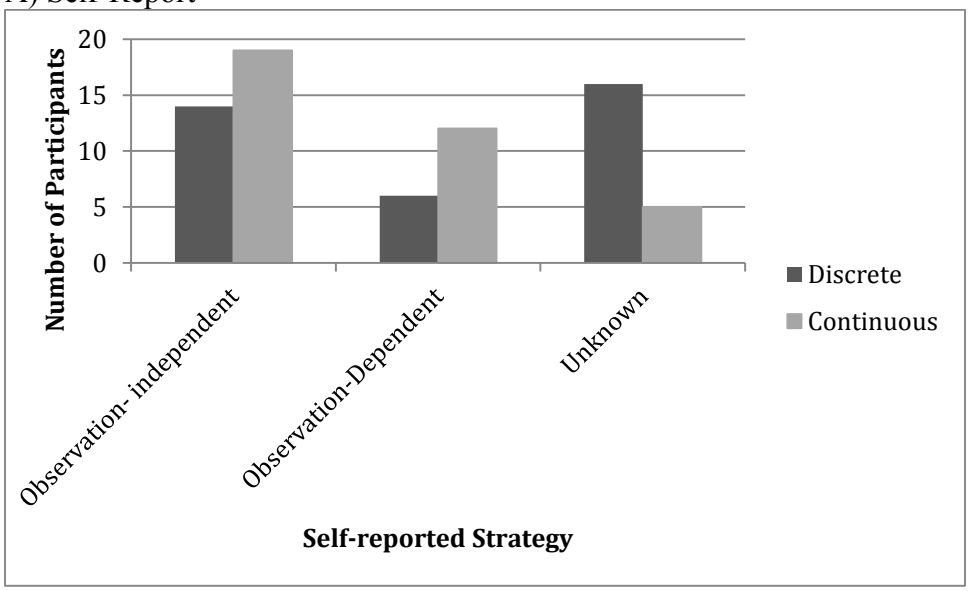


Figure W1. Comparison of intervention bias between self-reported intervention strategies and the results of the observational strategy analysis. The information contained in these distributions only refers to participants' interaction with the third patient scenario. $N = 72$.

A) Self-Report



B) Observational Strategy Analysis

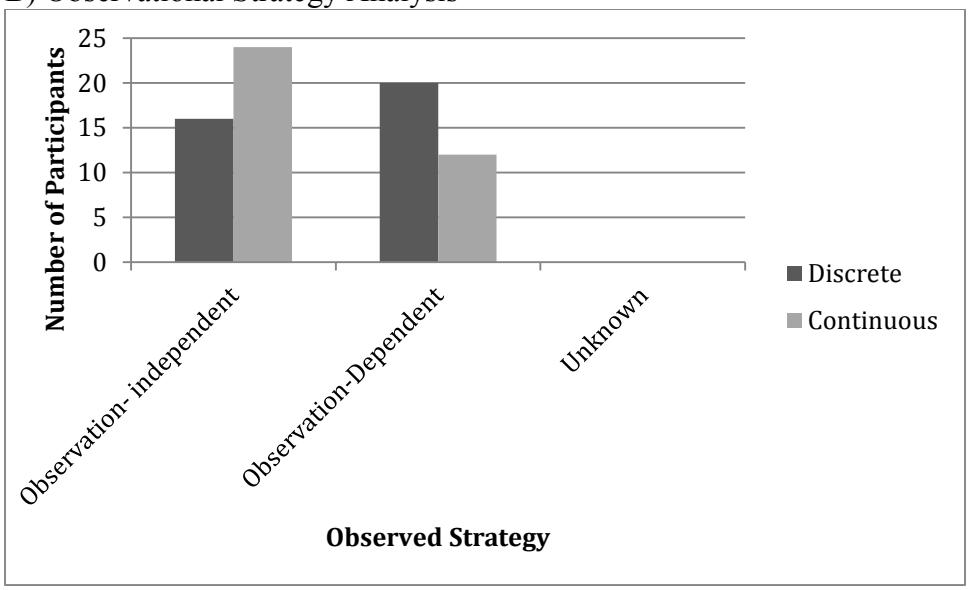


Figure W2. Comparison of observation dependence between self-reported intervention strategies and the results of the observational strategy analysis. The information contained in these distributions only refers to participants' interaction with the third patient scenario. $N = 72$.

Comparison to Observational Strategy Analysis

Next, I assessed whether the treatment intervention strategies that participants used with the third patient scenario were distributed in the same manner according to the self-report and observationally derived analyses. Unfortunately, the substantial number

of self-report responses that were coded as *unknown* prohibited a direct comparison of the distributions. Given this, I compared the distributions using correlation coefficients calculated from chi-square values describing the degree to which intervention strategies were evenly distributed across the strategy categorizations. I excluded the unknown categorization when calculating the chi-square values for the self-report analyses, as this category was distinct to the self-report analyses. The outcome of these analyses indicate that there were no significant differences between the self-reported and the observationally derived data distributions (Table W1 and W2), which validates both the type of categorizations (observation dependence and intervention bias) and the quantitative techniques used within the observationally derived strategy analysis.

Table W1
Comparison of Self-reported and Observationally Derived Intervention Strategies with the Discrete System Condition

	n	$X^2(1)$	*r*	*z*	*p*
Intervention Bias					
Self -report	21	10.29	0.70	0.63	0.53
Observed	36	12.67	0.59		
Observation Dependence					
Self -report	20	3.20	0.40	1.05	0.29
Observed	36	0.44	0.11		

Note. Self-report did not consider unknown categorization; *p*-value reflects a two-tailed analysis.

Table W2
Comparison of Self-reported and Observationally Derived Intervention Strategies with the Continuous System Condition

	n	$X^2(1)$	*r*	*z*	*p*
Intervention Bias					
Self-report	23	4.52	0.44	-1.18	0.24
Observed	36	16.17	0.67		
Observation Dependence					
Self-report	31	1.58	0.23	-0.45	0.65
Observed	36	4.00	0.33		

Note. Self-report did not consider unknown categorization; *p*-value reflects a two-tailed analysis.

Strategy Shift across Patient Scenarios

Less than half of the participants in the discrete (n = 16) and continuous (n = 12) system conditions reported altering their treatment intervention strategy across the three patient scenarios. Sixteen of these participants reported that they adjusted their treatment intervention strategy based on the underlying experimental manipulation combination (i.e., blood pressure behavior related to each patient scenario). Eight participants adjusted their treatment intervention strategy because they felt it would help them to realize how the treatment was influencing the patient's blood pressure, and the remaining three participants adapted their treatment intervention strategy based on their expectations about type of treatment (exercise, healthy diet, or medication).

APPENDIX X

Effect of Contingency and Delay on Intervention Strategy with Data Confounded by System Condition

Effect of Contingency on Intervention Strategy

I used Fisher's Exact Tests to assess the consistency of participants' intervention bias and observation dependence relative to changes in contingency. For these analyses, I separated the data according to the system condition, contingency, and delay to ensure data independence given the experimental design (i.e., each participant interacted with each delay and contingency manipulation in various combinations) and because strategic differences between the system conditions discussed in the results section.

I expected the contingency manipulations would affect intervention strategy differently according to the respective differences in either *P(e|c)* or *P(e|~c)*. I predicted that participants would be more reactive to changes in their patients' blood pressure observations and would apply the treatment selection less frequently as *P(e|c)* declined (i.e., 75|0 to 25|0) and as *P(e|~c)* increased (i.e., 75|0 to 75|75). Participants' intervention strategy only partially reflected the anticipated differences in contingency. Specifically, participants in the continuous system condition adjusted their observation dependence between the 25|0 and 75|0 contingency manipulations. Differences in strategy between the 25|0 and 75|75 contingency manipulations with both systems were unexpected, but reflect the fact the participants treated the 75|0 and 75|75 contingency manipulations in largely the same manner.

Discrete system. More participants applied unbiased treatment intervention strategies with the 75|75 contingency manipulation than with the 25|0 contingency manipulation with the discrete system condition. This was the only significant difference in intervention strategy with the discrete system condition. The same pattern was observed between the 75|0 and 25|0 contingency manipulations, but the difference did not reach significance. Participants' observation dependence was consistent despite changes in contingency with the discrete system condition. Table X1 presents the distribution of treatment intervention strategies and the results of the Fisher's Exact Tests used to assess differences as a function of the contingency manipulations with the discrete system condition.

Table X1
Distribution of Intervention Strategy by Contingency with the Discrete System Condition and Outcomes of Fisher's Exact Tests

	25|0	75|0	*p*	75|0	75|75	*p*	25|0	75|75	*p*
Unbiased	12	22	0.05[1]	22	24	0.45	12	24	< .01
Treatment-biased	21	14		14	10		21	10	
Independent	14	16	0.81	16	18	1.00	14	18	0.64
Dependent	22	20		20	21		22	21	

Note. Independent = Observation-Independent Intervention Strategy, Dependent = Observation-Dependent Intervention Strategy. All tests two-tailed.
[1] $p = 0.0547$.

Continuous system. Participants' intervention bias and observation dependence was evenly distributed across the contingency manipulations with a couple exceptions involving participants' observation dependence. Participants applied more observation-independent strategies with the 75|0 and 75|75 contingency manipulations relative to the 25|0 contingency manipulation. There were no differences in intervention bias with the continuous system condition with the data was separated by contingency. Table X2 presents the distribution of strategies and the results of the Fisher's Exact Tests used to

assess differences in strategy as a function of contingency with the continuous system condition.

Table X2
Distribution of Intervention Strategy by Contingency with the Continuous System Condition and Outcomes of Fisher's Exact Tests

	25|0	75|0	*p*	75|0	75|75	*p*	25|0	75|75	*p*
Unbiased	21	16	0.25	16	18	0.47	21	18	0.81
Treatment-biased	15	20		20	15		15	15	
Independent	16	30	<.01	30	28	0.77	16	28	<.01
Dependent	20	6		6	8		20	8	

Note. Independent = Observation-Independent Intervention Strategy, Dependent = Observation-Dependent Intervention Strategy. All tests two-tailed.

Effect of Delay on Intervention Strategy

I used Fisher's Exact Tests to assess whether participants intervention bias and observation dependence varied in combination with changes in delay. I expected longer delay intervals to result in more reactive, unstructured treatment selection behavior. Specifically, I expected participants to explore the withholding the treatment to a greater degree, which in turn would reduce the number of treatment selections participants applied in the learning task. Also, I anticipated that participants would become more sensitive to their patients' blood pressure observations with longer delay intervals and adapt their treatment selection processes in response. However, the results indicate intervention strategy did not differ as a function delay with either system condition. Tables X3 and X4 contains the distribution of intervention bias and observation dependence in relation to the delay manipulations and the results of the Fisher's Exact Tests used to assess whether intervention strategy differed between the delay manipulations.

Table X3

Distribution of Intervention Strategy by Delay with the Discrete System Condition and Outcomes of Fisher's Exact Tests

	One-trial	Two-trial	*p*	Two-trial	Three-Trial	*p*	One-trial	Three-Trial	*p*
Unbiased	17	19	0.63	19	23	0.46	17	23	0.23
Treatment-biased	18	15		15	12		18	12	
Independent	19	11	0.06	11	15	0.46	19	15	0.36
Dependent	17	25		25	21		17	21	

Note. One-trial = One-trial delay. Two-Trial = Two-Trial Delay. Three-trial = Three-trial delay. Independent = Observation-Independent Intervention Strategy, Dependent = Observation-Dependent Intervention Strategy. All tests two-tailed.

Table X4

Distribution of Intervention Strategy by Delay with the Continuous System Condition and Outcomes of Fisher's Exact Tests

	One-trial	Two-trial	*p*	Two-trial	Three-Trial	*p*	One-trial	Three-Trial	*p*
Unbiased	15	18	0.48	18	22	0.33	15	22	0.06
Treatment-biased	21	17		17	12		21	12	
Independent	25	23	0.80	23	26	0.46	25	26	1.00
Dependent	11	13		13	10		11	10	

Note. One-trial = One-trial delay. Two-Trial = Two-Trial Delay. Three-trial = Three-trial delay. Independent = Observation-Independent Intervention Strategy, Dependent = Observation-Dependent Intervention Strategy. All tests two-tailed.

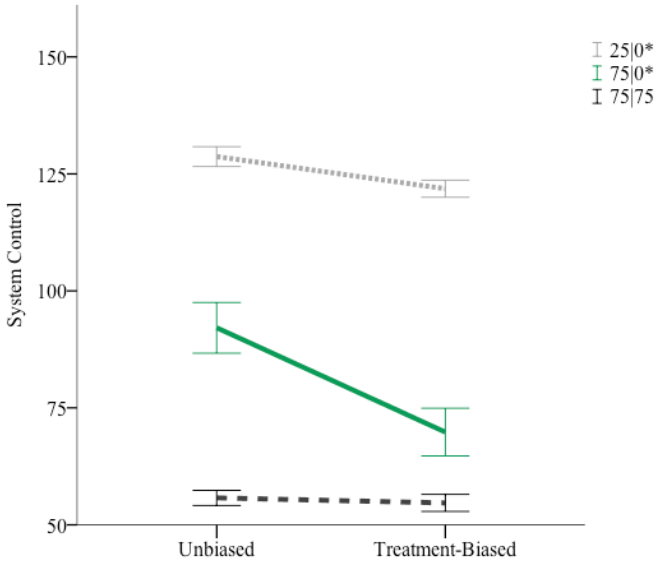
APPENDIX Y

Effect of Intervention Strategy on System Control

Contingency

By definition, the contingency manipulation constrains the possible influence of intervention on outcome (this is discussed in more detail within the proposed paradigm under observation dependence). Specifically, participants rarely observed an effect with the 25|0 contingency manipulation and frequently observed an effect with the 75|75 contingency manipulation regardless of intervention. In contrast, the 75|0 contingency manipulation presented more opportunity for intervention to alter the participants' ability to control of the system. So, unsurprisingly, participants' ability to control the system did differ in conjunction with intervention strategy with the 75|0 contingency manipulation, but not with the other contingency manipulations (Figures Y1a and Y1b). Participants' that applied treatment-biased and observation-independent intervention strategies lowered their patients' blood pressure levels to a greater degree than participants' that applied unbiased and observation-dependent intervention strategies with the 75|0 contingency manipulation. This is because both treatment-biased and observation-independent intervention strategies are associated with higher rates of treatment interventions, and application of the treatment intervention is the only way to lower the blood pressure level with the 75|0 contingency manipulation as there is no value associated with withholding the treatment intervention. Table Y1 presents the simple effect findings.

