Cognitive Psychology 79 (2015) 102–133

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Strategies to intervene on causal systems are adaptively selected

Anna Coenen *, Bob Rehder, Todd M. Gureckis

Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, United States

A R T I C L E   I N F O

Article history:
Accepted 18 February 2015
Available online 15 May 2015

Keywords:
Causal learning
Interventions
Hypothesis testing
Information gain
Self-directed learning

A B S T R A C T

How do people choose interventions to learn about causal systems? Here, we considered two possibilities. First, we test an information sampling model, *information gain*, which values interventions that can discriminate between a learner's hypotheses (i.e. possible causal structures). We compare this discriminatory model to a *positive testing strategy* that instead aims to confirm individual hypotheses. Experiment 1 shows that individual behavior is described best by a mixture of these two alternatives. In Experiment 2 we find that people are able to adaptively alter their behavior and adopt the discriminatory model more often after experiencing that the confirmatory strategy leads to a subjective performance decrement. In Experiment 3, time pressure leads to the opposite effect of inducing a change towards the simpler positive testing strategy. These findings suggest that there is no single strategy that describes how intervention decisions are made. Instead, people select strategies in an adaptive fashion that trades off their expected performance and cognitive effort.

1. Introduction

Causal knowledge underlies our intuitive grasp of physics ("Heat causes water to turn to steam."), technology ("This button causes it to go."), and helps us understand our fellow human beings

* Corresponding author.
E-mail addresses: anna.coenen@nyu.edu (A. Coenen), bob.rehder@nyu.edu (B. Rehder), todd.gureckis@nyu.edu (T.M. Gureckis).

http://dx.doi.org/10.1016/j.cogpsych.2015.02.004
0010-0285/

("Hunger causes her to be grumpy."). Often, the only way to find out about the causal structure of the world is by manipulating individual variables, and observing the effects of this manipulation. For example, banning sugary drinks can help decide whether they are a cause of diabetes. These decisions to manipulate a system are known as *interventions* (Pearl, 2000) and psychological research has recently explored how people use these interventions to learn (Bonawitz et al., 2010; Lagnado & Sloman, 2004; Schulz, Gopnik, & Glymour, 2007; Sloman & Lagnado, 2005; Waldmann & Hagmayer, 2005).

Most research into how people make causal intervention decisions has implicitly sought to identify the *single* strategy that characterizes people's choices best across one or more experiments. For example, one proposal is that people search for information that can *discriminate* between possible hypotheses about causal structure, for instance by using an *information gain* (IG) strategy (Bramley, Lagnado, & Speekenbrink, 2014; Nelson, 2005; Shafto, Goodman, & Griffiths, 2014; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003). Alternatively, in the broader hypothesis testing literature many studies argue that people seek information that yields positive evidence to *confirm* a single hypothesis, disregarding alternatives (e.g., Klayman & Ha, 1987; Nickerson, 1998; Wason, 1960). This mode of search is often referred to as the *positive test strategy* or PTS because it favors queries that are expected to yield a positive response ("yes", rather than "no") given a single hypothesis. A survey of the literature on information gathering during learning reveals forceful arguments for each of these alternatives (Gureckis & Markant, 2012; Nickerson, 1998) even though the division between these perspectives is not always precise (Navarro & Perfors, 2011; Oaksford & Chater, 1994).

The present paper begins from a slightly different perspective from this past work. In particular, we first ask if any single strategy model provides a plausible account of intervention-based causal learning. To that end, we describe a new hierarchical Bayesian method of identifying decision strategies during causal intervention learning. Using the model, we present evidence that individual participants adopt a *mixture* of strategies when learning through causal interventions (Experiment 1). Next, we ask if such mixtures are stable biases in the way people approach such tasks or if they change in response to environmental factors. Our second and third experiments show that strategy choice can change adaptively depending on the current task environment. Such adaptive adjustment of intervention-based strategies is unanticipated by single strategy models and suggests simple manipulations which might improve the quality of human reasoning.

1.1. Two perspectives on information gathering

Efficient learning from causal interventions is ultimately a problem of information search. The learner must decide which intervention to perform in order to gain information about a system's causal structure. The following section describes two theories of how people make such decisions and how they relate to the task of causal intervention learning.

1.1.1. Discriminatory: information gain

The first strategy considered here is based on a rational analysis of the structure learning task (Anderson, 1990; Chater & Oaksford, 2008; Marr, 1982). According to this perspective, people should choose interventions that will be maximally useful for distinguishing alternative hypotheses.

To illustrate, consider playing the children's game "Guess Who?". In this game, one player adopts a secret identity (e.g., a fictional character or a celebrity). The job of the other players is to reveal this identity as quickly as possible by asking questions that can be answered with a "yes" or "no". The space of possible hypotheses (identities) is large in the beginning, but can be reduced by asking revealing questions. For example "Is the character male?" is a useful question because (assuming the learner expects a roughly even split of males and females) either answer will cut the number of identities in half. In contrast, very specific questions like "Does the character have pointy ears?" is a lot less informative, because the likelihood of "yes" is very low, and a "no" does not reduce the hypothesis space by much (most people do not have pointy ears). Similarly, a too general query like "Does this character have eyes?" will do little to narrow down the number of plausible hypotheses, because it is true of most.

The intuitive difference between “good” and “bad” questions is captured in a number of formal models of information gathering (Nelson, 2005). The information gain model (IG) in particular has been advanced as a normative strategy for a wide number of information search tasks (e.g., Austerweil & Griffiths, 2011; Gureckis & Markant, 2009; Klayman & Ha, 1987; Lindley, 1956; Markant & Gureckis, 2012a, 2012b; Najemnik & Geisler, 2005; Nelson, Divjak, Gudmundsdottir, Martignon, & Meder, 2014; Oaksford & Chater, 1994, 1994), and was first applied to modeling causal interventions in the machine-learning literature (Murphy, 2001; Tong & Koller, 2001). It has also had an important influence on the psychology of learning from causal interventions (Bramley et al., 2014; Shafto et al., 2014; Steyvers et al., 2003). According to IG, interventions are made with the goal of decreasing the learner’s uncertainty about a causal system, given a range of possible structures that explain its behavior.

To illustrate the core psychological claims of IG it is helpful to review the underlying formalism. Consider a learner with a range of hypotheses about how the variables of a system might be connected with one another. Each hypothesis is a different causal structure. These structures can be represented as *causal Bayesian networks*, that is, directed acyclic graphs, in which the state of each variable is a function of the state of its direct parents (Pearl, 2000). IG prescribes that learners should choose the intervention that minimizes their uncertainty about which graph, $g \in G$, is most likely to underlie a set of observations of a system. A learner’s current uncertainty, $H(G)$, can be measured using the Shannon entropy defined over possible graphs:

$$H(G) = \sum_{g \in G} P(g) \log_2 \frac{1}{P(g)} \tag{1}$$

The probability of each graph, $P(g)$, is the learner’s subjective belief that graph g is the true generative process underlying the causal system. This value can be informed by the learner’s prior belief or previous observations of the causal system, for example. $H(G)$ equals zero if $P(g)$ equals 1 for one of the graphs (and 0 for all the others), that is, when there exists no uncertainty about which hypothesis is true. It reaches its maximum value when all graphs in G are considered to be equally likely.

A decision maker attempting to learn as much as possible about the system should choose an intervention a that maximally reduces $H(G)$. [1] The difference in uncertainty before and after an intervention is the information gain associated with that intervention. It depends on the type of action or intervention that was made, $a \in A$, as well as the outcome, $o \in O$, that occurred as a result of that action:

$$IG(a, o) = H(G) - H(G|a, o) \tag{2}$$

Because the outcome o is often uncertain before the intervention, a choice policy needs to take into account the information gain of all possible outcomes weighted by their probability of occurring, yielding the *expected information gain* of each possible intervention:

$$EIG(a) = H(G) - \sum_{o \in O} P(o|a) H(G|a, o) \tag{3}$$

where $P(o|a)$ is the probability of outcome o given action a. Calculating EIG requires knowing the new uncertainty after making intervention a and observing outcome o:

$$H(G|a, o) = \sum_{g \in G} P(g|a, o) \log_2 \frac{1}{P(g|a, o)} \tag{4}$$

where $P(g|a, o)$ is the probability of graph g given intervention a and resulting outcome o. To calculate $P(g|a, o)$, Bayes’ rule can be applied yielding $P(g|a, o) = P(o|g, a)P(g)/P(o|a)$. Finally, $P(o|a)$ can be computed by marginalizing over all possible graphs and their likelihood of producing outcome o given intervention a, $P(o|g, a)$.

Maximizing IG means choosing interventions with the highest expected information gain calculated using Eq. (3). However, it is likely that people do not always maximize information gain perfectly

[1] We follow prior work in assuming that such decision are *greedy* in that they choose to reduce uncertainty maximally on each trial, ignoring the possibility of multi-step decision strategies.

and may instead select interventions somewhat probabilistically. For example, applications of IG often assume that people choose options in proportion to their expected value using the softmax choice rule (Sutton & Barto, 1998):

$$P(a) = \frac{\exp(V(a)/\tau)}{\sum_i \exp(V(a_i)/\tau)} \tag{5}$$

where $V(a)$ is the value of an intervention (here: its expected information gain). The parameter τ captures the degree of probabilistic responding, ranging from complete guessing ($\tau \to \infty$) to always choosing the option with the highest value ($\tau \to 0$).

In sum, the IG model makes predictions about which intervention an individual should perform to maximally reduce uncertainty. Note that IG is inherently a *discriminatory* strategy because it favors interventions that lead to different outcomes under different hypotheses and avoids producing outcomes that are likely to occur under multiple hypotheses.

However, IG is by no means the only discriminatory sampling strategy that can be applied to causal interventions (other candidates are discussed for example in Nelson, 2005). We focus on it because it has been advanced as a strong candidate model in the causal reasoning and information search literature and because it provides an intuitive account of what it means to ask discriminatory questions about a set of hypotheses. It also turns out, however, that its predictions for the set of problems used in this article are very similar to at least one other prominent sampling norm, *probability gain* (whose empirical predictions will also be tested below). Thus, our conclusions about IG may generalize to a wider range of discriminatory models.

1.1.2. Confirmatory: positive test strategy

As mentioned, a large body of research has shown that people often seek confirmatory information that only pertains to one specific hypothesis. This strategy is often referred to as the positive test strategy (PTS) or *positivity bias* (Klayman & Ha, 1987). For example, in Wason's *2-4-6 task* (Wason, 1960) participants have to guess the rule that generated a sequence of numbers by proposing new sequences and receiving yes/no answers. Given the example sequence 2-4-6, participants often form a strong initial hypothesis that the rule is *increasing even numbers* and tend to test only positive examples of this hypothesis (e.g., 8-10-12) instead of counterexamples, like 1-2-3 or 6-4-2. In the 2-4-6 task, PTS leads participants to ignore alternative hypotheses about the rule, such as *increasing numbers* and hence positive testing is often treated as an obstacle to learning (Nickerson, 1998). Note, however, that there are conditions under which it can be a successful strategy (Ginzburg & Sejnowski, 1996; Klayman & Ha, 1989; Navarro & Perfors, 2011; Nelson, Tenenbaum, & Movellan, 2001; Oaksford & Chater, 1994).

Despite the large literature on PTS, it is unclear how this strategy might manifest itself in a causal structure learning task. We propose that PTS leads to a preference for interventions that have high *causal centrality* in a hypothesis that is currently under evaluation (Ahn, Kim, Lassaline, & Dennis, 2000; Kim & Ahn, 2002; Sloman, Love, & Ahn, 1998). Causal variables, or *nodes* in the causal graph, are considered central if they have a large number of direct and indirect descendant causal links. If learners make interventions on central nodes, they can gather positive evidence for a causal hypothesis by activating those links, that is by producing a large number of expected effects (assuming causal strengths are equal for all links). Because causal graph hypotheses can differ in the number of links that can be activated in principle, we will consider causal centrality relative to the total number of possible links in a given structure. Thus, the PTS value of intervening on a node n is determined by that node's maximum relative causal centrality over all graphs that are currently under consideration:

$$\mathrm{PTS}_n = \max_g \left[\frac{DescendantLinks_{n,g}}{TotalLinks_g}\right] \tag{6}$$

where descendant links are all links that lead to direct or indirect children of n.