A) Intervention Bias



B) Observation Dependence

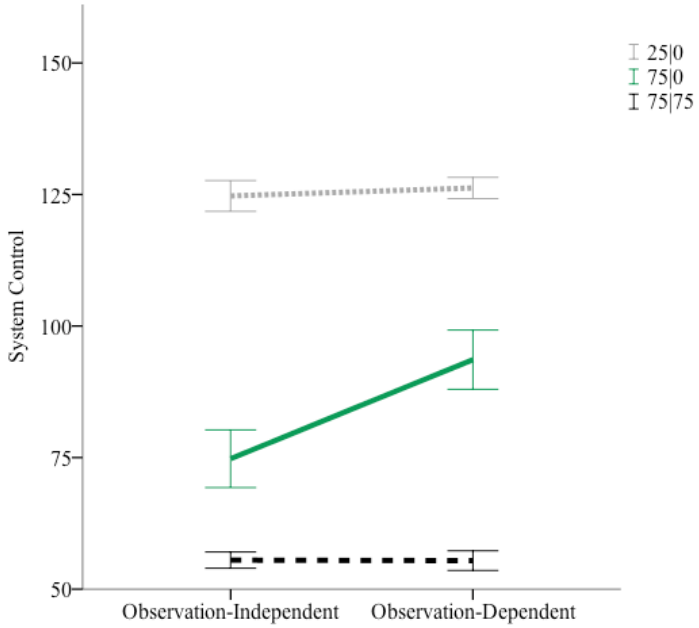


Figure Y1. Relationships between participants' system control and intervention strategy separated by contingency collapsing over system and delay. Error bars illustrate 95% confidence interval. Asterisks identify significant differences. A) Relationship between system control and intervention bias. B) Relationship between system control and observation dependence.

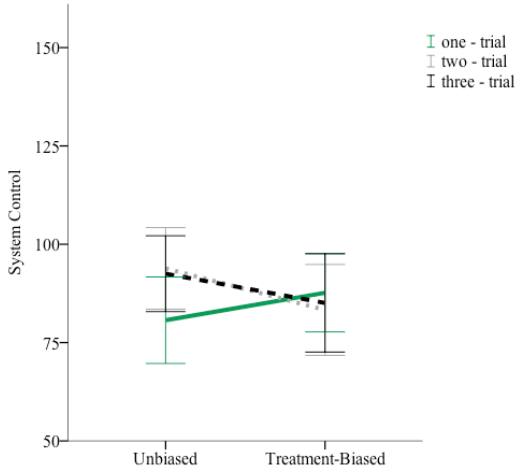
Table Y1
Effect of Strategy by Contingency Level on Control of the System (simple effects, between subjects)

	Unbiased			Treatment-biased						
	n	*M*	*SD*	*n*	*M*	*SD*	*F*	*df*	*p*	η^2
25|0	34	128.71	6.07	36	121.83	5.40	21.25	1, 67	< .01	.24
75|0	38	92.11	16.48	34	69.82	14.65	33.49	1, 69	< .01	.33
75|75	42	55.76	5.19	25	54.68	4.42	0.90	1, 64	.35	.01
	Observation-independent			Observation-dependent						
	n	*M*	*SD*	*n*	*M*	*SD*	*F*	*df*	*p*	η^2
25|0	30	124.73	7.82	42	126.24	6.51	1.03	1, 69	.31	.02
75|0	46	74.78	18.44	26	93.62	13.92	17.02	1, 69	< .01	.20
75|75	43	55.51	5.04	29	55.41	4.91	0.01	1, 69	.91	.00

Delay

Delay, in combination with intervention bias, did not affect participants' ability to control the system (Figure Y2a). However, delay, in combination with observation dependence shows some influence over participants' ability to control the system (Figure Y2b). Participants that applied observation-independent intervention strategies lowered patients' blood pressure levels to a greater degree than participants that applied observation-dependent intervention strategies with the two-trial and three-trial delay manipulations. Figure Y2 shows the relationships between system control and intervention strategy separated by delay collapsing over system and contingency, and Table Y2 presents the simple effect findings. Lower system control values reflect better performance.

A) Intervention Bias



B) Observation Dependence

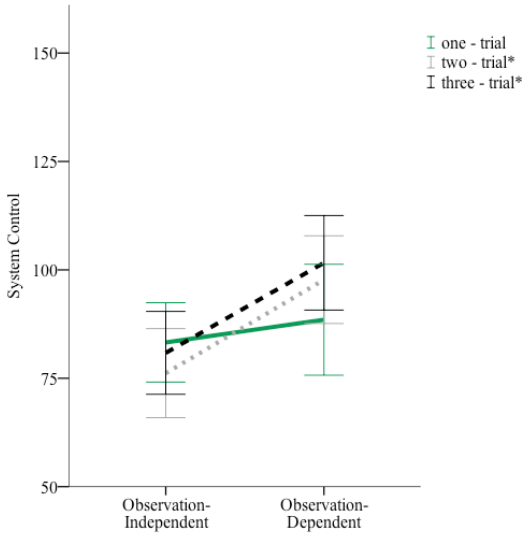


Figure Y2. Relationships between participants' system control and intervention strategy separated by delay collapsing over system and contingency. Error bars illustrate 95% confidence interval. Asterisks identify significant differences. A) Relationship between system control and intervention bias. B) Relationship between system control and observation dependence.

Table Y2

Effect of Strategy Given Delay on Control of the System (simple effects, between subjects)

	Unbiased			Treatment-biased					
Delay	*n*	*M*	*SD*	*n*	*M*	*SD*	*F*	*df*	*p*	$\square^2_{\square}$
One-trial	32	80.69	30.53	39	87.67	30.72	0.90	1, 69	.35	.01
Two-trial	37	93.86	31.12	32	83.31	32.06	1.86	1, 66	.18	.03
Three-trial	45	92.51	32.10	24	85.08	29.59	0.88	1, 66	.35	.01
	Observation-independent			Observation-dependent						
Delay	*n*	*M*	*SD*	*n*	*M*	*SD*	*F*	*df*	*p*	$\square^2_{\square}$
One-trial	44	83.27	30.11	28	88.50	32.96	0.44	1, 69	.51	.00
Two-trial	34	76.18	29.44	38	97.74	30.79	9.87	1, 69	< .01	.13
Three-trial	41	80.85	30.26	31	101.65	29.73	9.11	1, 69	< .01	.12

APPENDIX Z

Effect of Intervention Strategy and Delay on Causal Attributions

As with the data separated by contingency, causal attributions generally appear lower with unbiased and observation-dependent strategies across the delay manipulations. Despite a consistent pattern of behavior, causal attributions only significantly differed according to intervention bias with the three-trial delay manipulation (Figure Z1a). Also, higher causal attributions were associated with observation-independent strategies with the one-trial and two-trial delay manipulations (Figure Z1b), but attributions did not significantly differ with the three-trial delay manipulation regardless of observation dependence. Table Z1 presents the simple effect findings. Figure Z1 shows the relationships between participants' causal attributions and intervention strategy separated by delay collapsing over system and contingency, and

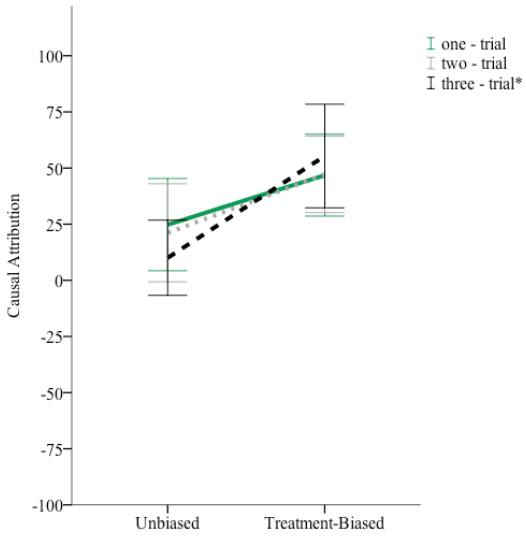
Table Z1

Effect of Strategy given Delay on Causal Attributions (simple effects, between subjects)

	Unbiased			Treatment-biased						
Delay	*n*	*M*	*SD*	*n*	*M*	*SD*	*F*	*df*	*p*	η^2_p
One-trial	32	24.78	57.02	39	46.82	56.26	2.59	1, 68	.11	.04
Two-trial	37	21.16	65.44	32	47.19	47.31	3.43	1, 66	.07	.05
Three-trial	45	10.04	55.81	24	55.37	54.57	10.32	1, 66	< .01	.14
	Observation-independent			Observation-dependent						
Delay	*n*	*M*	*SD*	*n*	*M*	*SD*	*F*	*df*	*p*	η^2_p
One-trial	44	55.30	48.50	28	6.64	57.27	14.71	1, 69	< .01	.18
Two-trial*	34	51.82	49.46	38	11.32	64.19	9.82	1, 69	< .01	.13
Three-trial	41	32.22	59.33	31	6.26	63.99	3.42	1, 69	.07	.05

Note. Significant system * strategy interaction with this delay manipulation.

A) Intervention Bias



Dependence

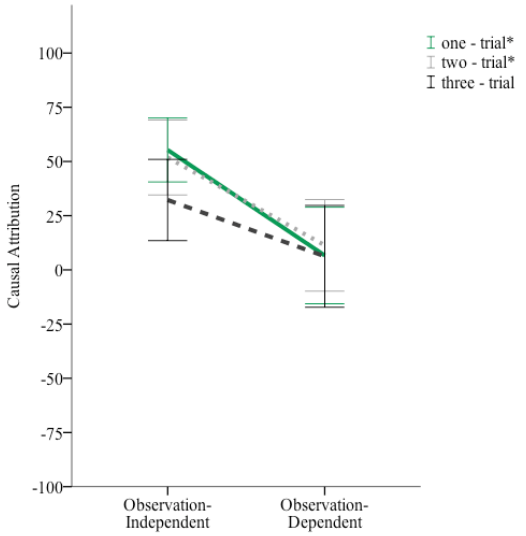


Figure Z1. Relationships between participants' causal attributions and intervention strategy separated by delay collapsing over system and contingency. Error bars illustrate 95% confidence interval. Asterisks identify significant differences. A) Relationship between causal attributions and intervention bias. B) Relationship between causal attributions and observation dependence.

APPENDIX AA

Intervention Strategy and System Condition ANOVAs

Tables AA1 and AA2 report the ANOVA findings pertaining to the interactions between intervention strategy and system condition. Tables AA3 and AA4 report the system condition by contingency and delay manipulation simple effects. The intervention strategy by contingency and delay manipulation simple effects are included in the body of the document, so are not repeated here.

Table AA1

The Effect of Treatment Intervention Strategies and System Condition separated by levels of Contingency and Delay on Causal Attributions

	F	*df*	*p*
*Treatment Bias*System Condition*			
Contingency			
25|0	2.31	1, 66	.13
75|0	1.86	1, 68	.18
75|75	0.00	1, 63	.99
Delay			
One-trial	0.03	1, 67	.86
Two-trial	2.57	1, 65	.11
Three-trial	1.72	1, 65	.19
*Observation Dependence*System Condition*			
Contingency			
25|0	1.41	1, 68	.24
75|0	3.12	1, 68	.08
75|75	0.01	1, 68	.91
Delay			
One-trial	0.34	1, 68	.57
Two-trial	7.79	1, 68	.01
Three-trial	0.23	1, 68	.63

Note. The following are the descriptive statistics pertaining to the significant interaction: Discrete System * Observation Dependence (Observation-Independent: $n = 11$, $M = 33.64$, $SD = 57.75$, Observation-Dependent: $n = 25$, $M = 28.48$, $SD = 52.86$), Continuous System * Observation Dependence (Observation-Independent: $n = 23$, $M = 60.52$, $SD = 43.70$, Observation-Dependent: $n = 13$, $M = -21.69$, $SD = 72.94$).

Table AA2

The Effect of Treatment Selection Strategies and System Condition separated by levels of Contingency and Delay on Control of the System

	F	*df*	*p*
*Treatment Bias*System Condition*			
Contingency			
25|0	4.08	1,66	.05*
75|0	1.73	1, 68	.19
75|75	0.08	1, 63	.78
Delay			
One-trial	0.17	1, 67	.68
Two-trial	2.41	1, 65	.13
Three-trial	0.34	1, 65	.57
*Observation Dependence*System Condition*			
Contingency			
25|0	0.64	1, 68	.43
75|0	0.08	1, 68	.78
75|75	1.37	1, 68	.25
Delay			
One-trial	2.56	1, 68	.11
Two-trial	2.38	1, 68	.13
Three-trial	0.04	1, 68	.85

Note. * p =.047. The following are the descriptive statistics pertaining to the significant interaction: Discrete System * Treatment Bias (Treatment-Biased: n = 21, M = 119.57, SD = 3.59, Unbiased: n = 13, M = 128.69, SD = 4.39), Continuous System * Treatment Bias (Treatment-Biased: n = 15, M = 125.00, SD = 6.00, Unbiased: n = 21, M = 128.71, SD = 7.02).