To illustrate, a node will have a value of 1 if, by intervening on it, all possible links of at least one graph can be activated (for example through the root node of a chain). If an intervention can activate at most one out of two links, it receives a score of .5. A score of zero means it yields no outcomes whatsoever (if the node has no children). The maximum operator in Eq. (6) means that nodes become attractive to intervene on if they have a high relative centrality in *at least one* hypothesis that is currently evaluated, irrespective of the differences between hypotheses. Although the maximum operator in Eq. (6) implies that learners consider both hypotheses in choosing an intervention, the intervention that is favored is one that confirms the largest proportion of links in one hypothesis (rather than, say, the largest proportion in all hypotheses), consistent with a preference for testing single hypotheses implied by PTS. Appendix B will present some minor variations of Eq. (6) (e.g., replacing the maximum operator with a sum) and assess whether they provide better accounts of the data from the upcoming experiments. To derive concrete choice probabilities from this model, the same softmax choice rule as described in Eq. (5) can be applied by substituting $V(a)$ with PTS scores from Eq. (6).

There exists a clear parallel between this strategy and positive test strategies in classic rule learning tasks if causal effects are taken as analogues to the positive responses ("yes") expected when testing a rule. The desire to "receive a yes response", which lies at the heart of positive testing, corresponds to a desire to "make effects happen" in the causal learning scenario. In reality, non-effects can be just as discriminating as effects, but they do not count as positive examples of a specific graph under this definition.

It is worth emphasizing again that PTS is not a discriminatory strategy. Interventions become attractive in virtue of their position in individual hypotheses, irrespective of whether they help distinguish hypotheses from each other. Furthermore, in all of the experiments reported here the average expected choice accuracy (when simulating an optimal learner) after selecting an intervention with IG is always as least as high as that obtained from PTS. This holds for different values of τ in Eq. (5), as will be reported in more detail below (see Table 1). While this seems intuitive, since PTS is not a discriminatory strategy, it is still worth pointing out, because the same might not hold in other domains of information search.

Eq. (6) is a first attempt to capture the essence of confirmatory information search for causal interventions, one that turns out to work well for our task and empirical results. Section 5 (and Appendix B) will discuss alternatives that may serve the same purpose.

1.2. Empirically distinguishing the models

To what degree do IG and PTS make different predictions? This question is important in light of recent analyses showing that discriminatory hypothesis testing and confirmation can make the same predictions in certain environments (Austerweil & Griffiths, 2011; Navarro & Perfors, 2011; Oaksford & Chater, 1994). Consider panel A of Fig. 1. Given the two hypothesized graphs depicted at the top of the panel (H1 and H2) and the opportunity to make a single intervention, both models predict a preference for intervening on the first two nodes (predictions plotted below the graphs are normalized values of IG and PTS scores calculated using Eqs. (3) and (6), respectively). According to IG, intervening on n1 or n2 could reveal the direction of the link between them. Since n3 has no causal effects under either hypothesis, intervening on this node is expected to have the same outcome. PTS equally values n1 and n2 because these nodes are causally central in either graph (n2 is central in H1 because it can affect both n1 and n3, n1 is causally central in H2 because it can affect n2 and n3). Thus, data from this problem cannot distinguish IG and PTS.

On the other hand, their predictions diverge on the example shown on the right (panel B). IG is agnostic between the three nodes, because intervening on each of them could lead to an outcome that is not expected under one of the hypotheses. However, PTS predicts a preference for the root nodes of the two graphs. In particular, for the top hypothesis (H1), n1 is preferred because it is more likely to

Table 1
Average posterior probability of the most likely graph after choosing a *single* intervention with a given strategy, averaged over the problems used in Experiment 1.

Model	arg max	$\tau = .2$	$\tau = .5$	$\tau = 1$	$\tau = 1.5$
IG	.91	.90	.86	.83	.83
PTS	.85	.86	.84	.83	.82
Random	.80	.80	.80	.80	.80

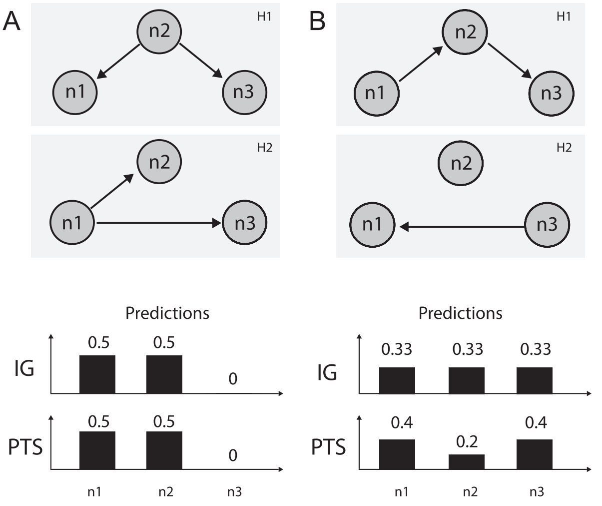

Fig. 1. Example problems with two hypotheses (H1 and H2), and predictions of IG and PTS. Panel A shows two structures for which IG and PTS predict the same interventions. Panel B shows two structures for which the predictions diverge, showing that the models are in principle distinguishable in certain cases.

activate all the links of the graph (i.e., intervening on n1 should affect n2 and, indirectly, n3). In H2, n3 is preferred for the same reason (it tests all the model's links, namely, one of them). Our experiments test a number of problems like this one, for which PTS and IG make different predictions.

1.3. Previous efforts to model intervention decisions

Few studies have directly examined the decision-making processes involved in causal intervention learning. One notable example is a study by Steyvers et al. (2003) who analyzed people's interventions in an "alien mind reading" task using IG. In this task, participants first observed the behavior of a causal system with three variables (aliens reading the content of each others' minds) a number of times, then indicated their favorite hypothesis (i.e., whose mind(s) each alien could read), and then chose one intervention (plant a thought in one alien's head and observe who reads it correctly). With three variables (aliens) participants could in principle entertain 18 different causal structure hypotheses (assuming only one or two links are possible). Although the authors endorsed IG as a good description of the intervention data, the model only fit when additional assumptions were introduced. Most importantly, it was assumed that people tested their favorite hypothesis against its subgraphs, with very low prior beliefs about any of the other 18 hypotheses. Given this post-hoc assumption (participants were not instructed to consider subgraphs) the predictions of this model actually strongly resembled those of PTS. That is, it favored interventions on central nodes that lead to many expected effects in a participant's favorite graph(s) (see Fig. 8 in Steyvers et al., 2003). A pure version of the IG model using the full posterior distribution over possible graphs after the observation phase did not fit the data well.

Lagnado and Sloman (2006) reported the frequencies with which participants chose different interventions (see Table 5 in their article). Again, the majority of interventions were made on high centrality nodes with many downstream effects, such as causes with multiple effects or root nodes of causal chains. A similar result was reported in a study by Hagmayer and Meder (2012) (see Table 1). While these patterns indicate the possible influence of PTS strategies, these papers did not explicitly model participants' intervention decisions.

One common feature of these experiments (also see Bramley et al., 2014; Sobel & Kushnir, 2006) is that the space of possible causal hypotheses that participants could entertain was left unconstrained or was quite large (e.g., the 18 possible graphs in Steyvers et al., 2003). Computing the full IG model for all hypotheses thus required simulating and comparing the outcomes of a very large number of causal graphs. Such computations may have made it difficult for participants to use IG in the first place and could have encouraged a preference for high-centrality interventions, as predicted by PTS. We address this issue in the present experiments by using a much simpler experimental task with a smaller number of explicitly enumerated hypotheses to make sure that it would be relatively easy to compute a discriminatory strategy.

1.4. The current study

In Experiment 1 we aimed to evaluate whether IG or PTS individually provide a credible fit to people's intervention decisions. Unlike past work, we took seriously the possibility that people might use a mixture of strategies and so developed a modeling framework that characterizes this mixture at an individual and group level. Using this framework, we then report two further experiments that manipulate aspects of the task and assess their effect on the parameters in the model. Experiment 2 investigated whether intervention decisions can be modified by changing the strategies' expected payoff. By selectively decreasing the expected success that PTS would yield in comparison to IG during one part of the experiment we could test if participants adaptively change their strategy to be more like PTS in a later part. Experiment 3 tested whether it is also possible to change participants' behavior to be more in line with PTS than IG by adding time pressure. This manipulation is based on the observation that IG-use was associated with longer response times in Experiment 1, suggesting that it might require more cognitive effort than PTS.

In all three experiments, participants were asked to make interventions on simple causal systems to learn their underlying structure. Participants were presented with a large number of problems that were chosen so that IG and PTS made different predictions on many of them, allowing us to distinguish the two models.

To avoid stacking the decks in favor of a simpler strategy (like PTS), our experimental protocol kept the complexity of computing IG at a minimum. Three aspect of our task distinguish our paradigm from previous work. First, we asked participants to consider only *two* causal hypotheses on each trial. This is the simplest case of structure comparison which allows us to model interventions without making many additional assumptions about participants' hypothesis space. Second, we used causal structures without any background causes. This drastically reduces the number of possible outcomes that could result from an intervention and so facilitates the simulation of outcomes implied by IG. Finally, each trial provided participants with a bonus payment that decreased with the number of interventions made on that trial. The intent of this bonus structure was to minimize random exploration and encourage a search for diagnostic interventions from the start.

1.5. Modeling strategy use

To analyze learners' interventions, we adopt a hierarchical Bayesian modeling approach. Hierarchical Bayesian models offer many advantages for analyzing data. For instance, they allow us to not only estimate participants' individual parameter values, like the τ parameter in Eq. (5), but also population-level distributions of these parameters, so-called *hyperparameters*. As a consequence, hierarchical Bayesian models can capture individual differences while still allowing inferences about the group level (Lee, 2011). An advantage of using *Bayesian* inference is that it allows one to assess the credibility of a wide range of possible parameter values, rather than dealing only with point estimates (Kruschke, 2010). We exploit the advantages of hierarchical Bayesian modeling to identify which strategy or strategies describes people's behavior best, and what factors influence strategy use.

First, to test how well each strategy fits the data, the hierarchical approach lets us conduct posterior predictive checks (Gelman, Meng, & Stern, 1996) of the individual strategies (IG and PTS), as well as a mixture of the two. This method uses the population-level hyperparameters from a fitted model to generate choice data of groups of simulated participants with the same sample size (N) as the experiment. These simulated data sets provide a range of behavioral patterns that could have plausibly been observed if people were actually choosing interventions in line with the fitted model. A comparison of the simulated and empirical data can reveal whether a model provides a plausible account of people's behavior.

Second, the hierarchical model allows us to compare hyperparameters across experiments and experimental manipulations. Because we will model behavior as a mixture of PTS and IG, we can, for example, compare the population-level mixture weight in different experimental conditions to check for differences in strategy use (this method bears some similarity to an ANOVA within a frequentist null-hypothesis testing framework, see Kruschke, 2013).

Note that despite adopting a fully Bayesian analysis approach, we sometimes show maximum-likelihood (ML) estimates of parameter values rather than estimates from the Bayesian model. We do so because the latter are generally subject to *shrinkage* imposed by the hyperparameters in the hierarchical model, that is, they tend to become less extreme compared to ML estimates. While this can be a desirable property of hierarchical models, it can lead to an overcorrection of extreme values, especially when data are sparse (Scheibehenne & Pachur, 2013).

2. Experiment 1

Experiment 1 represents an exploratory study of learning via causal interventions. In addition to comparing model predictions, we analyze different ad-hoc dependent measures with the aim of understanding participants' decision strategies. These preliminary analyses set the stage for the experimental studies reported in Experiments 2 and 3.

2.1. Method

2.1.1. Participants

We recruited 105 participants (50 women and 55 men) aged 18–64 ($M = 34.3$ years, $SD = 12$) via Amazon Mechanical Turk (AMT). Participation was restricted to AMT users who reported living in the US. Participants were paid $2 for participation with the option of earning up to another $1 bonus based on their performance in the task. All participants were assigned to a single experimental condition.

2.1.2. Stimuli and materials

On each trial, participants were presented with a simple causal system and asked to learn how it worked. The systems were described as computer chips with three components (nodes), which could either be on or off as indicated by their color (green[2] or red, respectively). Within a given chip some components were causally connected so that the value of one node could influence the state of another node. The locations of the nodes of each new chip were randomly distributed into five possible positions to control for the influence of spatial position on participants' intervention choices.

Two possibilities or *hypotheses* of how the chip worked were shown as diagrams above the chip. These diagrams were spatially congruent with the placement of components on the chip, and causal links were indicated with arrows (see Fig. 2). On each trial, one of the two hypotheses was randomly selected to be the true underlying structure of the test chip. Participants interacted with the chip to determine which hypothesis described its operation (see below).

All possible three-node structures with one or two links were tested, yielding four basic structure types: Common-cause structures (e.g. A affects both B and C), common-effect structures (e.g. both A and B can affect C), chain structures (e.g. A affects B and B affects C), and one-link structures (A affects B, C is independent). For common-effect structures, we used the *noisy-or* integration function, which states that multiple causes are independently sufficient to make the common effect occur (Cheng, 1997).

The four structure types were paired with one another to yield 27 structure comparisons, which made up the two hypotheses for a given trial (see Fig. 3). Participants completed all 27 comparisons once in a pseudo-random order (see below). All links had causal strengths of 0.8, such that there was

[2] For interpretation of color in Fig. 2, the reader is referred to the web version of this article.

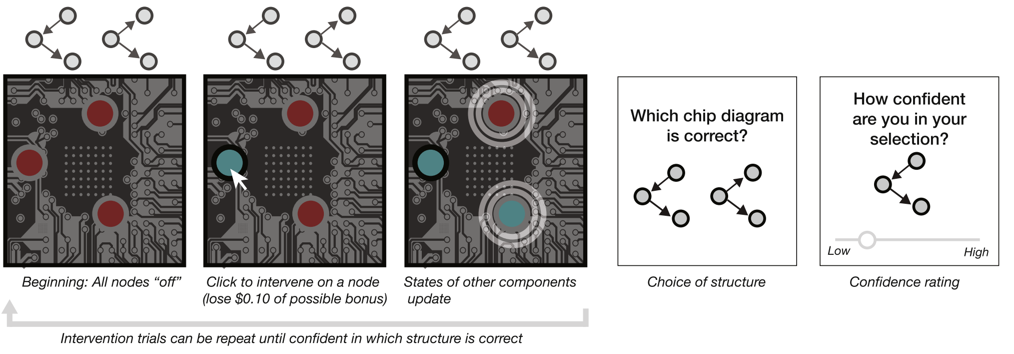

Fig. 2. Intervention phase of Experiment 1 which was repeated for each of the 27 problem types. The true underlying causal graph was selected randomly. Participants could make as many interventions as they wished, but lost $0.10 of a potential bonus payment with each intervention.

an 8/10 chance that an active parent node would turn on its direct descendants. No background causes existed that could turn on nodes spontaneously (i.e., nodes could only be activated through an intervention or through an active parent node). Some of the comparisons strongly distinguished IG from the PTS strategy (highlighted in gray in Fig. 3), while for others the models made similar predictions (see example in Fig. 1).

To test for learning effects over the course of the experiment, the 27 comparisons were divided into three groups of nine, so that behavior for each of the groups could be compared when a group was presented early (trials 1–9), in the middle (trials 10–18), or late (trials 19–27) in the experiment session. The order of the groups was counterbalanced between participants and the order of the comparisons within each group was randomized for each participant.

2.1.3. Expected strategy performance

To get a sense of the performance that can be expected in this task, Table 1 displays the average posterior of the most likely graph after making a single intervention with IG, PTS, or a random strategy. Note the possible range of values is .5 (complete uncertainty) to 1 (complete certainty). Assuming a learner makes just this one intervention, higher values would also correspond to higher accuracy in choosing the underlying graph. When interventions with the highest IG or PTS value (arg max) are chosen, the table shows that IG leads to higher average posterior beliefs. The same holds for probabilistic decisions with low decision noise, τ (from Eq. (5)). This is not surprising, given that IG aims

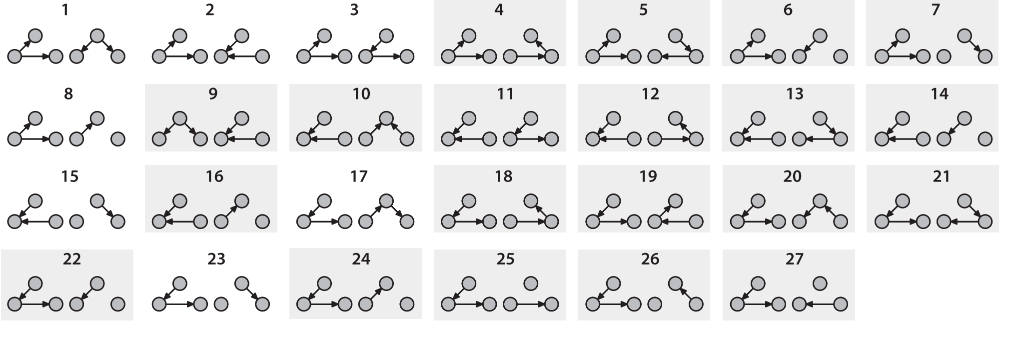

Fig. 3. All 27 problem types used in the experiment. Circles represent nodes in the causal graph (i.e., elements of the circuit board) and arrows indicate causal links (i.e., an arrow pointing from one node to another indicates that the first node can cause the other node to activate). Each numbered comparison represents a problem for which people were asked to decide between two causal hypotheses. Comparisons highlighted with a gray box in-principle distinguish the IG and PTS models.

to discriminate between hypotheses whereas PTS does not. Nevertheless, PTS and even a random strategy yield above chance performance (average posterior beliefs of .8 or higher).

2.2. Procedure

2.2.1. Instructions and training

Participants were asked to imagine working in a computer chip factory. Their job was to help test and identify a range of computer chips that were mixed up during an accident in the factory. Each chip was described as coming from one of two possible areas of the factory, corresponding to two possible structure hypotheses.

Participants received detailed instructions on the task, including the causal strengths of the links, the noisy-or integration function in common effect structures, and the rules for obtaining the bonus payment. They also performed a minimum of ten interventions on a simple two-node chip with one causal link to get some experience with making interventions and experiencing the probabilistic nature of the links. After the instructions and practice phase, they had to pass a short quiz before being allowed to proceed to the main task. If they failed the quiz, they had to view all instructions again, and re-take the quiz until they passed.

2.2.2. Intervention phase

See Fig. 2 for a visualization of the intervention phase. A short video of the intervention phase can be found here: http://gureckislab.org/annacoenen/videos/ChipTask_Exp1.mp4. Participants tested 27 chips corresponding to the 27 comparisons in Fig. 3. They were told that each chip was of one of two types, which were presented using arrow diagrams. These diagrams remained at the top of the screen the entire trial. To test a chip, participants could make as many interventions as they wished, but had to make at least one.

A trial began with all components of the chip switched off (red). Participants could then intervene on one component by clicking on it and thereby turning it on (green). When a component was activated, a black circle appeared around it to indicate the intervention. After a short (500 ms) delay an animated white ring appeared around *all* other components to indicate that they were updated as a consequence of the intervention. Components that were turned on by the intervention changed their color to green, all other components remained red. All components had to be reset to their original state (off) using a button press before another intervention could be made.

When participants felt that they had identified the chip type, they proceeded by clicking a button. The two arrow diagrams were then magnified on the screen, and participants indicated their choice by clicking on one of them. Participants rated their confidence about this decision using a continuous slider that ranged from "not at all confident" to "very confident". They received feedback on whether they chose the correct structure.

2.2.3. Incentive structure

To ensure that participants chose interventions carefully, they were told they could win a bonus of up to $1 from one randomly chosen trial at the end of the experiment. The bonus was only paid if they chose the correct structure on that trial, and it was further reduced by $0.10 for every additional intervention they made after the first one (which was obligatory). Thus, participants were incentivized to respond accurately while using only a small number of interventions.

2.3. Results

2.3.1. Graph choices

Participants were highly accurate in identifying the causal graphs. The percentage of correct choices averaged across individuals was 87% ($SD = 0.14$, $MD = 92\%$). This accuracy was achieved with only a small number of interventions. On average, participants made 1.56 ($SD = 0.59$) interventions during a single chip test, and on most trials stopped after only one intervention. The relative simplicity of the task (only two hypotheses and no background causes) and cost structure (that rewarded fewer interventions) probably contributed to this efficiency.

Participants' confidence ratings mirrored their choices, with higher confidence ratings on correct ($M = 80.22$), versus incorrect trials ($M = 72.62$), $t(89) = 3.66$, $p < .001$.

2.3.2. Learning with experience

To assess whether participants' accuracy improved over trials and with more feedback, we compared their accuracy on each of the three groups of nine comparisons and for each group position. Average accuracy on the first, second, and third group was 0.84, 0.87, and 0.89, respectively. In a logistic regression on accuracy with predictors for each participant, group, and group position, the effect of group position was significant, $\chi^2(2) = 11.7$ $p < 0.05$. However, overall the magnitude of improvement was relatively small.

2.3.3. Interventions

Our main question was how people decide which node(s) to intervene on given a pair of hypotheses. Due to the low total number of interventions (the modal number was one), we focus our analysis on the *first* intervention. This also avoids the complexities involved in accounting for how learners' interventions are affected by the outcome of earlier interventions.

Our first analysis considered choice patterns for each of the 27 problem types, that is, the proportion of individuals who chose each node on their first (and often only) intervention. Our aim was to present not only frequencies but also the variability in those frequencies across subjects, in order to characterize other choice distributions that might plausibly be observed if the same experiment were run again. Because the choice data are multinomial, there does not exist an easily interpretable measure of variability, such as a standard deviation for normally distributed data. We therefore used a bootstrapping method to generate other plausible choice patterns. For each of the 27 problem types, we repeatedly re-sampled (with replacement) participants' choices. These bootstrapped samples give an estimate of the range of other outcomes that we might expect if the experiment was repeated. The distribution of the samples is presented in a simplex plot for each problem. An example is shown as the blue cloud of dots in Fig. 4 (bottom left panel) and in subsequent figures. The white dot in the middle of the blue samples corresponds to the actual proportions with which any given node was chosen by participants in the experiment for the given problem. In this example the majority of participants selected n1, with the remaining individuals split roughly evenly between n2 and n3. Fig. 5 shows the bootstrapped choice data for all problem types.

In the following sections, we compare these empirical data to the predictions of IG, PTS, and a combination of the two, using a method of posterior predictive model assessment. Table 2, which will be revisited below, additionally shows more standard measures of model assessment, such as log-likelihood and BIC, along with the best fitting estimates of model parameters.

2.3.4. Information gain

To evaluate IG's ability to account for the data, we derived its predictions for each of the 27 problems according to Eq. (3). We assumed that participants had a uniform prior belief over the two structures before making an intervention (consistent with the instructions). When fitting the model, it will appear to account for some problems better than others. To ensure that model-fit discrepancies cannot be explained by chance variation, we fit IG using a hierarchical Bayesian model and then used a method of posterior-predictive model assessment (Gelman et al., 1996), as mentioned above.

The model includes a single free parameter τ in the softmax decision rule (see Eq. (5)), which increases with the degree to which behavior resembles guessing rather than choosing the option with the highest expected IG. Each participant's τ value was sampled from a population-level gamma distribution with two parameters, α (shape) and β (rate). A gamma distribution was chosen because τ can range from 0 to infinity. The values of α and β as well as each participants' τ value were fitted using the hierarchical Bayesian model summarized in Fig. 16 in Appendix A.