Table AA3

Effect of System on Causal Attributions (simple effects, between subjects) separated by levels of Contingency and Delay

	Discrete			Continuous						
	n	*M*	*SD*	*n*	*M*	*SD*	*F*	*df*	*p*	η_p^2
Contingency										
25|0	34	2.88	51.44	36	-23.67	47.50	3.74	1, 67	.06	.05
75|0	36	36.19	57.75	36	68.75	31.33	6.55	1, 69	.01	.09
75|75	34	52.65	50.88	33	56.94	49.82	0.06	1, 64	.81	.00
Delay										
One-trial	35	35.34	57.62	36	38.39	57.70	0.01	1, 68	.91	.00
Two-trial	34	32.71	50.40	35	33.74	66.71	0.00	1, 66	.99	.00
Three-trial	35	24.06	62.80	34	27.62	55.97	0.05	1, 66	.82	.00

Note. Findings from ANOVA that also considered intervention bias

Table AA4

Effect of System on Control of the System (simple effects, between subjects) separated by levels of Contingency and Delay

	Discrete			Continuous						
	n	*M*	*SD*	*n*	*M*	*SD*	*F*	*df*	*p*	η_p^2
Contingency										
25|0	34	123.06	5.92	36	127.17	6.78	4.32	1, 67	.04	.06
75|0	36	85.17	17.30	36	78.00	20.45	0.90	1, 69	.35	.01
75|75	34	55.12	5.00	33	55.61	4.42	0.31	1, 64	.58	.01
Delay										
One-trial	35	84.43	29.50	36	84.61	32.09	0.00	1, 68	.97	.00
Two-trial	34	90.12	29.17	35	87.86	34.50	0.06	1, 66	.82	.00
Three-trial	35	88.71	31.16	34	91.18	31.72	0.11	1, 66	.74	.00

Note. Findings from ANOVA that also considered intervention bias

APPENDIX AB

Theoretical Predictions relative to Observed Causal Attributions

I used one-sample *t*-tests to compare the observed causal attributions to theoretical predictions separated by treatment selection strategy (i.e., treatment bias, observation dependence) and contingency. I separated the observed data according to intervention strategy to account for the significant simple effects of intervention strategy and contingency on causal attribution. In general, higher causal attributions were associated with treatment-biased and observation-independent intervention strategies. For each contingency manipulation, I collapsed the causal attribution data across delay manipulations, as there were no significant delay main effects with causal attributions as the dependent measure. For brevity, I will only discuss the results relative to a subset of the theoretical predictions, but provide tables containing the comprehensive set of analyses at the end of this Appendix (Tables AB5-AB12). Somewhat counter-intuitively, non-significant differences with these analyses indicate that the theoretical predictions satisfactorily describe participants' causal attributions. These are designated with asterisks.

Although main effect trends are generally consistent with existing empirical work, a single reasoning theory does not successfully predict the observed causal attributions across the contingency manipulations for either system condition. This finding is consistent with the observed differences in causal attributions and the effect of reasoning strategy on the outcome measure, as well as the differences in participants' self-reported

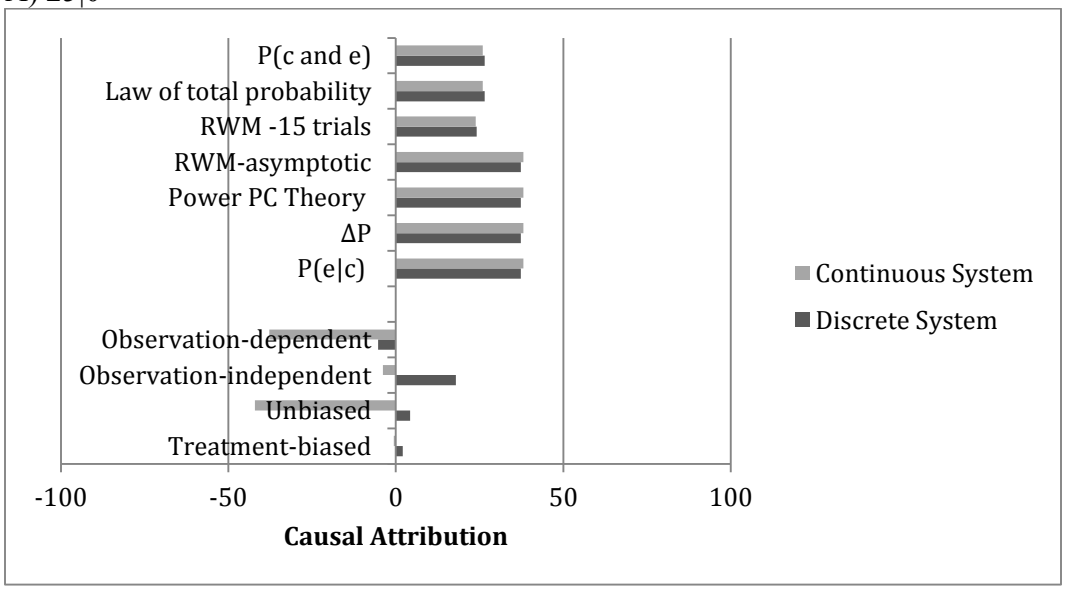
assessment technique in the judgment task. Also, system effects on causal attribution are apparent with this analysis. The theoretical predictions accounted for more causal attributions associated with the 25|0 contingency manipulation in the discrete system condition (which were less extreme), an equivalent amount of causal attributions with the 75|0 contingency manipulation (despite the more extreme causal attributions with the continuous system condition), and the pattern of findings was similar between the discrete and continuous systems with the 75|75 contingency manipulation. The theoretical predictions that left out probabilistic information, including base rates, were more successful at predicting the observed causal attributions.

Predicted Values

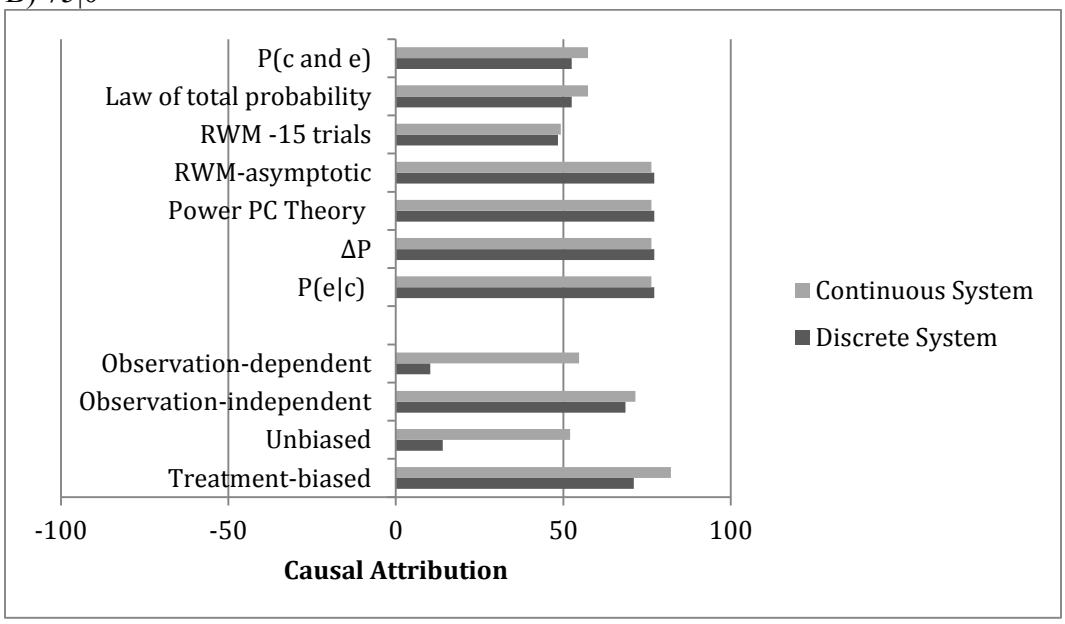
I calculated a set of predicted values (using the mean actual contingency values) for popular theories within the causal reasoning literature including Bayes Theorem (i.e., *P(e|c)), ΔP* (i.e., *P(e|c)-P(e|~c))*, Power PC Theory *(P(e|c)-P(e|~c))/(1-P(e|~c))*, and Rescorla-Wagner Model (i.e., RWM; see Appendix AC for calculations), as well as probabilistic values representative of reasoning strategies described by participants in their debrief questionnaire responses including *P(e)* and *P(c and e)*. *P(e)*, or the law of total probability, is representative of a reasoning strategy that focuses on the overall effect, meaning the number of times (discrete system) or the degree to which (continuous system) the patients' blood pressure lowered. *P(c and e)* is characteristic of a reasoning strategy in which the participant focuses on the joint probability of a cause (i.e., the number application of the treatment selection) and effect (i.e., number of times the patients' blood pressure lowers) without the consideration of base rate information (i.e., *P(e)* or *P(c))*.

Figure AB1 depicts the predicted causal attribution values alongside the mean of the observed causal attribution values separated by treatment intervention strategy for each system condition (the specific values are presented in Appendix AD). Several of the theoretical predictions converge at two predicted values with the 25|0 and 75|0 contingency manipulations, but all of the predictions vary with respect to the 75|75 contingency manipulation. In the subsequent analyses, I code these identical predictions as Group 1 (Bayes Theorem *(P(e|c)), ΔP*, RWM-asymptotic, and Power PC Theory) and Group 2 (*P(c and e)* and the law of total probability).

A) 25|0



B) 75|0



C) 75|75

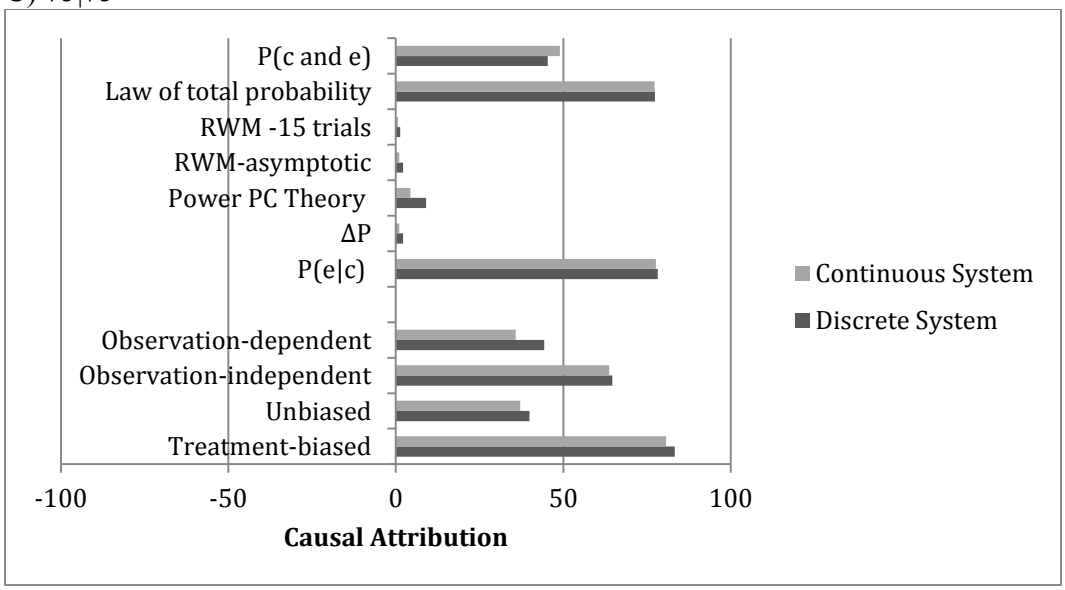


Figure AB1. Theoretical Predictions Compared to Mean Causal Attribution Values separated by Treatment Intervention Strategy (Observation-dependent, Observation-independent, Unbiased, Treatment-biased). Bayes theorem = *(P(e|c)); ΔP* Model = *P(e|c)-P(e|~c));* Power PC Theory = *(P(e|c)-P(e|~c))*/(1-*P(e|~c)*); RWM = Rescorla-Wagner Model, see Appendix AC for calculations; Law of total probability = *P(e)*

Discrete System

With the discrete system, a single reasoning theory was unable to consistently predict the observed causal attributions across the treatment intervention strategies or the

contingency manipulations. Although the theoretical predictions accounted for the causal attributions of participants that applied treatment-biased and observation-independent intervention strategies across the contingency manipulations, there was less correspondence between the predictions and causal attributions associated with unbiased and observation-dependent intervention strategies. Tables AB1 and AB2 contain the outcomes of the one-sample *t*-tests with the discrete system condition that compared the observed causal attributions to various theoretical and probabilistic predictions.