To obtain predictions from this model we then repeatedly sampled from the posterior-predictive distribution of α and β to simulate different populations of subjects. For each population, we then generated 105 (N in the experiment) values of τ and used these to generate choice data for each problem. These simulated choice distributions are shown alongside the bootstrapped empirical data as red dots

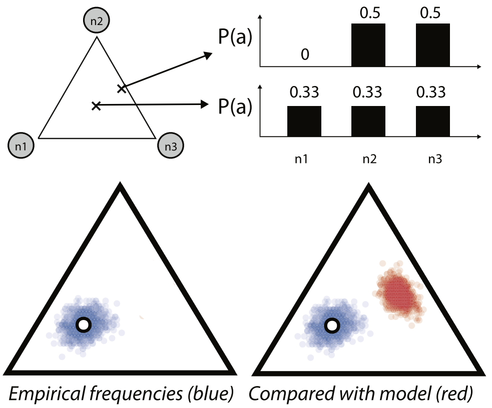

Fig. 4. How to interpret simplex plots that appear later in the paper. *Top*: The corners of the space represent each of the three possible intervention choices on any trial. Points within the simplex correspond to the probabilities, $P(a)$, of intervening on each of the three nodes in a causal graph averaged over participants. Any possible decision preference across the three nodes can be summarized as a unique point in this space. The middle of the simplex is the point of indifference between the three choice options. A point closer to any corner of the simplex represents a strong preference for a particular node. Two examples are plotted in the space corresponding to the middle point (indifference between the three options) and indifference between only n2 and n3. *Bottom*: The left panel shows an example of bootstrapped empirical data. The white dot in the center of the cloud represents the actual choice proportions made by participants in the experiment. The blue cloud of points (lighter color when viewed in grayscale) are bootstrapped samples showing the uncertainty in this empirical estimate (see text). The red cloud of points (darker color in grayscale) in the right panel represent a range of model predictions (also explained in the text). If the model predictions (red cloud) overlap with the blue cloud (empirical data) the model provides a credible account of the empirical data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

in Fig. 5. In sum, each red dot represents the choice proportions for a simulated sample of learners (drawn from a simulated population), of the same size as our experiment population. Rather than a single point, the model predictions reflect a distribution of possible outcomes given plausible assumptions about how parameters might vary in the population. One major advantage of visualizing the model predictions in this manner is that it allows us to assess the fit of a single model and to decide whether it offers a credible fit to the data. Comparing the distribution of bootstrapped empirical data (blue dots) and model simulations (red dots) invites a comparison of the two through visual inspection. When the red dots mostly overlap with the blue dots, the model provides a credible account of people's choices. This comparison is analogous to comparing and finding the overlap of two confidence intervals of normally distributed data, while staying true to the multinomial data of this experiment.

As can be seen from Fig. 5, there is a large overlap of the model and data in a number of problem types (such as Problems 1, 3, 6, 10 and 11). However, there also exist multiple problems for which the model and the data differ considerably (like Problems 4, 12, 14, 16, 21 and 27). This analysis shows that the IG model does not fully explain how participants chose their causal interventions in this task, even after taking into account the variability of the data and a range of possible model predictions.

To examine whether people's propensity to use IG changed with experience, we also compared the log likelihood of the IG model fit to each participant during the early, middle, and late comparisons phases of the experiment. Using a repeated measures ANOVA, we found no significant main effect of time on the log likelihood as a function of block, $F(2,2726) = 2.207$, $p > 0.11$. The effect did approach significance, however, so there may exist a trend towards higher IG use over time.

2.3.5. Other discriminatory strategies

Although we explored a range of possible modifications to the IG strategy as well as one alternative discriminatory model (probability gain; Baron, 1985; Nelson, 2005, for details see Appendix B), our results suggest that *any* discriminatory strategy would have difficulty explaining people's choices for certain problem types in this experiment. The reason is that participants frequently made

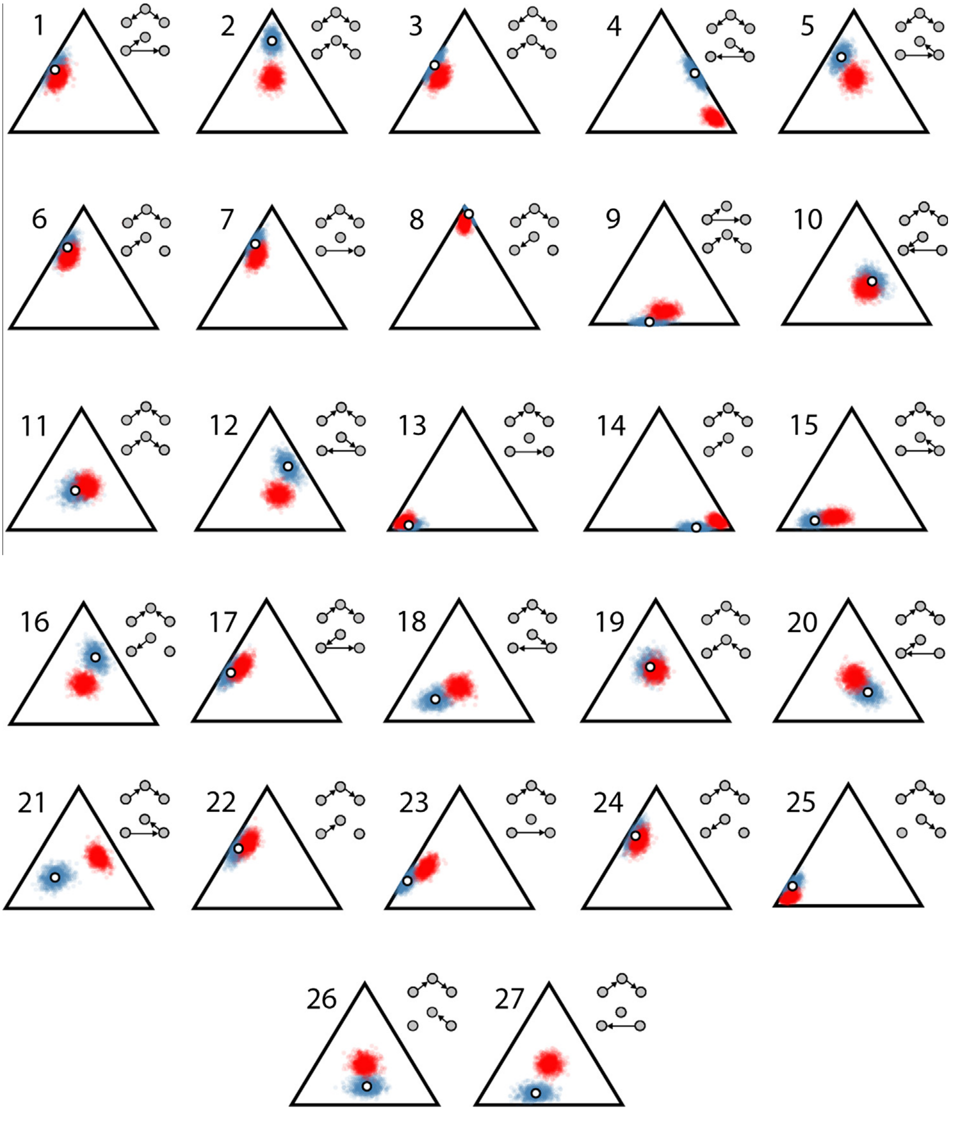

Fig. 5. Intervention choices and predictions of the IG model by problem type. The corners of each triangle correspond to nodes in the causal graph that participants intervened on (see Fig. 4). White dots indicate the actual choice frequencies. Bootstrapped samples of these choices are shown in blue (lighter color when viewed in grayscale). Samples from the IG model's posterior are shown in red (darker in grayscale). Where the two point clouds do not overlap (e.g. in Problems 4, 21, or 27), it is very unlikely that the observed data could have been generated by participants adhering to IG. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

interventions that tend to lead to the *same* outcome under both hypotheses, which thus failed to discriminate between them. Consider for example Problem 21, shown in the left panel of Fig. 6: IG predicts that learners should avoid intervening on n1, which is the root of both chain graphs, because it will probably lead to the same outcome (all nodes ON), irrespective of which hypothesis is true. Yet, most participants chose to intervene on n1. Problem 14 (right panel in Fig. 6) provides another example: In this problem, intervening on n1 will *always* lead to the same outcome for both hypotheses (it has zero EIG), yet participants intervened on this node frequently. Although less pronounced than in

Table 2
Overview of the models fit with maximum-likelihood estimation using the multinomial likelihood of participants' choice distribution on each of the 27 problems. To calculate BIC, N was the number of interventions (27) times the number of participants (105). The combined model far outperforms the two individual models, even taking into account the additional parameter, θ.

Model	Log likelihood	BIC	R^2	τ	θ
Random	−880	1760	–	–	
IG	−398	804	.66	.37	–
PTS	−375	758	.61	.30	–
Combined	−221	446	.86	.22	.40

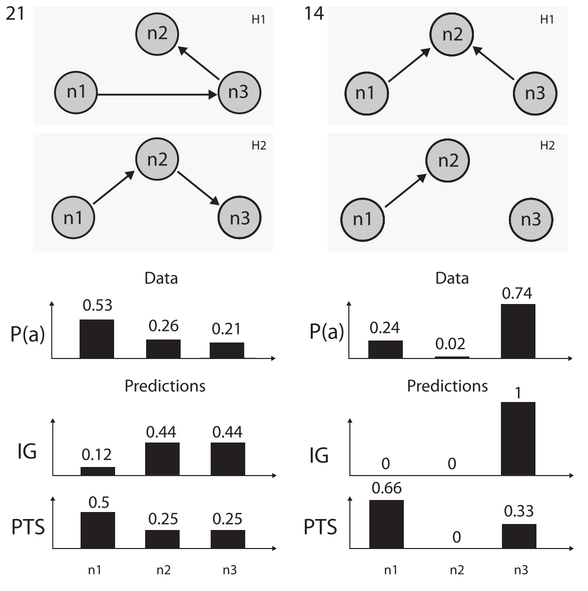

Fig. 6. Example Problems 21 (left panel) and 14 (right panel): $P(a)$ corresponds to participants' average choice probability for each node. Model predictions are normalized to sum to 1. In both problem types, participants deviated from the IG strategy with a higher than predicted preference for the root node `n1` despite the fact that this intervention will likely lead to a non-discriminating outcome. This divergence from IG is more pronounced in Problem 21 (left), but perhaps more surprising in Problem 14, because intervening on `n1` has no informational value whatsoever, yet 24% of participants chose this intervention.

Problem 21, this finding is unlikely to have arisen by chance from a population of IG users, as indicated by the barely overlapping model and empirical distributions for Problem 14 shown in Fig. 5.

In summary, with an experiment designed to make it particularly easy to use a discriminatory strategy like IG (at least relative to past work), participants' choices deviated from the IG model on a number of problem types.

2.3.6. Positive testing strategy

We next evaluated if participants' intervention decisions followed a PTS strategy. We calculated the predictions of PTS using Eq. (6). Fig. 7 shows bootstrapped samples (blue dots) from participants' actual choices (white dot) compared to samples from the posterior of the PTS model (red dots). Predictive samples were obtained in the same way as for IG.

PTS also fits the data reasonably well on some problems, but diverges on others. Crucially, the model fit the data better on many problems which were poorly predicted by IG, for example Problems 21, 26 and 27. To illustrate this point more generally, Fig. 8 shows the relationship between the log likelihood of the empirical data given the IG model for each problem and the rank correlation of the preference order of IG and PTS over the three nodes (i.e. interventions) on each problem type. It

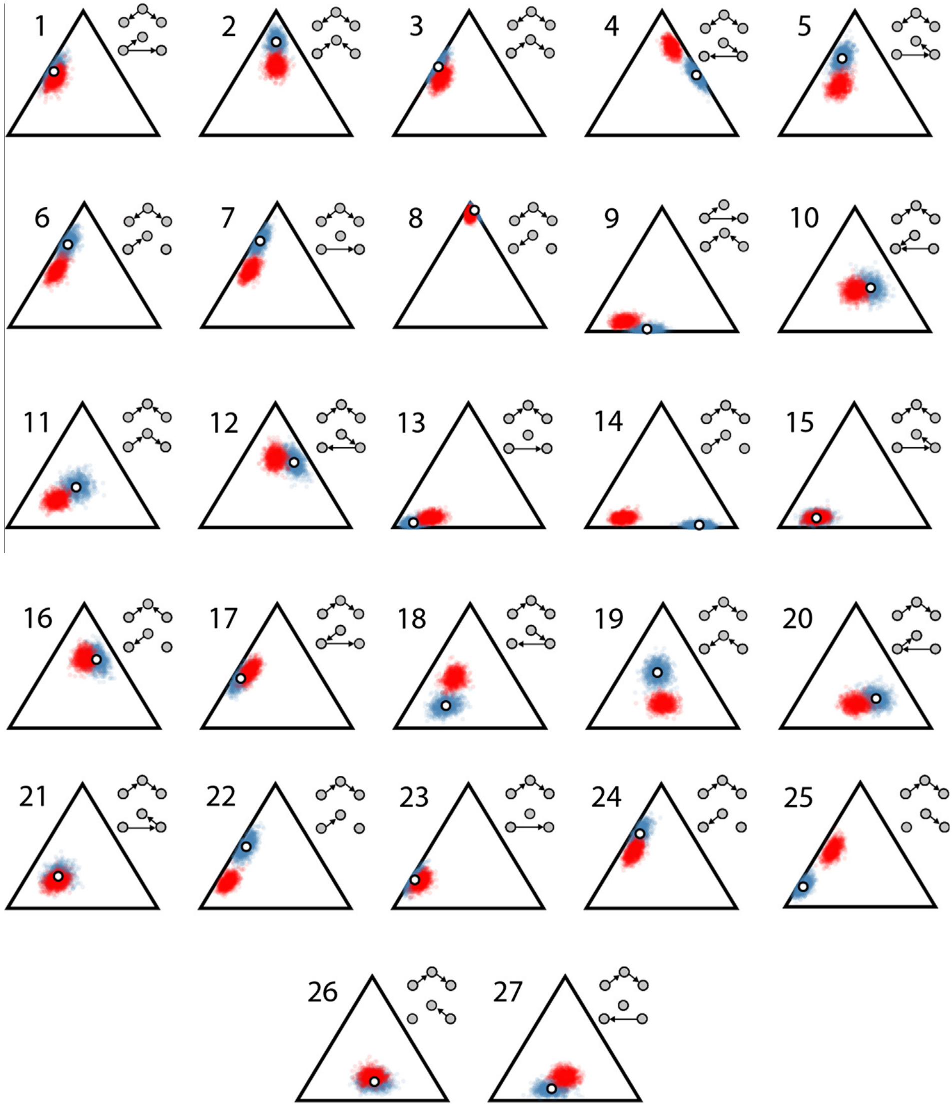

Fig. 7. Intervention choices and predictions of the PTS model, by problem type. Again, a lack of overlap of the bootstrapped data (blue) and model predictions (red), indicate that the model is highly unlikely to have generated the data, for example in Problems 14, 22 and 25. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

shows that when the models make very different predictions (low correlation), IG does not fit the data well, $r(25) = .68$, $p < 0.001$. This pattern suggests that deviations from IG on some problem types are not just due to random variation in the data. Instead, each model particularly suffers on problems where the other models make very different predictions.

2.3.7. Other non-discriminatory strategies

In addition to the above PTS model, we fit a range of alternative non-discriminatory models to the data. For example, we considered if participants' choices could be explained by a desire to merely turn

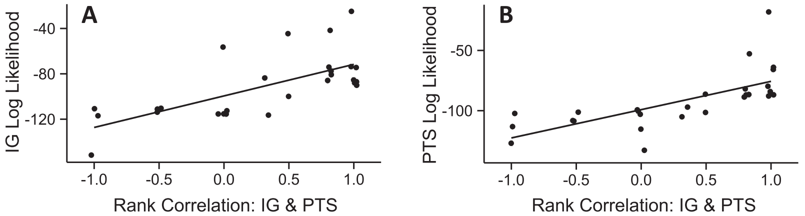

Fig. 8. Relationship between the goodness-of-fit of IG (A) and PTS (B) and agreement of the two models (Kendall's τ rank correlation), by problem type. Both models fit better when their predictions are in line with each other (high rank correlation). When the two models disagree, they fit less well, indicating that participants' divergence from each model is not purely due to random noise, but that the two strategies might, in fact, be complementary to each other.

on as many nodes as possible. No alternative we tested explained the data better than PTS and none of them complemented the predictions of IG better than PTS. For more details on these alternative models, see Appendix B.

2.3.8. A combined model of intervention choice

To test the possibility that participants are guided by both types of reasoning, we fit a third model which represents a linear combination of IG and PTS with a mixture weight θ for the two strategies. To compute the likelihood of an intervention in this model, the softmax rule in Eq. (5) takes weighted sums of the scores from IG and PTS, where the mixture weight θ can take values between 0 and 1. When $\theta = 1$ the strategy reduces to IG, when $\theta = 0$, it reduces to PTS. Both θ and τ were fit individually for each subject, as shown in the Bayesian hierarchical model in Fig. 9. The mixture weight θ was fit using a Beta distribution that was reparameterized by its mean μ and standard deviation κ. Fig. 10 shows posterior samples of this model compared to the data in the same way as for the two individual models. The combined model shows a much larger overlap of model predictions with the data on most problem types compared to the individual models. In fact, none of the empirical data from individual problems appears implausible in relation to the model (i.e., the red and blue distributions largely overlap) although, to be fair there is less overlap in Problems 18, 19, and 21.

$$\tau_i \sim Gamma(\alpha, \beta)$$
$$\theta_i \sim Beta(\mu\kappa, (1-\mu)\kappa)$$
$$p_{ij} = \frac{exp((\theta_i EIG_j + (1-\theta_i)PTS_j)/\tau_i)}{\sum exp((\theta_i EIG_j + (1-\theta_i)PTS_j)/\tau_i))}$$
$$y_{ij} \sim Categorical(p_{ij})$$

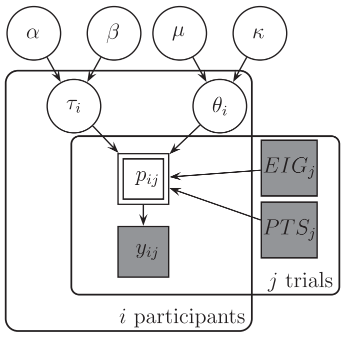

Fig. 9. Hierarchical Bayesian model of the combination of IG and PTS. Each trial, j, corresponds to one problem type for which each participant chose one intervention, y. IG_j And PTS_j are three-vectors with model scores for the three possible intervention on problem j. p_{ij} is a three-vector of choice probabilities for each intervention. τ_i and θ_i are fit for each participant and capture the noisiness of their choices (higher τ indicates behavior closer to guessing), and the strategy weight (high θ indicates behavior in line with IG), respectively. Hyperparameters, α, β, μ, and κ capture the population-level distributions of these parameters.

To ensure that a similar conclusion could have been derived without the use of a complex hierarchical modeling framework, we also compared the fit of all three models using maximum-likelihood estimation, without fitting individual participants' parameter values (that is, using a single τ parameter in the two individual models, and one τ and θ parameter in the combined model). Table 2 shows the log likelihoods of the IG, PTS, and combined models (with one τ parameter in the individual models, and τ and θ in the combined), along with the best-fitting parameter values. Again, the combined model leads to a large improvement of the model fit compared to the individual models. This is also implicated by the best-fitting value of τ, which is lower in the combined model, showing that participants' choices were closer aligned with the model.

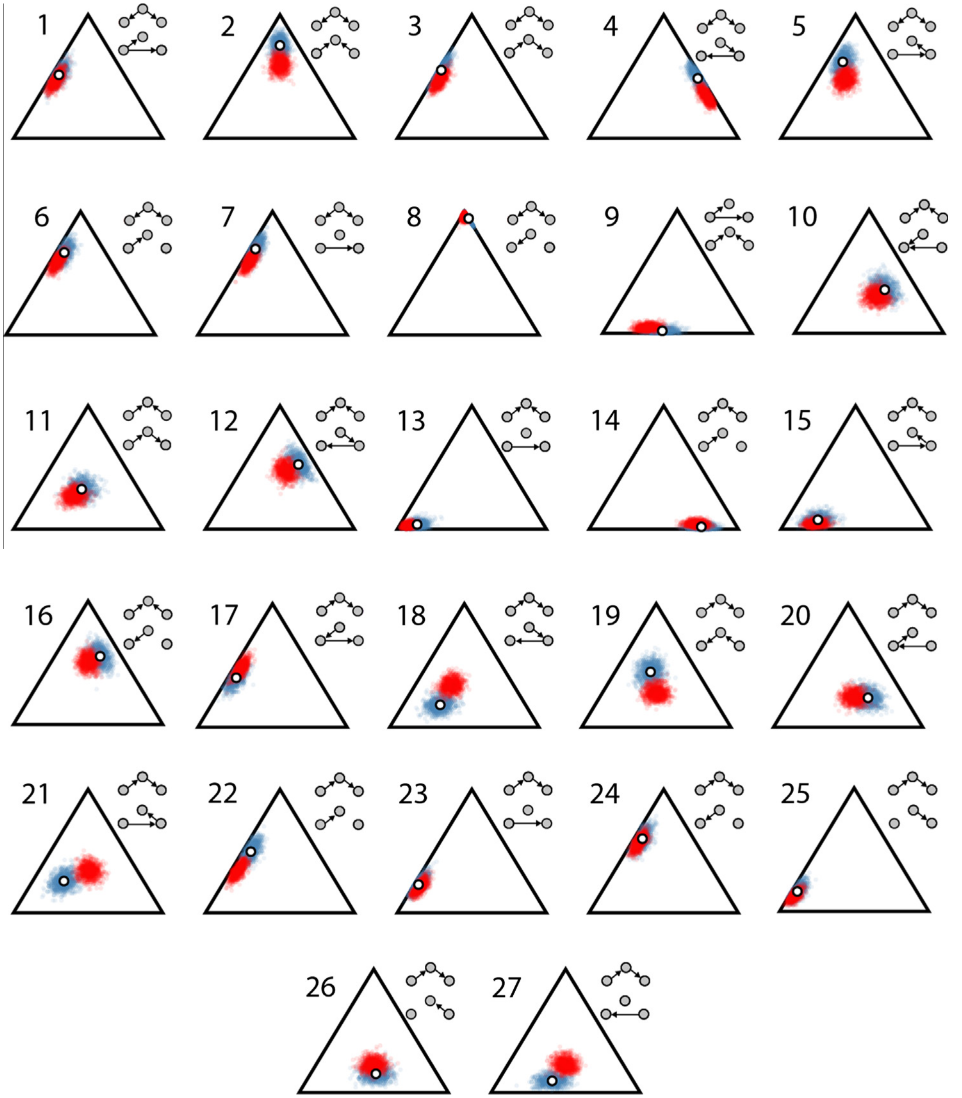

Fig. 10. Intervention choices and predictions of a combined model of IG and PTS, by problem type. Although there are some borderline cases (e.g. Problems 18 and 27), the overlap of the model predictions (red) and the bootstrapped data (blue) has increased compared to the individual models. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Based on the clear superiority of the combined model, in the following analyses we conduct additional tests of the hypothesis that people are driven by both discriminatory and confirmatory aspects of intervention selection. We examine whether the agreement or conflict between the strategies affects how long it took to plan an intervention and how participants differed in the degree to which they used one strategy compared to the other.

2.3.9. Response time

If people are influenced by both strategies, then cases where the two strategies disagree become particularly interesting, because learners potentially had to resolve the conflict before making a decision. We tested the possibility that interventions took longer when IG and PTS make different predictions. For each problem type, we calculated the agreement of the two models using the Kendall τ rank correlation coefficient of their prediction for each node. The relationship between model agreement and the median response time before making an intervention in each problem is shown in Fig. 11. As predicted, the correlation was negative, $r(25) = -0.58$, $p < 0.005$, such that participants responded more quickly when the two models agreed and more slowly when they disagreed.