Table AB1

Theoretical Predictions relative to Observed Causal Attributions with the Discrete System Condition separated by Intervention Bias

		Unbiased					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		4.23	43.46				
Group 1	37.29			-2.74	12	.02	-0.76
Group 2	26.48			-1.85	12	.09*	-0.51
75|0		14.05	62.05				
Group 1	77.18			-4.77	21	< .01	-1.02
Group 2	52.48			-2.91	21	.01	-0.62
75|75		39.88	55.52				
ΔP	2.17			3.33	23	< .01	0.68
P(e)	77.40			-3.31	23	< .01	-0.68
P(c and e)	45.34			-0.48	23	.63*	-0.10
		Treatment-biased					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		2.05	56.84				
Group 1	37.29			-2.84	20	.01	-0.62
Group 2	26.48			-1.97	20	.06*	-0.43
75|0		71.00	24.96				
Group 1	77.18			-0.93	13	.37*	-0.25
Group 2	52.48			2.78	13	.02	0.74
75|75		83.30	11.61				
ΔP	2.17			22.11	9	< .01	6.99
P(e)	77.40			1.61	9	.14*	0.51
P(c and e)	45.34			10.34	9	< .01	3.27

Note. Predicted = Predicted theoretical value. Group 1 = Bayes Theorem *(P(e|c)), ΔP*, RWM-asymptotic, Power PC Theory; Group 2 = Law of total probability*, P(c and e).*

* Non-significant findings indicate that the predicted value was equivalent to the observed causal attributions.

Table AB2

Theoretical Predictions relative to Observed Causal Attributions with the Discrete System Condition separated by Observation Dependence

		Observation-independent					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		17.92	48.47				
Group 1	37.29			-1.38	11	.19*	-0.40
Group 2	26.48			-0.61	11	.55*	-0.18
75|0		68.56	27.01				
Group 1	77.18			-1.28	15	.22*	-0.32
Group 2	52.48			2.38	15	.03	0.60
75|75		64.64	36.49				
ΔP	2.17			12.18	13	< .01	1.71
P(e)	77.40			0.53	13	.61*	-0.35
P(c and e)	45.34			4.86	13	< .01	0.53
		Observation-dependent					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		5.32	52.22				
Group 1	37.29			-3.83	21	< .01	-0.82
Group 2	26.48			-2.86	21	.01	-0.61
75|0		10.30	63.08				
Group 1	77.18			-4.74	19	< .01	-1.06
Group 2	52.48			-2.99	19	.01	-0.67
75|75		44.25	58.35				
ΔP	2.17			2.69	19	.02	0.72
P(e)	77.40			-3.00	19	.01	-0.57
P(c and e)	45.34			-0.88	19	.39*	-0.02

Note. Predicted = Predicted theoretical value. Group 1 = Bayes Theorem *(P(e|c)), ΔP*, RWM-asymptotic, Power PC Theory; Group 2 = Law of total probability*, P(c and e).*

***** Non-significant findings indicate that the predicted value was equivalent to the observed causal attributions.

Predicted values compared to causal attributions with the 75|0 contingency manipulation. The predicted value associated with Bayes Theorem *(P(e|c)), ΔP*, Power PC Theory, and Rescorla-Wagner Model at its asymptotic level (Group 1) was equivalent to causal attributions associated with treatment-biased and observation-independent intervention strategies with the 75|0 contingency manipulation. In contrast, causal

attributions were more equivalent to the reasoning theories' predictions that omitted base rate information (i.e., Group 2 - *P(e)* and *P*(*c and e))* with the 25|0 and 75|75 contingency manipulations. None of the theoretical predictions matched causal attributions associated with unbiased and observation-dependent intervention strategies with the 75|0 contingency manipulation.

Predicted values compared to causal attributions with the 25|0 contingency manipulation. The predicted value of *P(e)* and *P*(*c and e)* (Group 2) was equivalent to causal attributions associated with the treatment-biased, unbiased, and observation-independent intervention strategies for the 25|0 contingency manipulation. Causal attributions that followed observation-independent intervention strategies were also equivalent to the predictions of base rate inclusive models (i.e., Group 1 - *(P(e|c)), ΔP*, Power PC Theory, and Rescorla-Wagner Model at asymptotic).

Predicted values compared to causal attributions with the 75|75 contingency manipulation. With 75|75 contingency manipulation, causal attributions associated with treatment-biased and observation-independent intervention strategies were equivalent to the overall probability of the effect (i.e., *P(e)).* Conversely, the joint probability of a cause and effect (i.e., *P*(*c and e))* matched the causal attributions of participants that used unbiased and observation-dependent intervention strategies. The *ΔP* prediction, and the predictions of other models that were closer to zero (see Appendix AD), did not account for causal attributions across the treatment intervention strategies with the 75|75 contingency manipulation.

Continuous System

As with the discrete system condition, no reasoning theory consistently predicted the observed causal attributions across the contingency manipulations with the continuous system condition. The theoretical predictions accounted for more causal attributions with the 75|0 contingency manipulation with the continuous than the discrete system condition. However, none of the predictions accounted for the causal attributions associated with the 25|0 contingency manipulation. Tables AB3 and AB4 contain the complete series of comparisons between the observed causal attributions and the theoretical and probabilistic predictions with the continuous system condition.

Table AB3

Theoretical Predictions relative to Observed Causal Attributions with the Continuous System Condition separated by Intervention Bias

		Unbiased					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		-37.81	43.85				
Group 1	38.09			-7.93	20	< .01	-1.73
Group 2	25.99			-6.67	20	< .01	-1.45
75|0		52.06	40.07				
Group 1	76.29			-2.42	15	.03	-0.60
Group 2	57.35			-0.53	15	.60*	-0.13
75|75		37.11	56.81				
ΔP	1.06			2.69	17	.02	0.63
P(e)	77.28			-3.00	17	.01	-0.71
P(c and e)	48.94			-0.88	17	.39*	-0.21
		Treatment-biased					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		-3.87	46.64				
Group 1	38.09			-3.48	14	< .01	-0.90
Group 2	25.99			-2.48	14	.03	-0.64
75|0		82.10	10.90				
Group 1	76.29			2.38	19	.03	0.53
Group 2	57.35			10.15	19	< .01	2.27
75|75		80.73	25.34				
ΔP	1.06			12.18	14	< .01	3.14
P(e)	77.28			0.53	14	.61*	0.14
P(c and e)	48.94			4.86	14	< .01	1.25

Note. Predicted = Predicted theoretical value. Group 1 = Bayes Theorem *(P(e|c)), ΔP*, RWM-asymptotic, Power PC Theory; Group 2 = Law of total probability*, P(c and e).*

* Non-significant findings indicate that the predicted value was equivalent to the observed causal attributions.

Table AB4

Theoretical Predictions relative to Observed Causal Attributions with the Continuous System Condition separated by Observation Dependence

		Observation-independent					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0[a]		-0.63	39.07				
Group 1	38.09			-3.96	15	< .01	-0.99
Group 2	25.99			-2.73	15	.02	-0.68
75|0[b]		71.57	33.32				
Group 1	76.29			-0.78	29	.44*	-0.14
Group 2	57.35			2.34	29	.03	0.43
75|75[c]		63.72	46.34				
ΔP	1.06			6.76	24	< .01	1.35
P(e)	77.28			-1.46	24	.16*	-0.29
P(c and e)	48.94			1.60	24	.12*	0.32
		Observation-dependent					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		-42.10	46.34				
Group 1	38.09			-7.74	19	< .01	-1.73
Group 2	25.99			-6.57	19	< .01	-1.47
75|0		54.67	12.18				
Group 1	76.29			-4.35	5	< .01	-1.78
Group 2	57.35			-0.54	5	.61*	-0.22
75|75		35.75	57.51				
ΔP	1.06			1.71	7	.13*	0.60
P(e)	77.28			-2.04	7	.08*	-0.72
P(c and e)	48.94			-0.65	7	.54*	-0.23

Note. Predicted = Predicted theoretical value. Group 1 = Bayes Theorem *(P(e|c)), ΔP*, RWM-asymptotic, Power PC Theory; Group 2 = Law of total probability*, P(c and e).*

* Non-significant findings indicate that the predicted value was equivalent to the observed causal attributions.

Predicted values compared to causal attributions with the 75|0 contingency manipulation. Consistent with the discrete system condition findings, the predicted value associated with *P(e|c), ΔP*, Power PC Theory, and Rescorla-Wagner Model at its asymptotic level (Group 1) was equivalent to causal attributions of participants that applied observation-independent invention strategies with the 75|0 contingency

manipulation. However, the predicted value of Group 1 did not account for causal attributions following treatment-biased strategies with the continuous system condition, nor did any of the predicted values. Also, a departure from the discrete system findings, causal attributions of participants that applied unbiased and observation-dependent intervention strategies with the 75|0 contingency manipulation were equivalent to the theoretical prediction of *P(e)* and *P(c and e)* (Group 2).

Predicted values compared to causal attributions with the 25|0 contingency manipulation. None of the theoretical predictions matched the observed causal attributions with the 25|0 contingency manipulation.

Predicted values compared to causal attributions with the 75|75 contingency manipulation. *P(e)* and *P(c and e)* were equivalent to causal attributions following observation-independent and observation-dependent intervention strategies with the 75|75 contingency manipulation. Neither *P(e)* nor *P(c and e)* take into account the specific value associated with the probability of the effect in the absence of the causal variable (e.g., *P(e|~c))*. The *ΔP* prediction was statically consistent with causal attributions associated with observation-dependent strategies, as well. However, this outcome may be attributable to low power given the limited number of participants that applied observation-dependent strategies with the continuous system condition. With respect to the strategic dimension of intervention bias, only *P(e)* predicted causal attributions that followed treatment-biased selections strategies and only *P(c and e)* predicted causal attributions that followed applied unbiased-treatment intervention strategies.

Comprehensive Tables

Table AB5
Theoretical Predictions relative to Observed Causal Attributions with Unbiased Intervention Strategies in the Discrete System Condition

		Unbiased					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0							
Group 1	37.29	4.23	43.46	-2.74	12	.02	-0.76
Group 2	26.48			-1.85	12	.09*	-0.51
RWM -15 trials	24.10			-1.65	12	.13*	-0.46
75|0		14.05	62.05				
Group 1	77.18			-4.77	21	< .01	-1.02
Group 2	52.48			-2.91	21	.01	-0.62
RWM -15 trials	48.40			-2.60	21	.02	-0.55
75|75		39.88	55.52				
P(e|c)	78.17			-3.38	23	< .01	-0.69
ΔP	2.17			3.33	23	< .01	0.68
Power PC Theory	9.04			2.72	23	.01	0.56
RWM-asymptotic	2.17			3.33	23	< .01	0.68
RWM -15 trials	1.32			3.40	23	< .01	0.69
P(e)	77.40			-3.31	23	< .01	-0.68
P(c and e)	45.34			-0.48	23	.63*	-0.10

Note. Predicted = Predicted theoretical value. Group 1 = Bayes Theorem *(P(e|c)), ΔP*, RWM-asymptotic, Power PC Theory; Group 2 = Law of total probability*, P(c and e)*. *Findings that are statistically non-significant indicate that the predicted value was equivalent to the observed causal attributions.

Table AB6

Theoretical Predictions relative to Observed Causal Attributions with Treatment-biased Intervention Strategies in the Discrete System Condition

		Treatment-biased					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		2.05	56.84				
Group 1	37.29			-2.84	20	.01	-0.62
Group 2	26.48			-1.97	20	.06*	-0.43
RWM -15 trials	24.10			-1.78	20	.09*	-0.39
75|0		71.00	24.96				
Group 1	77.18			-0.93	13	.37*	-0.25
Group 2	52.48			2.78	13	.02	0.74
RWM -15 trials	48.40			3.39	13	.01	0.91
75|75		83.30	11.61				
P(e|c)	78.17			1.40	9	.20*	0.44
ΔP	2.17			22.11	9	< .01	6.99
Power PC Theory	9.04			20.24	9	< .01	6.40
RWM-asymptotic	2.17			22.11	9	< .01	6.99
RWM -15 trials	1.32			22.34	9	< .01	7.06
P(e)	77.40			1.61	9	.14*	0.51
P(c and e)	45.34			10.34	9	< .01	3.27

Note. Predicted = Predicted theoretical value. Group 1 = Bayes Theorem *(P(e|c)), ΔP*, RWM-asymptotic, Power PC Theory; Group 2 = Law of total probability*, P(c and e).* *Findings that are statistically non-significant indicate that the predicted value was equivalent to the observed causal attributions.

Table AB7

Theoretical Predictions relative to Observed Causal Attributions with Observation-independent Intervention Strategies in the Discrete System Condition

		Observation-independent					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		17.92	48.47				
Group 1	37.29			-1.38	11	.19*	-0.40
Group 2	26.48			-0.61	11	.55*	-0.18
RWM -15 trials	24.10			-0.44	11	.67*	-0.13
75|0		68.56	27.01				
Group 1	77.18			-1.28	15	.22*	-0.32
Group 2	52.48			2.38	15	.03	0.60
RWM -15 trials	48.40			2.99	15	.01	0.75
75|75		64.64	36.49				
P(e|c)	78.17			-1.39	13	.19*	-0.37
ΔP	2.17			12.18	13	< .01	1.71
Power PC Theory	9.04			5.70	13	< .01	1.52
RWM-asymptotic	2.17			6.41	13	< .01	1.71
RWM -15 trials	1.32			6.49	13	< .01	1.74
P(e)	77.40			0.53	13	.61*	-0.35
P(c and e)	45.34			4.86	13	< .01	0.53

Note. Predicted = Predicted theoretical value. Group 1 = Bayes Theorem *(P(e|c)), ΔP*, RWM-asymptotic, Power PC Theory; Group 2 = Law of total probability*, P(c and e).* *Findings that are statistically non-significant indicate that the predicted value was equivalent to the observed causal attributions.

Table AB8

Theoretical Predictions relative to Observed Causal Attributions with Observation-dependent Intervention Strategies in the Discrete System Condition

		Observation-dependent					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		-5.32	52.22				
Group 1	37.29			-3.83	21	< .01	-0.82
Group 2	26.48			-2.86	21	.01	-0.61
RWM -15 trials	24.10			-2.64	21	.02	-0.56
75|0		10.30	63.08				
Group 1	77.18			-4.74	19	< .01	-1.06
Group 2	52.48			-2.99	19	.01	-0.67
RWM -15 trials	48.40			-2.70	19	.01	-0.60
75|75		44.25	58.35				
P(e|c)	78.17			-2.60	19	.02	-0.58
ΔP	2.17			2.69	19	.02	0.72
Power PC Theory	9.04			2.70	19	.01	0.60
RWM-asymptotic	2.17			3.23	19	< .01	0.72
RWM -15 trials	1.32			3.29	19	< .01	0.74
P(e)	77.40			-3.00	19	.01	-0.57
P(c and e)	45.34			-0.88	19	.39*	-0.02

Note. Predicted = Predicted theoretical value. Group 1 = Bayes Theorem *(P(e|c)), ΔP*, RWM-asymptotic, Power PC Theory; Group 2 = Law of total probability*, P(c and e)*. *Findings that are statistically non-significant indicate that the predicted value was equivalent to the observed causal attributions.

Table AB9

Theoretical Predictions relative to Observed Causal Attributions with Unbiased Intervention Strategies in the Discrete System Condition

		Unbiased					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		-37.81	43.85				
Group 1	38.09			-7.93	20	< .01	-1.73
Group 2	25.99			-6.67	20	< .01	-1.45
RWM -15 trials	23.83			-6.44	20	< .01	-1.41
75|0		52.06	40.07				
Group 1	76.29			-2.42	15	.03	-0.60
Group 2	57.35			-0.53	15	.60*	-0.13
RWM -15 trials	49.30			0.28	15	.79	0.07
75|75		37.11	56.81				
P(e|c)	77.66			-3.03	17	.01	-0.71
ΔP	1.06			2.69	17	.02	0.63
Power PC Theory	4.39			2.44	17	.03	0.58
RWM-asymptotic	1.06			2.69	17	.02	0.63
RWM -15 trials	0.64			2.72	17	.02	0.64
P(e)	77.28			-3.00	17	.01	-0.71
P(c and e)	48.94			-0.88	17	.39*	-0.21

Note. Predicted = Predicted theoretical value. Group 1 = Bayes Theorem *(P(e|c)), ΔP*, RWM-asymptotic, Power PC Theory; Group 2 = Law of total probability*, P(c and e).* *Findings that are statistically non-significant indicate that the predicted value was equivalent to the observed causal attributions.

Table AB10

Theoretical Predictions relative to Observed Causal Attributions with Treatment-biased Intervention Strategies in the Discrete System Condition

		Treatment-biased					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		-3.87	46.64				
Group 1	38.09			-3.48	14	< .01	-0.90
Group 2	25.99			-2.48	14	.03	-0.64
RWM -15 trials	23.83			-2.30	14	.04	-0.59
75|0		82.10	10.90				
Group 1	76.29			2.38	19	.03	0.53
Group 2	57.35			10.15	19	< .01	2.27
RWM -15 trials	49.30			13.46	19	< .01	3.01
75|75		80.73	25.34				
P(e|c)	77.66			0.47	14	.65	0.12
ΔP	1.06			12.18	14	< .01	3.14
Power PC Theory	4.39			11.67	14	< .01	3.01
RWM-asymptotic	1.06			12.18	14	< .01	3.14
RWM -15 trials	0.64			12.24	14	< .01	3.16
P(e)	77.28			0.53	14	.61*	0.14
P(c and e)	48.94			4.86	14	< .01	1.25

Note. Predicted = Predicted theoretical value. Group 1 = Bayes Theorem *(P(e|c)), ΔP*, RWM-asymptotic, Power PC Theory; Group 2 = Law of total probability*, P(c and e).* *Findings that are statistically non-significant indicate that the predicted value was equivalent to the observed causal attributions.

Table AB11

Theoretical Predictions relative to Observed Causal Attributions with Observation-independent Intervention Strategies in the Discrete System Condition

		Observation-independent					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		-0.63	39.07				
Group 1	38.09			-3.96	15	< .01	-0.99
Group 2	25.99			-2.73	15	.02	-0.68
RWM -15 trials	23.83			-2.50	15	.02	-0.63
75|0		71.57	33.32				
Group 1	76.29			-0.78	29	.44*	-0.14
Group 2	57.35			2.34	29	.03	0.43
RWM -15 trials	49.30			3.66	29	< .01	0.67
75|75		63.72	46.34				
P(e|c)	77.66			-1.50	24	.15*	-0.30
ΔP	1.06			6.76	24	< .01	1.35
Power PC Theory	4.39			6.40	24	< .01	1.28
RWM-asymptotic	1.06			6.76	24	< .01	1.35
RWM -15 trials	0.64			6.81	24	< .01	1.36
P(e)	77.28			-1.46	24	.16*	-0.29
P(c and e)	48.94			1.60	24	.12*	0.32

Note. Predicted = Predicted theoretical value. Group 1 = Bayes Theorem *(P(e|c)), ΔP*, RWM-asymptotic, Power PC Theory; Group 2 = Law of total probability*, P(c and e).* *Findings that are statistically non-significant indicate that the predicted value was equivalent to the observed causal attributions.

Table AB12

Theoretical Predictions relative to Observed Causal Attributions with Observation-independent Intervention Strategies in the Discrete System Condition

		Observation-dependent					
	Predicted	*M*	*SD*	*t*	*df*	*p*	*d*
25|0		-42.10	46.34				
Group 1	38.09			-7.74	19	< .01	-1.73
Group 2	25.99			-6.57	19	< .01	-1.47
RWM -15 trials	23.83			-6.36	19	< .01	-1.42
75|0		54.67	12.18				
Group 1	76.29			-4.35	5	< .01	-1.78
Group 2	57.35			-0.54	5	.61*	-0.22
RWM -15 trials	49.30			1.08	5	.33*	0.44
75|75		35.75	57.51				
P(e|c)	77.66			-2.06	7	.08*	-0.73
ΔP	1.06			1.71	7	.13*	0.60
Power PC Theory	4.39			1.54	7	.17*	0.55
RWM-asymptotic	1.06			1.71	7	.13*	0.60
RWM -15 trials	0.64			1.73	7	.13*	0.61
P(e)	77.28			-2.04	7	.08*	-0.72
P(c and e)	48.94			-0.65	7	.54*	-0.23

Note. Predicted = Predicted theoretical value. Group 1 = Bayes Theorem *(P(e|c)), ΔP*, RWM-asymptotic, Power PC Theory; Group 2 = Law of total probability*, P(c and e).* *Findings that are statistically non-significant indicate that the predicted value was equivalent to the observed causal attributions.

APPENDIX AC

Method Used to Generate Predicted RWM Values for Theoretical Comparison

RWM asymptotic

Approach used by Buehner, Cheng, and Clifford (2003) and was proposed in Wasserman et al. (1993).

$$V_{asymptopic} = \frac{a}{a+b} - \frac{c}{c+d}$$

where

- $a = P(E|C)$
- b = P(~E|C)
- c = $P(E|\sim C)$
- d = P(~E|~C)

RWM – 15 trials

Tables AC1 through AC6 illustrate the application of the following equations to generate the predicted values given 15 trials of observations. I used the same alpha and beta values to generate these predictions as Buehner, Cheng, and Clifford (2003).

$\Delta V = \alpha\beta(\lambda - V_{initial})$
and

$$V_{end\ trial} = V_{initial} + \Delta V$$

where

- $V_{intital}^{t} = V_{end\ trial}^{t-1}$
- $V_{intital}^{t=1} = 0.00$
- ΔV = change in association strength
- α = salience of blood pressure observation
- β = salience of treatment selection
- λ = causal contingency value or ΔP value

Table AC1

Generation of RWM Predicted Value for the 25|0 Contingency Manipulation with the Discrete System Condition

Trial	Treatment	Initial V_t	ΔV	EndTrial V_t
1	1	0.00	0.03	0.03
2	1	0.03	0.03	0.06
3	1	0.06	0.03	0.08
4	1	0.08	0.02	0.11
5	1	0.11	0.02	0.13
6	1	0.13	0.02	0.15
7	1	0.15	0.02	0.16
8	1	0.16	0.02	0.18
9	1	0.18	0.02	0.20
10	1	0.20	0.01	0.21
11	1	0.21	0.01	0.22
12	0	0.22	0.00	0.23
13	0	0.23	0.00	0.23
14	0	0.23	0.00	0.24
15	0	0.24	0.00	0.24

Note. I used 15 trials because participants were not shown the outcome of their treatment selection in the 16th trial and 11 treatment selections given the probability that a treatment applied with the 25|0 causal contingency manipulation with the continuous system condition was .71. Treatment Selections: 1= apply treatment, 0 = do not apply treatment. $\alpha = 0.1$, β(treatment) = .8, β(~treatment) = .3, $\lambda = .37$.

Table AC2

Generation of RWM predicted Value for the 75|0 Contingency Manipulation with the Discrete System Condition

Trial	Treatment	Initial V_t	ΔV	EndTrial V_t
1	1	0.00	0.06	0.06
2	1	0.06	0.06	0.12
3	1	0.12	0.05	0.17
4	1	0.17	0.05	0.22
5	1	0.22	0.04	0.26
6	1	0.26	0.04	0.30
7	1	0.30	0.04	0.34
8	1	0.34	0.03	0.38
9	1	0.38	0.03	0.41
10	1	0.41	0.03	0.44
11	0	0.44	0.01	0.45
12	0	0.45	0.01	0.46
13	0	0.46	0.01	0.47
14	0	0.47	0.01	0.48
15	0	0.48	0.01	**0.48**

Note. I used 15 trials because participants were not shown the outcome of their treatment selection in the 16th trial and 10 treatment selections given the probability that a treatment applied with the 75|0 causal contingency manipulation with the continuous system condition was .75. Treatment Selections: 1= apply treatment, 0 = do not apply treatment. $\alpha = 0.1$, β(treatment) = .8, β(~treatment) = .3, $\lambda = .77$.

Table AC3

Generation of RWM Predicted Value for the 75|75 Contingency Manipulation with the Discrete System Condition

Trial	Treatment	Initial V_t	ΔV	EndTrial V_t
1	1	0.0000	0.0017	0.0017
2	1	0.0017	0.0016	0.0033
3	1	0.0033	0.0015	0.0048
4	1	0.0048	0.0014	0.0062
5	1	0.0062	0.0012	0.0074
6	1	0.0074	0.0011	0.0085
7	1	0.0085	0.0011	0.0096
8	1	0.0096	0.0010	0.0106
9	1	0.0106	0.0009	0.0115
10	0	0.0115	0.0003	0.0118
11	0	0.0118	0.0003	0.0121
12	0	0.0121	0.0003	0.0123
13	0	0.0123	0.0003	0.0126
14	0	0.0126	0.0003	0.0129
15	0	0.0129	0.0003	**0.0132**

Note. I used 15 trials because participants were not shown the outcome of their treatment selection in the 16th trial and 9 treatment selections given the probability that a treatment applied with the 75|75 causal contingency manipulation with the continuous system condition was .58. Treatment Selections: 1= apply treatment, 0 = do not apply treatment. $\alpha = 0.1$, β(treatment) = .8, β(~treatment) = .3, $\lambda = .02$.

Table AC4

Generation of RWM Predicted Value for the 25|0 Contingency Manipulation with the Continuous System Condition

Trial	Treatment	Initial V_t	ΔV	EndTrial V_t
1	1	0.00	0.03	0.03
2	1	0.03	0.03	0.06
3	1	0.06	0.03	0.08
4	1	0.08	0.02	0.11
5	1	0.11	0.02	0.13
6	1	0.13	0.02	0.15
7	1	0.15	0.02	0.17
8	1	0.17	0.02	0.18
9	1	0.18	0.02	0.20
10	1	0.20	0.01	0.21
11	0	0.21	0.00	0.22
12	0	0.22	0.00	0.22
13	0	0.22	0.00	0.23
14	0	0.23	0.00	0.23
15	0	0.23	0.00	**0.24**

Note. I used 15 trials because participants were not shown the outcome of their treatment selection in the 16th trial and 10 treatment selections given the probability that a treatment applied with the 25|0 causal contingency manipulation with the continuous system condition was .68. Treatment Selections: 1= apply treatment, 0 = do not apply treatment. $\alpha = 0.1$, β(treatment) = .8, β(~treatment) = .3, $\lambda = .38$.

Table AC5

Generation of RWM Predicted Value for the 75|0 Contingency Manipulation with the Continuous System Condition

Trial	Treatment	Initial V_t	ΔV	EndTrial V_t
1	1	0.00	0.06	0.06
2	1	0.06	0.06	0.12
3	1	0.12	0.05	0.17
4	1	0.17	0.05	0.22
5	1	0.22	0.04	0.26
6	1	0.26	0.04	0.30
7	1	0.30	0.04	0.34
8	1	0.34	0.03	0.37
9	1	0.37	0.03	0.40
10	1	0.40	0.03	0.43
11	1	0.43	0.03	0.46
12	0	0.46	0.01	0.47
13	0	0.47	0.01	0.48
14	0	0.48	0.01	0.48
15	0	0.48	0.01	**0.49**

Note. I used 15 trials because participants were not shown the outcome of their treatment selection in the 16th trial and 11 treatment selections given the probability that a treatment applied with the 75|0 causal contingency manipulation with the continuous system condition was .75. Treatment Selections: 1= apply treatment, 0 = do not apply treatment. $\alpha = 0.1$, β(treatment) = .8, β(~treatment) = .3, $\lambda = .76$.