Using linear regression, we also controlled for two other nuisance variables that could have mediated this relationship between model agreement and RT. First, since the problem types differed in complexity (graphs could have either one or two links), we included the total number of links of both hypothesized graphs in the regression analysis. Second, participants' general uncertainty about which intervention to choose, irrespective of whether this uncertainty arises from strategy conflict, might have increased RTs. We therefore also included the Shannon entropy over the preference scores for the three nodes in each problem, as predicted by the combined model (using the best-fitting parameter values of θ and τ at the population level). Even controlling for these variables, the negative relationship between RT and model agreement survived, $t(23) = -3.465$, $p < 0.05$. Additionally, the total number of links increased response time, $t(23) = 3.623$, $p < 0.01$. Uncertainty was not a significant predictor, $t(23) = -1.559$, $p > 0.10$.

In summary, we find evidence that strategy conflict is correlated with response time, which provides indirect evidence that participants had to resolve mental conflict between the two strategies.

2.3.10. Individual variability

The θ parameter in the combined model provides an estimate of participants' preferred strategy ($\theta = 0$ means perfect PTS, $\theta = 1$ means perfect IG). Fig. 12A shows a histogram of the best-fitting values of θ for each participant based on maximum-likelihood estimation. Interestingly, rather than dividing into two groups, many participants fall on a continuum between the two strategies. Thus, behavior does not only resemble a strategy mixture in the aggregate; it does so at the individual level, as well.

To get a sense of how closely the strategy types matched individual behavior, Fig. 12B shows the relationship between the maximum likelihood values of θ and τ for each participant. Recall that high

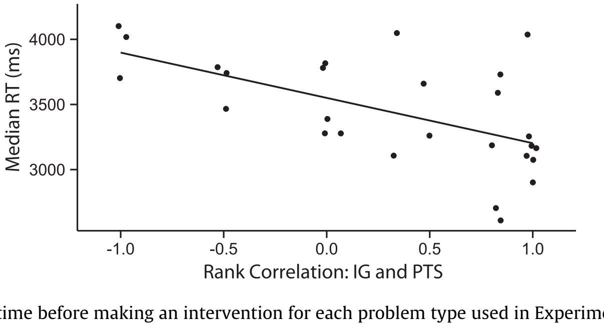

Fig. 11. Median response time before making an intervention for each problem type used in Experiment 1, by agreement of IG and PTS (Kendall's τ rank correlation). When the models make different predictions (low rank correlation), RTs were longer, indicating that those problems made it more difficult for participants to choose an intervention.

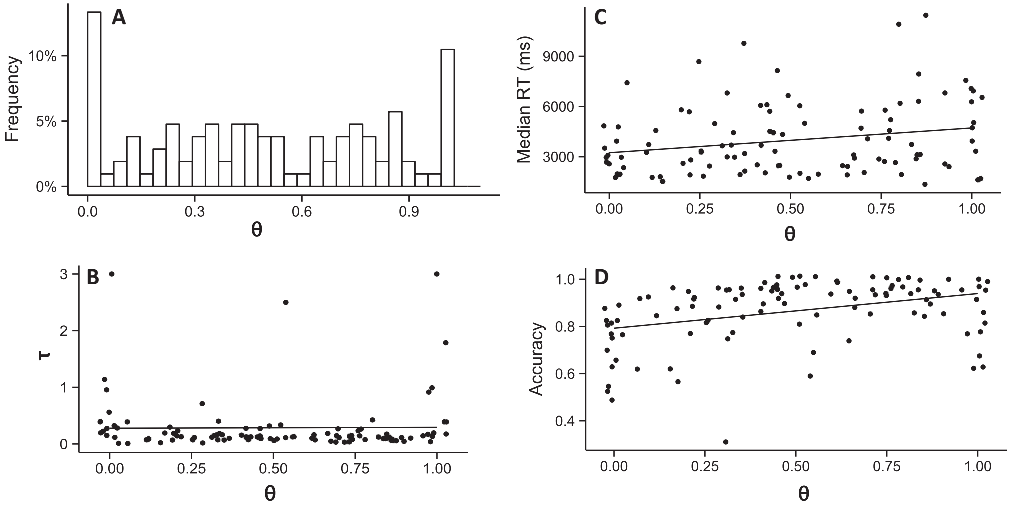

Fig. 12. *Panel A*: Histogram of maximum likelihood estimates of θ in Experiment 1. When $\theta = 1$, a participant's strategy is best fit by a pure IG model, when $\theta = 0$, it is best captured by PTS alone. *Panel B*: Relationship of the ML estimates of θ and τ by participant. Note that τ values greater than 3 were set equal to 3 for the purpose of this plot. This setting already starts resembling a pure guessing strategy. For example, for values [.1,.9,0.], the softmax function using $\tau = 3$ yields choice probabilities [.31,.40,.29]. *Panels C and D*: Relationship between best-fitting θ of each participant and their response time and accuracy.

levels of τ reflect increased guessing. There is no significant trade-off between the two variables, $r(103) = -0.07$, $p > 0.05$, suggesting that particular fitted values of θ do not reflect increased random choosing.

Next, we asked whether and how a participant's intervention strategy, measured by θ, was related to other variables. Specifically we wanted to know if it correlated with response time and overall accuracy.

We expected that the IG strategy, which requires comparison between hypotheses and simulation of outcomes from each intervention, would be associated with longer RTs than PTS, which only requires finding nodes with a large proportion of children in either hypothesis. Fig. 12C shows the relationship between θ and median RT per participant. As expected, participants whose behavior was better accounted for by IG took significantly longer to choose interventions, $r(103) = .23$, $p < 0.05$.

We also considered the relationship between strategy use and accuracy. From model simulations (see Table 1), we expect learners choosing interventions with IG to be more likely to choose the correct structure for reasonably low values of τ. Fig. 12D shows the relationship between θ and mean accuracy, which is positive as expected. We also expected the degree to which a participant behaved randomly, measured by τ, to have a negative impact on the quality of information and thus accuracy. We therefore calculated the correlation between θ and residual accuracy after accounting for each participant's value of τ, which not surprisingly remains positive, $r(102) = 0.44$, $p < 0.001$.

2.4. Discussion

Three important observations emerged from our first study. First, the results show that neither the IG model nor PTS alone could predict causal intervention decisions across a range of structure learning problems. This finding is at odds with single-strategy theories that have so far dominated the literature on causal intervention learning.

Second, a linear combination of the two strategies greatly improved the model fit, suggesting that both discriminatory and confirmatory aspects can play into people's interventions. The success of the combined model also tallies with recent work on information search that has demonstrated mixtures of strategy use, albeit in non-causal environments (Markant & Gureckis, 2012b). Although use of a confirmatory strategy may be unsurprising in light of the literature on rule learning (e.g., Klayman & Ha, 1987), this article is the first to define what positive testing of causal rules consists of and establish that learners in fact conduct such tests.

Third, Experiment 1 found large individual variability in how well learners were described by the two models and, moreover, that most individuals were influenced by both strategies. That is, rather than distinct groups of positive testers and discriminatory learners most participants behaved as if they used a mixture of both. One open question that remained from this first experiment is what guides the degree to which participants use one strategy or another. One possibility is that the tendency to use PTS reflects a stable bias in how people approach such tasks that may only be overridden with extensive instruction.

However, it is worth noting that in Experiment 1, even if participants used PTS to choose interventions, they could still in principle learn the correct structure in most problem types with only few interventions. That is, using PTS only incurred a small cost in accuracy compared to IG (similar to a "flat-maximum" phenomenon in economic decisions where different strategies all perform relatively similarly). Accordingly, one alternative hypothesis is that learners only adopt a more effortful IG strategy when its benefits exceed the cost of one that is more cognitively frugal. One might speculate that everyday reasoning problems rarely necessitate the additional effort expended to engaging in discriminatory reasoning. However, when faced with those that do, people may adapt their strategies accordingly. Such a cost-benefit perspective on strategy selection is often adopted in other areas of the cognitive and decision sciences (Payne, Bettman, & Johnson, 1988) but has been less prominent in the literature on causal learning.

We now directly explore these issues by manipulating factors that we hypothesized might alter participants' intervention strategies either to favor IG (Experiment 2) or PTS (Experiment 3).

3. Experiment 2

Experiment 2 tests the hypothesis that people will adopt more discriminatory strategies when the cost of using a confirmatory strategy is greater than it was in Experiment 1. We essentially repeated the design from Experiment 1 with an additional between-subjects manipulation. Participants first completed a set of intervention problems that were designed to make PTS a lot less effective than IG (PTS− condition) or equally effective (PTS= condition). PTS can be made less effective by choosing problems for which it yields non-diagnostic outcomes. In other words, the PTS− condition presents an environment in which the benefits of using IG may outweigh the cognitive effort required to discriminate between hypotheses. In the PTS= condition both strategies result in diagnostic outcomes, so there is no clear incentive to choose IG over PTS.

Both groups of participants were then tested on a set of transfer problems taken from Experiment 1. If strategy use is a stable trait or bias (or if our cost manipulation is ineffective), no difference between the conditions should be found. But if strategy use is adaptive, transfer-phase interventions will be more in line with IG in the PTS− condition. Apart from demonstrating adaptation, the latter result would also provide corroborating evidence for our interpretation of Experiment 1 by showing that interventions we claim reflect IG and PTS respond to an experimental manipulation in the expected way.

3.1. Method

3.1.1. Participants

We recruited 122 participants via Amazon Mechanical Turk. Compensation and incentive structure were the same as in Experiment 1. Participants were randomly assigned to either the PTS− ($N = 62$) or PTS= ($N = 60$) condition.

3.1.2. Stimuli and materials

Participants completed a total of 40 problems. The first half of the experiment presented 20 problems consisting of pairs of four-node causal networks (see Fig. 13). In the *PTS−* condition, problems

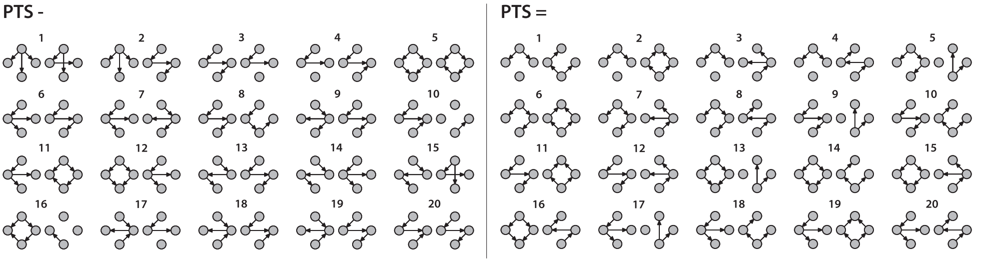

Fig. 13. New problem types used in the two conditions of Experiment 2. The problems in the PTS− group were designed to yield particularly non-informative outcomes from interventions chosen with PTS, that is, by intervening on root nodes the outcome would often be ambiguous regarding the true underlying graph. This is not the case for problems in the PTS= group, in which PTS interventions are highly informative.

were designed such that PTS interventions would often lead to outcomes that do not discriminate the hypotheses. A simulation of an optimal learner choosing interventions on these problems resulted in only 62% accuracy after one intervention using PTS compared to 91% using IG (assuming the learner always chooses the option with the highest IG/PTS score on each trial). In the *PTS*= condition, simulated accuracy of PTS after one intervention was 93%, compared to 95% with IG. To compare, in Experiment 1 PTS would have led to accuracy of 85% and IG to 92% (see also Table 1).

Participants were then tested on a subset of 20 problems used in Experiment 1. This *transfer set* included problems for which IG and PTS made different predictions, as judged by the rank order of predicted preferences for each node. They are highlighted with a gray box in Fig. 3. We only used a subset of problems from Experiment 1 to prevent the experiment getting too long without sacrificing the discriminability of the strategies.

3.1.3. Procedure

The procedure was the same as in Experiment 1, except that participants were first given the new four-node problems followed by the transfer problems, both in randomized order. There was no difference in the procedure between the two experimental conditions besides the set of novel problems and no indicator of the transition between phases other than the transition from four- to three-node problems. Another short video showing a problem with two four-node graphs can be found here: http://gureckislab.org/annacoenen/videos/ChipTask_Exp2.mp4.