Table AC6

Generation of RWM Predicted Value for the 75|75 Contingency Manipulation with the Continuous System Condition

Trial	Treatment	Initial V_t	ΔV	EndTrial V_t
1	1	0.0000	0.0008	0.0008
2	1	0.0008	0.0008	0.0016
3	1	0.0016	0.0007	0.0023
4	1	0.0023	0.0006	0.0029
5	1	0.0029	0.0006	0.0035
6	1	0.0035	0.0005	0.0040
7	1	0.0040	0.0005	0.0045
8	1	0.0045	0.0005	0.0050
9	1	0.0050	0.0004	0.0054
10	0	0.0054	0.0001	0.0056
11	0	0.0056	0.0001	0.0057
12	0	0.0057	0.0001	0.0058
13	0	0.0058	0.0001	0.0060
14	0	0.0060	0.0001	0.0061
15	0	0.0061	0.0001	**0.0062**

Note. I used 15 trials because participants were not shown the outcome of their treatment selection in the 16th trial and 9 treatment selections given the probability that a treatment applied with the 75|75 causal contingency manipulation with the continuous system condition was .63. Treatment Selections: 1= apply treatment, 0 = do not apply treatment. $\alpha = 0.1$, β(treatment) = .8, β(~treatment) = .3, $\lambda = .01$.

APPENDIX AD

Predicted Causal Attributions

In order to compare the observed findings to theoretical predictions, I calculated a set of predicted values (using the mean actual contingency value for each contingency manipulation) for popular theories within the causal reasoning literature including Bayes Theorem (i.e., *P(e|c)), ΔP* (i.e., *P(e|c)-P(e|~c))*, Power PC Theory *(P(e|c)-P(e|~c))/(1-P(e|~c))*, and Rescorla-Wagner Model (i.e., RWM; see Appendix AC for calculations), as well as probabilistic values representative of reasoning strategies described by participants in their debrief questionnaire responses including *P(e)* and *P(c and e).* Table AD1 contains these predicted values alongside the mean of the observed causal attribution values for each contingency manipulation divided by intervention strategy.

Table AD1

Predicted and Observed Causal Attributions separated by Contingency and System Conditions

	Discrete System			Continuous System		
	25|0	75|0	75|75	25|0	75|0	75|75
	Mean observed causal attribution value					
Treatment-biased	2.05	71.00	83.30	-0.63	82.10	80.73
Unbiased	4.23	14.05	39.88	-42.10	52.06	37.11
Observation-independent	17.92	68.56	64.64	-3.87	71.57	63.72
Observation-dependent	-5.32	10.30	44.25	-37.81	54.67	35.75
	Predicted causal attribution value					
P(e|c)	37.29	77.18	78.17	38.09	76.29	77.66
ΔP	37.29	77.18	2.17	38.09	76.29	1.06
Power PC Theory	37.29	77.18	9.04	38.09	76.29	4.39
RWM-asymptotic	37.29	77.18	2.17	38.09	76.29	1.06
RWM -15 trials	24.10	48.40	1.32	23.83	49.30	0.64
Law of total probability	26.48	52.48	77.40	25.99	57.35	77.28
P(c and e)	26.48	52.48	45.34	25.99	57.35	48.94

Note. Intervention Strategies = Treatment-biased, Unbiased, Observation-independent, Observation-dependent. Bayes theorem = *(P(e|c)); ΔP* Model = *P(e|c)-P(e|~c));* Power PC Theory = *(P(e|c)-P(e|~c))/(1-P(e|~c));* RWM = Rescorla-Wagner Model, see Appendix AC for calculations; Law of total probability = *P(e)*

APPENDIX AE

Additional Findings from Demographic and Debrief Analyses

Treatment Onset and Offset

I asked participants to quantify their expectations toward treatment onset and offset (i.e., the amount of time before the treatment takes effect and the amount of time it takes for the treatment to dissipate) to verify that the number of trials, as well as delay manipulations, fell within this expected timeframe. The majority of participants stated that treatment onset (X^2 (1, n = 52) = 27.77, p < .01, *Cramer's V* = .73) and offset (X^2 (1, n = 51) = 7.08, p = .01, *Cramer's V* = .37) typically occur within the 16 day/trial timeframe of this study. These biases were formed independent of their assigned system condition (Onset: t(50) = -1.06, p = .29, 95% CI [-8.08, 2.49], d = -.30 Offset: t(49) = -0.45, p = .66, 95% CI [-13.24, 8.40], d = -.13). Interestingly, a substantial number of participants expressed reluctance in generalizing both onset (n = 13) and offset (n = 16) timeframes for treatments, stating that temporal dynamics is confounded with the treatment type and the individual. Figure AE1 summarizes participants' general expectations about the treatment onset and offset durations.

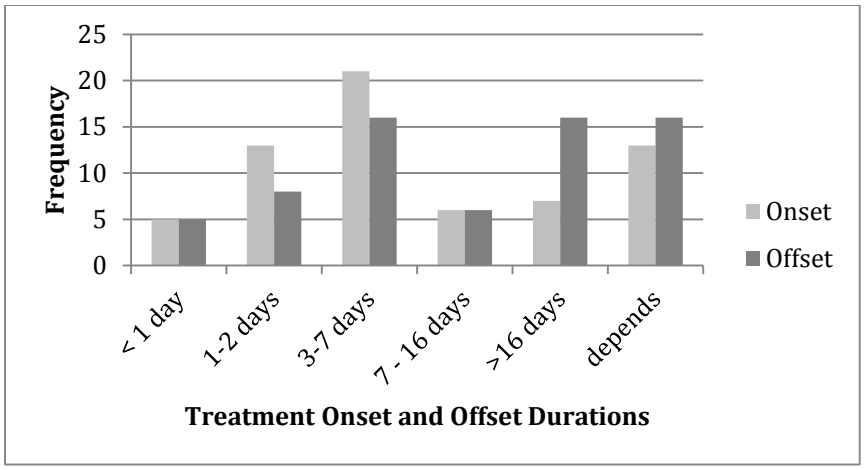


Figure AE1. Expectations about treatment onset and offset. $N = 65$.

Task Difficulty

I asked participants to rate the task difficulty of their final patient scenario on a 7-point graphic ratings scale. There was no significant difference in the task difficult between the discrete ($M = 4.17$, $SD = 1.72$) and continuous ($M = 3.97$, $SD = 1.93$) system conditions: $t(70) = 0.45$, $p = .65$, two-tailed, 95% CI [-0.67, 1.05], $d = 0.11$. I combined the system condition data and performed a 3 (contingency) * 3 (delay) ANOVA. There were no significant differences (Table AE1).

Table AE1

*3 (Contingency) * 3 (Delay) ANOVA - Task Difficult Differences between Experimental Manipulations*

	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Contingency	6.78	2	3.39	1.02	.37	.03
Delay	4.69	2	2.35	0.71	.50	.02
Contingency * Delay	15.06	4	3.76	1.14	.35	.06
Error	208.13	63	3.30			

Note. System data was combined. MSerror and *df* for delay contingency and delay was 3.33 and 67, respectively given no significant interaction.

HBQ Items to Number of Treatment Applications

Relative to our study, participants with higher ratings on the Health Belief Questionnaire (HBQ) items may apply the treatment less frequently when determining how the treatments influence the patients' blood pressure levels. Table AE2 contains the correlational analyses between the HBQ items and the total number of treatment applications participants applied across their three patient scenarios. The item *medicines do no harm than good* was negatively correlated with the number of treatment applications, meaning the participants that agreed that medicines do more applied the treatment less frequently. None of the other items were correlated with the number of treatment applications measure.

Table AE2
Correlational Analyses between Health Belief Questionnaire Items and Total Number of Apply Treatment Interventions

	1. Too Many	2. Natural	3. Harm	4. Stop
1. Too Many				
2. Natural	.352**			
3. Harm	.212	.242*		
4. Stop	.152	.242*	.250*	
5. Apply Treatments	.172	.126	-.243*	-.009

Note. $N = 72$. 1. Too many = Doctors use too many medicines. 2. Natural = Natural remedies are safer than medicines. 3. Harm = Medicines do more harm than good. 4. Stop = People who take medicines should stop their treatment for a while every now and again. 5. Apply Treatments = Sum of the apply treatment interventions across the three patient scenarios. *Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).

APPENDIX AF

Self-reported Assessment Technique in the Judgment Task

Differences in Assessment Technique between System Conditions

Table AF1 presents the outcomes of chi-square goodness of fit tests used to assess differences between the system conditions. Substantially more participants stated that they used information from specific trials when assigning causality with the continuous than with the discrete system. None of the other assessment techniques significantly varied between the system conditions. Potentially more of differences in assessment technique between the system conditions would reach significance if participants were prompted after each judgment task, rather than only after their last patient scenario.

Table AF1
Distribution of Self-Reported Assessment Techniques in the Judgment Task and Outcomes of Chi-Square Goodness of Fit Tests

	Discrete System	Continuous System	$X^2(1)$	p	V
Specific	2	10	5.33	.02	.67
Tally	12	8	0.80	.37	.26
Overall	5	8	0.69	.41	.23
Intuition	6	2	2.00	.16	.50
Unknown	11	8	0.47	.49	.16

Note. Specific = consideration of specific observations, Tally = number of successful or normal blood pressure observations, Overall = consideration of the overall effect of the treatment. *V = Cramer's V.*

Implicit (unknown) vs. Explicit Judgment Task Technique

I assessed if causal attribution differed as a function of the interaction between participants' intervention strategy in the learning task and whether they had an explicit

assessment technique in the judgment task. Unknown responses and references to intuition and were coded as implicit, and the remaining responses were coded as explicit. Table AF2 contains the outcomes of two 2 (intervention strategy) by 2 (judgment task) ANOVAs. None of the interactions were statistically significant. Also, I considered whether the contingency and delay manipulations interacted with the explicitness of participants' technique in the judgment task, again using causal attribution as the dependent measure. There were no significant interactions with these analyses (Table AF3).

Table AF2

*2 (Intervention Strategy) * 3 (Explicitness of Judgment Task Technique) ANOVAs*

	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Intervention Bias * Judgment Task	1287.09	1	1287.09	0.36	.55	< .01
Error	241421.12	67	3603.30			
Observation Dependence * Judgment Task	3566.29	1	3566.29	1.08	.30	0.01
Error	224233.79	68	3297.56			

Note. System data was combined.

Table AF3

*2 (Intervention Strategy) * 3 (Explicitness of Judgment Task Technique) ANOVAs*

	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2
Contingency * Judgment Task	1682.36	2	841.18	0.19	.83	0.01
Error	289112.34	66	4380.49			
Delay * Judgment Task	10750.81	2	5375.41	1.31	.28	0.04
Error	271051.60	66	4106.84			

Note. System data was combined.

APPENDIX AG

Actual Observed Contingency with for Each Treatment Intervention Distribution

Table AG1
Actual Observed Contingency with for Each Treatment Intervention Distribution

Intervention Distribution[1]		Contingency Specification		
Treatment	~Treatment	25|0	75|0	75|75
1	15	1.00|0	1.00|0	1.00|0.73
2	14	0.50|0	0.50|0	0.50|0.71
3	13	0.33|0	0.67|0	0.67|0.77
4	12	0.25|0	0.75|0	0.75|0.75
5	11	0.40|0	0.80|0	0.80|0.73
6	10	0.33|0	0.67|0	0.67|0.70
7	9	0.29|0	0.71|0	0.71|0.78
8	8	0.25|0	0.75|0	0.75|0.75
9	7	0.33|0	0.78|0	0.78|0.71
10	6	0.30|0	0.70|0	0.70|0.67
11	5	0.27|0	0.73|0	0.73|0.80
12	4	0.25|0	0.75|0	0.75|0.75
13	3	0.31|0	0.77|0	0.77|0.67
14	2	0.29|0	0.71|0	0.71|0.50
15	1	0.27|0	0.73|0	0.73|1.00
16	0	0.25|0	0.75|0	0.75|?

Note. Treatment = total number of applications of the treatment. ~Treatment = total number of instances that the treatment was withheld.
[1]Several of these treatment distributions were not observed. The minimum number of treatment selections applied by participants in the discrete and continuous system condition was three and six, respectively.

APPENDIX AH

Runs Tests for Randomness

A series of Wald-Wolfowitz runs tests for randomness on each string of intervention selections suggests that participants' intervention strategy was largely random with both the discrete and continuous system conditions. Table AH1 summarizes the number of non-random strategies organized by experimental manipulation combination. Tables AH2, AH3, and AH4 contain the data and findings pertaining to the discrete system condition. Tables AH5, AH6, and AH7 contain the data and findings pertaining to the continuous system condition. It is important to note that closer examination of the intervention selections identified by the runs tests as random, revealed patterns of behavior that were directly tied to specific observations.

Table AH1

Summary of the Runs Tests for Randomness – Number of Non-Random Intervention Strategies

	Discrete		Continuous	
	Table	z distribution	Table	z distribution
Across Manipulations	9	16	23	36
25|0*				
One-trial delay	0	2	1	2
Two-trial delay	1	1	0	2
Three-trial delay	0	2	3	4
75|0*				
One-trial delay	4	4	2	4
Two-trial delay	1	1	3	5
Three-trial delay	0	2	3	3
75|75 *				
One-trial delay	0	0	4	6
Two-trial delay	0	1	4	4
Three-trial delay	3	3	3	6

Note. $N_{discrete} = 108$, $N_{continuous} = 108$, Discrete = discrete system condition, Continuous = continuous system condition, z distribution = number of instances with $p < .05$ using a z distribution (typically used for samples of over 20), table = number of instances with $p < .05$ using the critical values table for when observations are less than or equal to 20 (Swed & Eisenhart, 1943)

Table AH2

Runs Tests for Randomness – Consideration of Intervention Strategies from the Discrete System Condition with 25|0 Contingency

	Treatment	~Treatment	R	$\bar{R}$	s_R	Z
25|0, one-trial delay	3	13	5	5.88	1.12	-0.78
	14	2	5	4.50	0.76	0.65
	13	3	6	5.88	1.12	0.11
	15	1	3	2.88	0.33	0.38
	16	0	1	1.00	0.00	NA
	11	5	9	7.88	1.64	0.69
	11	5	7	7.88	1.64	-0.53
	12	4	8	7.00	1.41	0.71
	13	3	3	5.88	1.12	-2.56*
	11	5	7	7.88	1.64	-0.53
	13	3	3	5.88	1.12	-2.56*
	13	3	7	5.88	1.12	1.00
25|0, two-trial delay	12	4	6	7.00	1.41	-0.71
	12	4	5	7.00	1.41	-1.41
	14	2	5	4.50	0.76	0.65
	10	6	5	8.50	1.80	-1.94
	9	7	7	8.88	1.90	-0.99
	11	5	3	7.88	1.64	-2.97*°
	13	3	7	5.88	1.12	1.00
	10	6	7	8.50	1.80	-0.83
	13	3	7	5.88	1.12	1.00
	12	4	7	7.00	1.41	0.00
	16	0	1	1.00	0.00	NA
	16	0	1	1.00	0.00	NA
25|0, three-trial delay	10	6	7	8.50	1.80	-0.83
	4	12	4	7.00	1.41	-2.12*
	11	5	7	7.88	1.64	-0.53
	13	3	3	5.88	1.12	-2.56*
	13	3	7	5.88	1.12	1.00
	6	10	8	8.50	1.80	-0.28
	8	8	11	9.00	1.93	1.04
	6	10	6	8.50	1.80	-1.39
	12	4	7	7.00	1.41	0.00
	13	3	7	5.88	1.12	1.00
	6	10	10	8.50	1.80	0.83
	13	3	7	5.88	1.12	1.00

Note. *$p < .05$ using z distribution (typically used for samples of over 20), °$p<.05$ using critical values table for when observations are less than or equal to 20 (Swed & Eisenhart, 1943), Treatment = the number of apply treatment interventions, ~ Treatment = the number of withhold treatment interventions, R = the observed number of runs, $\bar{R}$ = the expected number of runs, s_R = the standard deviation of the number of run.

Table AH3

Runs Tests for Randomness – Consideration of Intervention Strategies from the Discrete System Condition with 75|0 Contingency

	Treatment	~Treatment	R	$\bar{R}$	s_R	Z
75|0, one-trial delay	12	4	9	7.00	1.41	1.41
	15	1	3	2.88	0.33	0.38
	15	1	3	2.88	0.33	0.38
	16	0	1	1.00	0.00	NA
	9	7	14	8.88	1.90	2.70*°
	11	5	8	7.88	1.64	0.08
	9	7	10	8.88	1.90	0.59
	12	4	7	7.00	1.41	0.00
	9	7	14	8.88	1.90	2.70*°
	14	2	3	4.50	0.76	-1.96
	7	9	2	8.88	1.90	-3.62*°
	12	4	3	7.00	1.41	-2.83*°
75|0, two-trial delay	9	7	9	8.88	1.90	0.07
	10	6	7	8.50	1.80	-0.83
	14	2	3	4.50	0.76	-1.96
	16	0	1	1.00	0.00	NA
	12	4	3	7.00	1.41	-2.83*°
	8	8	9	9.00	1.93	0.00
	10	6	9	8.50	1.80	0.28
	10	6	7	8.50	1.80	-0.83
	11	5	8	7.88	1.64	0.08
	10	6	9	8.50	1.80	0.28
	9	7	9	8.88	1.90	0.07
	10	6	9	8.50	1.80	0.28
75|0, three-trial delay	5	11	11	7.88	1.64	1.90
	10	6	9	8.50	1.80	0.28
	9	7	10	8.88	1.90	0.59
	16	0	1	1.00	0.00	NA
	13	3	3	5.88	1.12	-2.56*
	15	1	2	2.88	0.33	-2.65*
	9	7	9	8.88	1.90	0.07
	12	4	9	7.00	1.41	1.41
	11	5	9	7.88	1.64	0.69
	8	8	11	9.00	1.93	1.04
	9	7	10	8.88	1.90	0.59
	7	9	10	8.88	1.90	0.59

Note. *$p < .05$ using z distribution (typically used for samples of over 20), °$p<.05$ using critical values table for when observations are less than or equal to 20 (Swed & Eisenhart, 1943), Treatment = the number of apply treatment interventions, ~ Treatment = the number of withhold treatment interventions, R = the observed number of runs, $\bar{R}$ = the expected number of runs, s_R = the standard deviation of the number of run.

Table AH4

Runs Tests for Randomness – Consideration of Intervention Strategies from the Discrete System Condition with 75|75 Contingency

	Treatment	~Treatment	R	$\bar{R}$	s_R	Z
75|75, one-trial delay	5	11	8	7.88	1.64	0.08
	16	0	1	1.00	0.00	NA
	10	6	6	8.50	1.80	-1.39
	5	11	8	7.88	1.64	0.08
	5	11	7	7.88	1.64	-0.53
	5	11	8	7.88	1.64	0.08
	5	11	8	7.88	1.64	0.08
	11	5	8	7.88	1.64	0.08
	8	8	8	9.00	1.93	-0.52
	16	0	1	1.00	0.00	NA
	5	11	8	7.88	1.64	0.08
	16	0	1	1.00	0.00	NA
75|75 two-trial delay	8	8	8	9.00	1.93	-0.52
	1	15	2	2.88	0.33	-2.65*
	15	1	3	2.88	0.33	0.38
	16	0	1	1.00	0.00	NA
	4	12	9	7.00	1.41	1.41
	7	9	8	8.88	1.90	-0.46
	11	5	8	7.88	1.64	0.08
	15	1	3	2.88	0.33	0.38
	5	11	8	7.88	1.64	0.08
	12	4	6	7.00	1.41	-0.71
	7	9	10	8.88	1.90	0.59
	6	10	10	8.50	1.80	0.83
75|75, three-trial delay	8	8	11	9.00	1.93	1.04
	8	8	8	9.00	1.93	-0.52
	14	2	5	4.50	0.76	0.65
	11	5	2	7.88	1.64	-3.58*°
	15	1	3	2.88	0.33	0.38
	6	10	8	8.50	1.80	-0.28
	10	6	11	8.50	1.80	1.39
	7	9	9	8.88	1.90	0.07
	16	0	1	1.00	0.00	NA
	6	10	5	8.50	1.80	-1.94
	10	6	3	8.50	1.80	-3.05*°
	8	8	3	9.00	1.93	-3.11*°

Note. *$p < .05$ using z distribution (typically used for samples of over 20), °$p<.05$ using critical values table for when observations are less than or equal to 20 (Swed & Eisenhart, 1943), Treatment = the number of apply treatment interventions, ~ Treatment = the number of withhold treatment interventions, R = the observed number of runs, $\bar{R}$ = the expected number of runs, s_R = the standard deviation of the number of run.

Table AH5

Runs Tests for Randomness – Consideration of Intervention Strategies from the Continuous System Condition with 25|0 Contingency

	Treatment	~Treatment	R	$\bar{R}$	s_R	Z
25|0, one-trial delay	7	9	7	8.88	1.90	-0.99
	10	6	2	8.50	1.80	-3.61*°
	12	4	8	7.00	1.41	0.71
	12	4	9	7.00	1.41	1.41
	13	3	5	5.88	1.12	-0.78
	14	2	3	4.50	0.76	-1.96
	9	7	7	8.88	1.90	-0.99
	10	6	9	8.50	1.80	0.28
	12	4	4	7.00	1.41	-2.12*
	12	4	5	7.00	1.41	-1.41
	14	2	3	4.50	0.76	-1.96
	15	1	3	2.88	0.33	0.38
25|0, two-trial delay	6	10	11	8.50	1.80	1.39
	10	6	9	8.50	1.80	0.28
	10	6	5	8.50	1.80	-1.94
	11	5	8	7.88	1.64	0.08
	11	5	7	7.88	1.64	-0.53
	12	4	5	7.00	1.41	-1.41
	8	8	5	9.00	1.93	-2.07*
	9	7	9	8.88	1.90	0.07
	10	6	7	8.50	1.80	-0.83
	13	3	3	5.88	1.12	-2.56*
	14	2	5	4.50	0.76	0.65
	16	0	1	1.00	0.00	NA
25|0, three-trial delay	6	10	8	8.50	1.80	-0.28
	8	8	8	9.00	1.93	-0.52
	9	7	14	8.88	1.90	2.70*°
	11	5	10	7.88	1.64	1.29
	11	5	5	7.88	1.64	-1.75
	14	2	5	4.50	0.76	0.65
	8	8	6	9.00	1.93	-1.55
	9	7	6	8.88	1.90	-1.51
	10	6	7	8.50	1.80	-0.83
	11	5	4	7.88	1.64	-2.36*°
	13	3	2	5.88	1.12	-3.45*°
	13	3	3	5.88	1.12	-2.56*

Note. *$p < .05$ using z distribution (typically used for samples of over 20), °$p<.05$ using critical values table for when observations are less than or equal to 20 (Swed & Eisenhart, 1943), Treatment = the number of apply treatment interventions, ~ Treatment = the number of withhold treatment interventions, R = the observed number of runs, $\bar{R}$ = the expected number of runs, s_R = the standard deviation of the number of run.

Table AH6

Runs Tests for Randomness – Consideration of Intervention Strategies from the Continuous System Condition with 75|0 Contingency

	Treatment	~Treatment	R	$\bar{R}$	s_R	Z
75|0, one-trial delay	9	7	13	8.88	1.90	2.17*
	11	5	11	7.88	1.64	1.90
	12	4	5	7.00	1.41	-1.41
	14	2	3	4.50	0.76	-1.96*
	16	0	1	1.00	0.00	NA
	16	0	1	1.00	0.00	NA
	8	8	14	9.00	1.93	2.59*°
	9	7	15	8.88	1.90	3.22*°
	10	6	10	8.50	1.80	0.83
	11	5	6	7.88	1.64	-1.14
	13	3	6	5.88	1.12	0.11
	16	0	1	1.00	0.00	NA
75|0, two-trial delay	10	6	5	8.50	1.80	-1.94
	11	5	7	7.88	1.64	-0.53
	12	4	5	7.00	1.41	-1.41
	13	3	2	5.88	1.12	-3.45*°
	13	3	5	5.88	1.12	-0.78
	15	1	2	2.88	0.33	-2.65*
	6	10	8	8.50	1.80	-0.28
	8	8	2	9.00	1.93	-3.62*°
	13	3	2	5.88	1.12	-3.45*°
	13	3	3	5.88	1.12	-2.56*
	16	0	1	1.00	0.00	NA
	16	0	1	1.00	0.00	NA
75|0, three-trial delay	6	10	6	8.50	1.80	-1.39
	10	6	13	8.50	1.80	2.50*°
	12	4	2	7.00	1.41	-3.54*°
	14	2	5	4.50	0.76	0.65
	15	1	3	2.88	0.33	0.38
	16	0	1	1.00	0.00	NA
	8	8	4	9.00	1.93	-2.59*°
	9	7	10	8.88	1.90	0.59
	10	6	5	8.50	1.80	-1.94
	11	5	5	7.88	1.64	-1.75
	15	1	3	2.88	0.33	0.38
	16	0	1	1.00	0.00	NA

Note. *$p < .05$ using z distribution (typically used for samples of over 20), °$p<.05$ using critical values table for when observations are less than or equal to 20 (Swed & Eisenhart, 1943), Treatment = the number of apply treatment interventions, ~ Treatment = the number of withhold treatment interventions, R = the observed number of runs, $\bar{R}$ = the expected number of runs, s_R = the standard deviation of the number of run.

Table AH7

Runs Tests for Randomness – Consideration of Intervention Strategies from the Continuous System Condition with 75|75 Contingency

	Treatment	~Treatment	R	$\bar{R}$	s_R	Z
75|75, one-trial delay	6	10	2	8.50	1.80	-3.61*°
	12	4	3	7.00	1.41	-2.83*°
	13	3	6	5.88	1.12	0.11
	14	2	3	4.50	0.76	-1.96
	15	1	2	2.88	0.33	-2.65*
	16	0	1	1.00	0.00	NA
	5	11	2	7.88	1.64	-3.58*°
	5	11	8	7.88	1.64	0.08
	8	8	4	9.00	1.93	-2.59*°
	9	7	8	8.88	1.90	-0.46
	13	3	3	5.88	1.12	-2.56*
	16	0	1	1.00	0.00	NA
75|75 two-trial delay	3	13	4	5.88	1.12	-1.67
	5	11	6	7.88	1.64	-1.14
	8	8	9	9.00	1.93	0.00
	9	7	2	8.88	1.90	-3.62*°
	16	0	1	1.00	0.00	NA
	16	0	1	1.00	0.00	NA
	5	11	8	7.88	1.64	0.08
	9	7	3	8.88	1.90	-3.09*°
	9	7	2	8.88	1.90	-3.62*°
	13	3	2	5.88	1.12	-3.45*°
	16	0	1	1.00	0.00	NA
	16	0	1	1.00	0.00	NA
75|75, three-trial delay	1	15	2	2.88	0.33	-2.65*
	3	13	3	5.88	1.12	-2.56*
	6	10	4	8.50	1.80	-2.50*°
	10	6	4	8.50	1.80	-2.50*°
	11	5	11	7.88	1.64	1.90
	14	2	2	4.50	0.76	-3.27*°
	6	10	12	8.50	1.80	1.94
	7	9	11	8.88	1.90	1.12
	8	8	9	9.00	1.93	0.00
	9	7	11	8.88	1.90	1.12
	15	1	2	2.88	0.33	-2.65*
	16	0	1	1.00	0.00	NA

Note. *$p < .05$ using z distribution (typically used for samples of over 20), °$p<.05$ using critical values table for when observations are less than or equal to 20 (Swed & Eisenhart, 1943), Treatment = the number of apply treatment interventions, ~ Treatment = the number of withhold treatment interventions, R = the observed number of runs, $\bar{R}$ = the expected number of runs, s_R = the standard deviation of the number of run.