3.2. Results

3.2.1. Accuracy

Accuracy on the 20 new four-node problems was lower for PTS− participants (80% accurate) than PTS= participants (92% accurate). This difference was unsurprising given that the expected accuracy from optimal learner simulations was higher on PTS= problems for both strategies. Additionally, participants who learned to use IG in the PTS− condition may have used PTS on early trials, leading to particularly low accuracy. Indeed there is some evidence (approaching significance) that accuracy increased with each trial in the PTS− condition ($R^2 = 0.18, F(1, 18) = 3.97, p = 0.061$), suggesting the presence of strategy learning. We do not find the same relationship in the PTS= condition ($R^2 = 0.06, F(1, 18) = 1.22, p = 0.283$).

On the 20 transfer set problems both groups achieved an average accuracy of 88% ($SD = 0.13$ and $SD = 0.12$ in PTS− and PTS=, respectively). That is, accuracy alone indicated no effect of the training manipulation. Recall, however, that both IG and PTS can lead to relatively high accuracy on these problems (91% and 85%, respectively; see Table 1). Thus, it is possible that the underlying strategy profiles in the two conditions were different even with similar performance levels. Also note that there are other drivers of accuracy, such as decision noise during interventions, the number of interventions

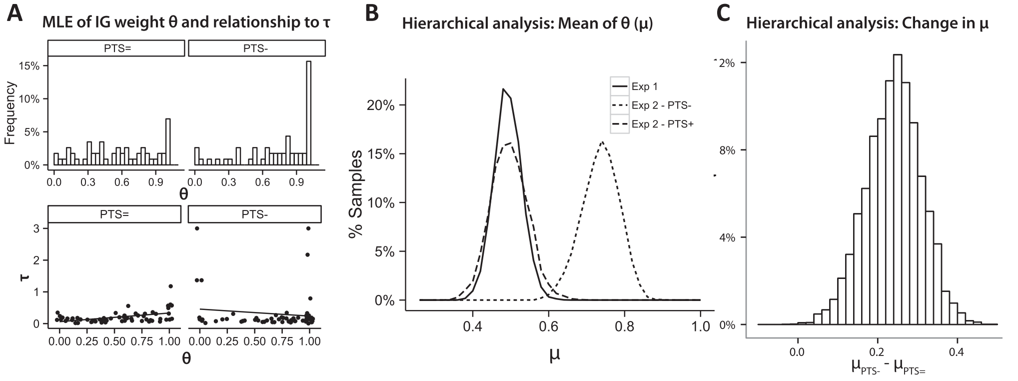

Fig. 14. (A) Top plot shows histograms of best-fitting θ parameters in both conditions of Experiment 2. High θ indicates a better match of the data to IG compared to PTS. Bottom plot shows the relationship between θ and τ. (B) Distribution of samples of the μ parameter (population mean of θ), fit to data in Experiment 1 and both conditions of Experiment 2. (C) Difference between samples of μ in PTS− and PTS= conditions of Experiment 2.

taken, and a participants' ability to correctly update their beliefs in light of an intervention outcome, which might have further diminished the possibility to detect an effect for the two groups.

3.2.2. Intervention strategy

To investigate whether the training manipulation affected participants' transfer strategy, we first determined the best-fitting mixture parameters θ for each participant using maximum-likelihood estimation. To make the two conditions comparable, we fit these parameters only to the 20 transfer problems. The distribution of best-fitting θ parameters in Experiment 2 is shown in Fig. 14A. In the PTS− condition θ values fall closer to IG (i.e., $\theta = 1.0$), compared to PTS=.

To assess whether, at the population level, this new distribution of strategy weights was different between the two conditions, and to compare the results to Experiment 1, we also fit the full hierarchical Bayesian model (see Fig. 9). Fig. 14B shows histograms of MCMC samples of the μ parameter from this model for the two new conditions and Experiment 1, calculated based on the transfer problems common to all three experimental conditions. These histograms approximate the posterior distribution of μ, which represents the population mean of θ. There are a number of advantages of using this population parameter,[3] rather than simply taking the mean of the individual maximum-likelihood values of θ and comparing them across manipulations. For example, the width of the posterior distribution (approximated by MCMC samples) indicates how confident we should be about a parameter value in the population.

Fig. 14B shows that μ is shifted considerably towards higher IG-use in the PTS− condition of Experiment 2, compared to PTS= and Experiment 1. Another way of testing whether this difference is credible involves determining the 95% Highest Density Interval (HDI) of the distribution of the *difference in* μ in the PTS− and PTS= conditions. To compute this difference, we took 10,000 samples from each model, paired the samples randomly, and computed $\mu_{PTS=} - \mu_{PTS-}$ (method is similar to Kruschke Kruschke, 2013). The resulting distribution is shown in Fig. 14C. Since the 95% HDI of this distribution of $\mu_{PTS=} - \mu_{PTS-}$ does not include 0, we can be confident to conclude that there is a credible difference at the population-level in the degree to which participants used IG in these two conditions.

[3] Besides those already mentioned, one advantage of fitting all parameters together in the same model is that the estimate of μ is automatically influenced more strongly by participants with low values of τ (i.e. those whose behavior did *not* resemble guessing). This is important because participants might have extreme values of θ when fit using MLE (close to 0 or 1), but still mostly choose randomly (high τ). Their data will have a smaller effect on the parameter estimates of θ and μ in the hierarchical model.

3.3. Discussion

This experiment found that participants were indeed more prone to behave in a discriminatory fashion after encountering problems for which PTS led to a lower expected payoff (PTS−) than ones for which it did not (PTS=). Importantly, this difference was found on the same set of transfer problems in both conditions. That learners experience with one problem set carried over to another suggests that strategy use is not a stable trait or bias but rather can adapt to different choice environments. As mentioned, the finding that IG/PTS use increased/decreased in the expected way to the training manipulation also corroborates our claim for the presence of those strategies in Experiment 1.

4. Experiment 3

Experiment 2 showed that people's intervention strategy can shift towards IG after experiencing a modest number of situations that disincentive positive testing. To further test the adaptive nature of intervention strategies we next consider a manipulation that might encourage the use of PTS and thereby lead to the opposite pattern.

Since we found that the tendency toward IG use was associated with longer response times (see Experiment 1), one factor that might guide strategy selection is the time available to decide to intervene. If the IG strategy, or discriminatory strategies in general, are more cognitively effortful they may take longer to compute, making PTS more attractive when time is limited or costly. In Experiment 3 we put participants under time pressure by adding a cost to the time taken to intervene.

We tested three groups in which the time to make an intervention was limited to 60, 8, or 4 s. Participants were incentivized to respond quickly since their potential bonus decreased as a linear function of time. We tested two short time windows (4 s and 8 s) to ensure that at least one of them would exert time pressure without inducing random guessing. We tested the longer time window (60 s) to test whether less time pressure would have a weaker effect on strategy use than the two extreme conditions.

If participants take their computational capacities into account in an adaptive fashion, and if PTS is easier to compute than IG, then we expected the shorter response deadlines to increase the use of PTS (i.e. result in overall lower estimates of θ and μ). On the other hand, if strategy selection is not adaptive, we expected that they would only increase the noisiness of choices (i.e. an increase in τ).

4.1. Method

4.1.1. Participants

We recruited 295 participants via Amazon Mechanical Turk.[4] Participants were randomly assigned to the three conditions ($N = 98$, $N = 98$, and $N = 99$ in the 4 s, 8 s, and 60 s conditions, respectively). They were paid the same pre-bonus amount as in Experiments 1 and 2 ($2), but the rules for the bonus payment differed, as explained below.

4.1.2. Stimuli and materials

Participants completed the same 27 problems used in Experiment 1. The screen was set up similar to previous experiments, but included an hourglass on the right hand side (replacing the yellow status bar used in Experiments 1 and 2, see Fig. 2).

4.1.3. Procedure

A short video of an intervention trial under time pressure (from the 8 s condition) can be found here: http://gureckislab.org/annacoenen/videos/ChipTask_Exp3.mp4. The procedure was the same as in Experiments 1 and 2, with a few exceptions.

[4] We decided to collect more data (~100 participants per condition) in this experiment compared to Experiment 2, because we expected the timing manipulation to potentially make the choice data noisier. Consequently, we leaned towards a similar sample size as in Experiment 1.

Table 3
RT percentiles in the three conditions of Experiment 3. Note that the difference between the 8 s and 60 s conditions is more pronounced than between the 4 s and 8 s conditions.

Condition	5th	25th	50th	75th	95th
4 s	1401	1772	2204	2866	5141
8 s	1510	1974	2574	3530	6382
60 s	1874	2587	3561	5050	8850

First, the bonus that could be gained on each intervention trial decreased with the time participants took to make it. Participants saw an hourglass on the right hand side of each chip they were testing. The "sand" in the hourglass started running out as soon as the chip appeared and stopped once an intervention was made. In the three experimental conditions, the sand ran out after either 60 s, 8 s, or 4 s.

Second, the potential bonus from each trial was computed by multiplying the maximum amount ($1) with the proportion of time that had elapsed when a participant made an intervention (i.e., the proportion of sand left in the top half of the hourglass). As before, the bonus was also conditional on selecting the correct chip diagram. Participants had to make an intervention and choose a chip diagram even when the time had expired (and so their potential bonus was $0).

Third, in contrast to the previous experiments, participants could only make *one* intervention on every chip they tested in order to avoid developing an additional tradeoff where repeated interventions were penalized.

4.2. Results

The following analyses excluded trials in which participants ran out of time before intervening, because after reaching a bonus of $0 there remains no incentive to decide quickly. As a consequence, four participants were completely excluded from the analyses (one in the 8 s condition, and three in the 4 s condition).[5] Note that the remaining 93 participants in the 4 s condition exceeded the time limit on average on only 1.3 trials (out of 27).

4.2.1. Timing manipulation

There was a main effect of timing condition on participants' average response time, $F(2,290) = 28.65, p < 0.001$. The average response times were 2637 ms, 3098 ms, and 4327 ms, in the 4 s, 8 s, and 60 s conditions, respectively (for percentiles, see Table 3). The average increase in response time from the 8 s to the 60 s conditions was more than twice as large than the increase from the 4 s to 8 s conditions.

4.2.2. Intervention choices

To compare strategies in the three conditions, we again found the best-fitting parameters for τ and θ, shown in Fig. 15A. We also fit the full hierarchical Bayesian model from Fig. 9 to extract the population-level parameter μ. Fig. 15B shows the distribution of samples of μ for the three new conditions, as well as Experiment 1. The plot shows that in all three time pressure conditions participants' behavior was better described PTS, compared to Experiment 1. Furthermore, the two fastest conditions (8 s and 4 s) showed a stronger PTS tendency than the 60 s condition indicating that participants were sensitive to not only the presence of, but also the amount of time pressure. The difference between 4 s and 8 s is less pronounced, consistent with the smaller difference in RTs in those conditions.

Importantly, this change in behavior towards PTS under time pressure was not driven by an increase in noisy choice behavior (i.e. random guessing). On the contrary, in the two fast conditions (4 s and 8 s), participants with high IG scores (high θ) actually looked noisier than those using PTS

[5] Note that none of the key results reported below depend on the exclusion of these participants. But since they strictly speaking should not have been influenced by the timing manipulation in the way we intended, they were excluded.