APPENDIX AI

Effect of System on Strategy with Data Separated by each Experimental Manipulation Combination

Table AI1

Distribution of Participants' Intervention Bias and Outcomes of Fisher's Exact Tests

	Discrete System		Continuous System		
	Unbiased	Biased	Unbiased	Biased	*p*
25|0 and					
One-trial delay	3	8	4	8	1.00
Two-trial delay	4	8	8	4	.22
Three-trial delay	6	5	9	3	.40
75|0 and					
One-trial delay	5	7	6	6	1.00
Two-trial delay	9	3	4	8	.10
Three-trial delay	8	4	6	6	.68
75|75 and					
One-trial delay	9	3	5	7	.21
Two-trial delay	6	4	6	5	1.00
Three-trial delay	9	3	7	3	1.00

Note. Biased = Treatment-biased intervention strategy. All tests two-tailed.

Table AI2

Distribution of Participants' Observation Dependence and Outcomes of Fisher's Exact Tests

	Discrete System		Continuous System		
	Independent	Dependent	Independent	Dependent	*p*
25|0 and					
One-trial delay	6	6	6	6	1.00
Two-trial delay	4	8	4	8	1.00
Three-trial delay	4	8	6	6	.68
75|0 and					
One-trial delay	10	2	9	3	1.00
Two-trial delay	3	9	10	2	.01
Three-trial delay	3	9	11	1	< .01
75|75 and					
One-trial delay	3	9	10	2	.01
Two-trial delay	4	8	9	3	.10
Three-trial delay	8	4	9	3	1.00

Note. Independent = Observation-independent intervention strategy; Dependent = Observation-dependent intervention strategy. All tests two-tailed.

APPENDIX AJ

Fisher's Exact Tests using Nominal Attributions

To help understand why the causal attributions were more extreme with the continuous system, I compared the abstract, nominal attributions related to the 25|0 and 75|0 contingency manipulations using Fisher's Exact Tests. I considered the number of participants that selected raised, no effect, and lowered separately as their nominal attribution in combination with the 25|0 and 75|0 contingency manipulations in both system conditions. The outcomes of these tests indicate that attributions are formed independent of system condition (Table AJ1).

Table AJ1
Frequency of Abstract, Nominal Attributions separated by the System Condition and Contingency and Outcomes of Chi-square Tests for Independence

	Discrete System		Continuous System			
	25|0	75|0	25|0	75|0	*p*	*p***
Raised	10	6	15	1	.08	
No Effect	14	6	14	0	NA	.11
Lowered	12	24	7	35	.11	

Note. *n(raised)* = 32, *n(no-effect)* = 34, *n(lowered)* = 78. ***p* = Fisher's Exact test calculated with adjusted data (i.e., +1 was added to every cell). Raised = nominal selection that the treatment raised the patient's blood pressure level, No effect = nominal selection that the treatment had no effect on the patient's blood pressure level, Lowered = nominal selection that the treatment lowered the patient's blood pressure level. All tests two tailed.

REFERENCES

Allan, L. G., Tangen, J. M., Wood, R., & Shah, T. (2003). Temporal contiguity and contingency judgments: A Pavlovian analogue. *Integrative Physiological & Behavioral Science*, *38*, 214-229. doi: 10.1007/BF02688855

Bradley, J. V. (1968). *Distribution-Free Statistical Tests.* Englewood Cliffs, NJ: Prentice-Hall.

Brehmer, B. (1989). Feedback delays and control in complex dynamic systems. Paper presented at the Computer-based management of complex systems in international system dynamics conference, Berlin: Springer-Verlag. Retrieved from http://www.systemdynamics.org/conferences/1989/proceed/parallel%20sessions%20pdf/brehm189.pdf

Buehner, M. J., & Cheng, P. W. (2005). Causal learning. *The Cambridge Handbook of Thinking and Reasoning*, 143-168.

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. (2003). Rethinking temporal contiguity and the judgment of causality: Effects of prior knowledge, experience, and reinforcement procedure. *The Quarterly Journal of Experimental Psychology*, *56*, 865-890. doi: 10.1080/02724980244000675

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

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17,* 391-416. doi: 10.1016/0010-0285(85)90014-3

Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality Social Psychology*, *58*, 545-567. doi: 10.1037/0022-3514.58.4.545

Chobanian, A. V., Bakris, G. L., Black, H. R., Cushman, W. C., Green, L. A., Izzo, J. L., …Wright, J. T. (2003). The seventh report of the Joint National Committee on prevention, detection, evaluation and treatment of high blood pressure: the JNC 7 report. *The Journal of the American Medical Association, 289,* 2560–2572. doi: 10.1001/jama.289.19.2560

Claflin, D. I., Garrett, T., & Buffington, M. L. (2005). A developmental comparison of trace and delay eyeblink conditioning in rats using matching interstimulus intervals. *Developmental Psychobiology*, *47*, 77-88. doi: 10.1002/dev.20068

Diehl, E., & Sterman, J. D. (1995). Effects of feedback complexity on dynamic decision making. *Organizational Behavior and Human Decision Processes*, *62*, 198-215. doi: 10.1006/obhd.1995.1043

Edwards, W. D. (1962). Dynamic decision theory and probabilistic information processing. *Human Factors, 4*, 59–73. doi: 10.1177/001872086200400201

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

Flach, J. M. (2015). [Possibilities for System Dynamics]. Unpublished figure.

Frosch, C. A., McCormack, T., Lagnado, D. A., & Burns, P. (2012). Are causal structure and intervention judgments inextricably linked? A developmental study. *Cognitive Science*, *36*, 261-285. doi: 10.1111/j.1551-6709.2011.01208.x

Gao, Y., Nitzany, E., & Edelman, S. Online learning of causal structure in a dynamic game situation. Retrieved July 4, 2013, from: http://mindmodeling.org/cogsci2012/papers/0076/paper0076.pdf

Gonzalez, C., & Dutt, V. (2011). A generic dynamic control task for behavioral research and education. *Computers in Human Behavior*, *27*, 1904-1914. doi: 10.1016/j.chb.2011.04.015

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, *37*, 620-629. doi: 10.1037/0012-1649.37.5.620

Greville, W. J., & Buehner, M. J. (2010). Temporal predictability facilitates causal learning. *Journal of Experimental Psychology General*, *139*, 756-771. doi: 10.1037/a0020976

Hagmayer, Y., Meder, B., Osman, M., Mangold, S., & Lagnado, D. (2010). Spontaneous causal learning while controlling a dynamic system. *Psychology*, *3*, 145-162. doi: 10.2174/1874350101003020145

Horne, R., Weinman, J., & Hankins, M. (1999). The beliefs about medicines questionnaire: The development and evaluation of a new method for assessing the cognitive representation of medication. *Psychology and Health, 14,* 1 - 24. doi: 10.1080/08870449908407311

Imhof, B., Scheiter, K., Edelmann, J., & Gerjets, P. (2012). How temporal and spatial aspects of presenting visualizations affect learning about locomotion patterns. *Learning & Instruction*, *22*, 193-205. doi: 10.1016/j.learninstruc.2011.10.006

Imhof, B., Scheiter, K., & Gerjets, P. (2011). Learning about locomotion patterns from visualizations: Effects of presentation format and realism. *Computers & Education*, *57*, 1961-1970. doi: 10.1016/j.compedu.2011.05.004

Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs: General and Applied*, *79*, 1-17. doi: 10.1037/h0093874

Jensen, E., & Brehmer, B. (2003). Understanding and control of a simple dynamic system. *System Dynamics Review, 19*, 119–137. doi: 10.1002/sdr.267

Johnson-Laird, P. N., Legrenzi, P., & Legrenzi, M. S. (1972). Reasoning and a sense of reality. *British Journal of Psychology*, *63*, 395-400. doi: 10.1111/j.2044-8295.1972.tb01287.x

Kahneman, D. (2003). A perspective on judgment and choice. *American Psychologist, 58,* 697–720. doi: 10.1037/0003-066x.58.9.697.

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*, 107-128. doi: 10.1037/h0034225

Kleinmuntz, D. N., & Thomas, J. B. (1987). The value of action and inference in dynamic decision making. *Organizational Behavior and Human Decision Processes*, *39*, 341-364. doi: 10.1016/0749-5978(87)90028-8

Lagnado, D. A., & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 451-460. doi: 10.1037/0278-7393.32.3.451

Lagnado, D.A. & Sloman, S.A. (2002). Learning causal structure. In W. Gray & C. D. Schunn (Eds.), *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society,* 828-833. Retrieved July 4, 2013, from http://www.ucl.ac.uk/lagnado-lab/david_lagnado.html.

Lagnado, D.A. & Speekenbrink, M. (2010). The influence of delays in real-time causal learning. *Psychology*, *3*, 184-195. doi:10.2174/1874350101003020184

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge, UK: Cambridge University Press.

Level test. (2013, October 30). Oxford, UK: Macmillan Publishers. Retrieved from http://www.macmillanreaders.com/tests/level-test

Lippa, K. D., Klein, H. A., & Shalin, V. L. (2008). Everyday expertise: Cognitive demands in diabetes self-management. *Human Factors*, *50*, 112-120. doi: 10.1518/001872008X250601

Mayrhofer, R., & Waldmann, M. R. (2011). Heuristics in covariation-based induction of causal models: Sufficiency and necessity priors. In C. H. Carlson & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, 3110–3115. Retrieved July 4, 2013, from http://mindmodeling.org/cogsci2011/papers/0713/paper0713.pdf

Novick, L. R., & Cheng, P. W. (2004). Assessing interactive causal influence. *Psychological Review*, *111*, 455-485. doi: 10.1037/0033-295X.111.2.455

Ostchega, Y., Yoon, S. S., Hughes, J., & Louis, T. (2008). Hypertension awareness, treatment, and control - continued disparities in adults: United States, 2005–2006. *NCHS Data Brief, 3,* 1-8. Retrieved July 4, 2013, from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19389317

Paich, M., & Sterman, J. D. (1993). Boom, bust, and failures to learn in experimental markets. *Management Science*, *39*, 1439-1458. doi: 10.1287/mnsc.39.12.1439

Perales, J. C., & Shanks, D. R. (2003). Normative and descriptive accounts of the influence of power and contingency on causal judgment. *The Quarterly Journal of Experimental Psychology*, *56*, 977-1007. doi: 10.1080/02724980244000738

Rottman, B. M., & Ahn, W.K. (2011). Effect of grouping of evidence types on learning about interactions between observed and unobserved causes. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *37*, 1432-1448. doi: 10.1037/a0024829

Rottman, B. M., & Ahn, W. K. (2009). Causal inference when observed and unobserved causes interact. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, 1477-1482.

Rottman, B. M., & Keil, F. C. (2012). Causal structure learning over time: Observations and interventions. *Cognitive Psychology*, *64*, 93-125. doi: 10.1016/j.cogpsych.2011.10.003

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*, 139-159. doi: 10.1080/14640748908401189

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*, 127-190. doi: 10.1037/0033-295X.84.2.127

Sterman, J. D. (1989). Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes*, *43*, 301-335. doi: 10.1016/0749-5978(89)90041-1

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science, 27*, 453-489. doi: 10.1207/s15516709cog2703_6

Swed, F.S. and Eisenhart, C. (1943). Tables for testing randomness of grouping in a sequence of alternatives. *Annals of Mathematical Statistics, 14,* 83-86. Retrieved June, 1 2015, from http://www4.nau.edu/shustercourses/BIO%20682/Documents/Handouts/runs%20test%20values.pdf

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124 – 1131. doi: 10.1126/science.185.4157.1124

Wasserman, E. A., Elek, S. M., Chatlosh, D. L., & Baker, A. G. (1993). Rating causal relations: Role of probability in judgments of response-outcome contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 174-188. doi: 10.1037/0278-7393.19.1.174

Wasserman, E. A., & Neunaber, D. J. (1986). College students' responding to and rating of contingency relations: The role of temporal contiguity. *Journal of the Experimental Analysis of Behavior*, *46*, 15-35. doi: 10.1901/jeab.1986.46-15

Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors*, *5*, 7-17. doi: 10.1177/001872086300500102

Yoon, S. S., Burt, V., Louis, T., & Carroll, M. D. (2012). Hypertension among adults in the United States, 2009–2010. *NCHS Data Brief, 107,* 1-8. Retrieved from http://www.cdc.gov/nchs/data/databriefs/db107.htm