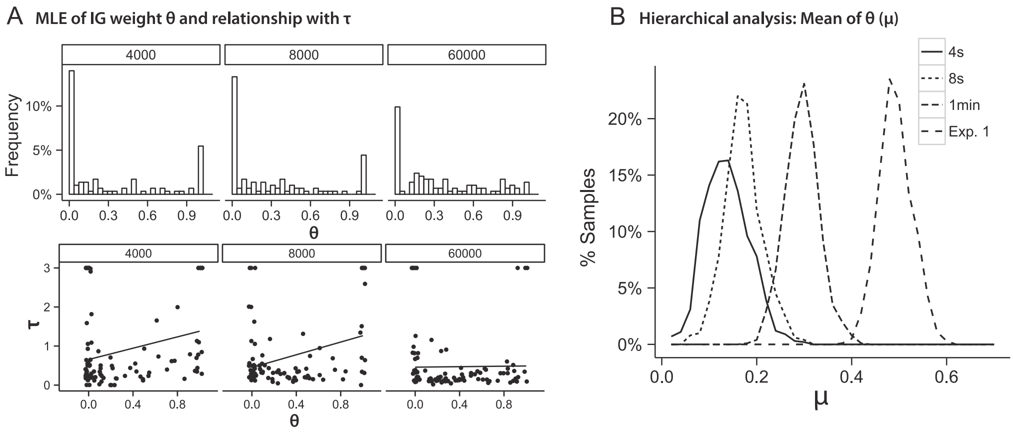

Fig. 15. (A) Top plot shows histograms of best-fitting mixture weight θ by condition of Experiment 3. The bottom shows the relationship between ML estimates of θ and τ (high values indicate choices that resemble guessing). (B) Histogram of samples of μ parameter from the hierarchical model in Fig. 9 for all conditions of Experiment 3, as well as Experiment 1.

(see Fig. 15, panel A). Thus, although the histogram of ML estimates of θ at the top of panel A shows that some participants were fit as extreme IG users, many of them were mostly choosing randomly.

4.3. Discussion

As expected, participants responded more quickly given shorter deadlines and conducted more positive tests as a consequence. These results provide further support for the flexibility of people's intervention strategies and show that the shift towards IG that was observed in Experiment 2 is not the only possible direction of strategy change. They also support the hypothesis that PTS is cognitively easier to use than discriminatory strategies.

5. General discussion

This article asked what strategies people use when planning causal interventions to test their hypotheses about the world. We considered two strategies. Information gain (IG) has been proposed as both as a normative account of information search (Murphy, 2001; Nelson, 2005; Tong & Koller, 2001), and a psychological model of causal learning via interventions (Steyvers et al., 2003). Positive test strategies (PTS), on the other hand, have been shown to underlie hypothesis testing behavior in a variety of domains, including rule learning and logical reasoning (Klayman & Ha, 1987; Jones & Sugden, 2001; Wason, 1960).

To summarize our main findings, Experiment 1 found that people did not behave as predicted by the IG model, even though our task was designed to minimize the cost of using a discriminatory strategy. Instead, the data were best explained by a *combined model* that describes behavior as a mixture of discriminatory (IG) and confirmatory (PTS) hypothesis testing. Two experimental manipulations then demonstrated that people's strategies change in response to other variables in the task. Experiment 2 showed that participants' tests became more discriminatory after they were first exposed to a range of problems with a low expected payoff of PTS. In other words, without otherwise changing the instructions of the task, participants behaved more in line with IG when previous experience showed them that PTS was inferior. Experiment 3 demonstrated the opposite pattern. When put under time pressure, participants were more prone to conduct positive tests.

Our findings add to the body of work showing that people can effectively reason about potential interventions in order to learn about the structure of causal systems (Bramley et al., 2014; Lagnado & Sloman, 2006; Meder, Hagmayer, & Waldmann, 2008; Sobel & Kushnir, 2006; Waldmann &

Hagmayer, 2005). In the present study, learners not only chose the correct structure (out of two possibilities) with high probability, they did so on the basis of very few interventions (usually one). That this high accuracy obtained even on early trials indicates that extensive experience with the task was not necessary for choosing effective interventions.

Two other prior studies have compared formal models to explain causal interventions. Contrary to our results, Steyvers et al. (2003, Experiments 2 and 3) found that a *rational test model*, a version of IG, fit the data well and thus concluded that "people's intervention choices may be explained as rational tests given their own subjective beliefs about causal structure" (p. 481). Unlike this study, however, our experiments were designed to test the predictions of IG without making assumptions about participants' hypothesis space. By providing two hypotheses and no prior observations, we strongly encouraged them to have an equal prior belief over the two hypotheses. In contrast, Steyvers and colleagues had to infer people's priors over the full hypothesis space, because participants only indicated their favorite graph after an observation phase. Interestingly, the auxiliary assumption needed to improve the fit of IG (that the favorite hypothesis is being compared to its subgraphs) yields strikingly similar predictions to the PTS strategy when applied to only one hypothesis: it chooses nodes in proportion to the number of effects that can be produced by intervening on them (see Fig. 8, last row in Steyvers et al., 2003). Therefore, the interventions observed in their study and ours are not inconsistent although we interpret them differently. Instead of attempting to discriminate hypotheses, we argue that many of their learners and ours were performing positive tests.

Bramley et al. (2014) used a sequential causal learning paradigm in which participants started with an unconstrained hypothesis space and could gradually narrow down the space of possible three-node structures through repeated interventions. They concluded that participants generally chose informative interventions but often ignored evidence from previous interventions. Similar to our study, this finding demonstrates that people are efficient but resource-limited causal learners. Like the authors note, it would be interesting to investigate in a more controlled setting what strategies people use in such complex, sequential intervention tasks. Given our own findings, we suspect that interventions will reflect both discriminatory and confirmatory strategies, but further research is needed to study the exact interplay of intervention choices, belief-updating, and cognitive constraints.

The adaptive nature of intervention choices. We identified two factors that impact strategy selection: First, Experiment 2 demonstrated that learning the strategies' expected payoffs encouraged more discriminatory sampling. Second, Experiment 3 showed that time pressure had the opposite effect of leading to an increase in PTS use. Both results can be understood from an adaptive view of strategy selection (e.g., Gigerenzer & Todd, 1999; Payne et al., 1988; Simon, 1956), which suggests that people select cognitive mechanisms on the basis of conditions in the environment as well as internal constraints on memory or computational capacity.

While Experiment 2 is not the first to demonstrate a behavior change from confirmatory to discriminatory, previous work on hypothesis testing has focused mostly on the impact of task framing or context. For example people behave in a more falsificationist manner in the classic Wason Card Selection task (Wason, 1983) when it is phrased as a cheater-detection problem, such as testing if people violate the social contract rule "If you take the benefit, then you pay the cost." (e.g., Cosmides, 1989; Griggs & Cox, 1982), or some other normative rule (Cheng & Holyoak, 1985). In contrast, the behavior change observed in Experiment 2 was based on *experience* with the task rather than the manner in which the problem was presented. This finding ties in with others showing that people select and adapt learning strategies based on their expected payoff in an environment (e.g., Bröder, 2003; Marewski & Schooler, 2011; Rieskamp & Otto, 2006; von Helversen & Rieskamp, 2008).

Comparatively fewer studies have addressed the impact of computational cost on strategy selection, as was our goal in Experiment 3. Some studies manipulate the cost of searching for information (e.g., Bröder, 2000; Newell & Shanks, 2003; Pachur & Hertwig, 2006; Rieskamp & Hoffrage, 2008), either directly or through time pressure, and find corresponding strategy changes toward less information-greedy algorithms. However, these studies target the external cost of collecting information, rather than *cognitive* resources. The present study offers a contribution to adaptive theories of judgment and decision-making by showing strategy changes as a consequence of manipulating cognitive capacity through time pressure.

For the most part, work on causal reasoning and structure learning has rarely adopted such an adaptive viewpoint on strategy selection (but see Meder, Gerstenberg, Hagmayer, & Waldmann, 2010; Rottman, 2014). Instead the focus has been to propose optimal principles of causal reasoning (e.g., Cheng, 1997; Gopnik et al., 2004; Griffiths & Tenenbaum, 2005; Rehder, 2014; Sobel, Tenenbaum, & Gopnik, 2004) to either explain people's behavior or to point out specific violations of these principles. We suggest that it would be worthwhile to investigate further the *repertoire* of strategies that people use to approach and simplify causal reasoning problems, and the factors that drive which strategies are selected.

Why PTS?. Why do people use PTS at all when IG, at least in the present task, leads to higher or equal expected accuracy? This question particularly applies to Experiment 1 in which participants were under no time pressure and yet a sizable portion were better fit by PTS. As multiple authors have pointed out, PTS can be a rational strategy if certain conditions are met, most importantly that the hypothesis space is *sparse* (Austerweil & Griffiths, 2008; Navarro & Perfors, 2011; Oaksford & Chater, 1994). Broadly speaking, sparsity holds when each hypothesis is consistent with a small number of observations that do not overlap much between hypotheses. If this requirement is satisfied, PTS often yields highly diagnostic answers. Similarly in our task, PTS is a good strategy if the hypotheses predict very *different effects* of an intervention. This was the case for problems used in the PTS= condition of Experiment 2 (see Fig. 13), for example.

One explanation for a PTS preference is therefore that the types of causal hypotheses we come across in every day life are sparse in the sense that effects tend to be produced by only a small number of causal mechanisms. The chip-testing paradigm may have contributed to an assumption of sparsity by evoking a particularly mechanistic domain of causal reasoning, encompassing causal systems like computers, smartphones, or other electronic devices. Relationships in this domain tend to be deterministic (button clicks and key presses usually create their intended effects), and there often exists only one way to produce an outcome (e.g., there is exactly one way to turn on most devices or to make them "work" in a desired way). Such a sparse hypothesis space may have encouraged the use of PTS. One interesting avenue for future work is thus to investigate if domain-knowledge can impact whether people choose a discriminatory or confirmatory strategy. To do so, the first step would be to identify people's intuitions about "causal sparsity" in different domains and then assess whether their default strategies vary in the way just described.

Limitations of PTS. One contribution of this article was to develop an interpretation of positive testing in causal intervention tasks (see Eq. (6)), which has not been attempted before. Although we found evidence for this particular interpretation in our experiments, it is beyond the scope of this article to provide a full test of its feasibility in other contexts. Several aspects of the strategy deserve further consideration.

First, we explain above that by choosing nodes with high *centrality*, PTS appears to favor interventions that can produce a large number of effects in one graph. But because the expected number of effects depends on not only a graphs structure but also its parameters (e.g., causal strengths and background causes) it would be possible, for example, to have a central node that produces few effects if intervened on, because the links to its child nodes are very weak. Since in our experiments centrality and expected number of effects overlapped (because causal strengths did not vary), we were unable to test whether people would still prefer central nodes if they led to fewer effects. It would be worth exploring this question in future work.

A second limitation is that the above definition of PTS assumes that *all* hypothesized graphs are treated equally when computing PTS scores. Since this article focused on single interventions given only two hypotheses, this assumption seems unproblematic. However, it is not unreasonable to think that with a larger hypothesis space only a subset of graphs might be considered at any point in time or that Eq. (6) might need to be extended by a weighting term, $P(g)$, to ensure that graphs with high priors are given higher preference. Again, future work could address this question and investigate how people use a PTS strategy when more hypotheses are being considered.

There are other properties of causal graphs and intervention types that could be interesting for future study. For example, one could ask if PTS generalizes to cases in which causal links are inhibitory. It is possible that inhibition is treated the same way as activation (such that non-activation gets effectively treated as "positive evidence"), but it might follow a different dynamic. It would also be interesting to see how the PTS strategy is used when learners can make multiple interventions at the same time (similar to the procedure used by Bramley et al., 2014). Learners might then intervene on multiple nodes if these are necessary to create all the effects in a given structure.

Cognitive mechanisms. With the combined model of intervention selection, we have developed a *descriptive* framework that captures to what degree people behave in accordance with a discriminatory IG strategy or a confirmation-based PTS model. However, it is still an open question what cognitive mechanisms give rise to these strategies and a mix of them.

For example, it is an open question how people determine the value of a node for confirming or discriminating hypotheses. They might use surface-level features of the causal networks, like root causes or links that differ between hypotheses, to determine a node's value, for example. It's also unclear how strategies are then combined to yield a choice. Rather than fully computing both strategies and weighting them, for example, it is possible that they are recruited sequentially or partially. The response time findings reported for Experiment 1 hint at the possibility that rather than merely combining the two, there can exist competition between them, such that decisions become more difficult when the two strategies disagree. This could happen, for example, if people started out using a simple confirmatory strategy and then (sometimes) check whether the resulting intervention plan also serves discriminatory purposes. A negative answer could lead the learner to initiate an IG-based search, leading to longer RTs. This sequential process is also compatible with the finding of higher PTS-use in Experiment 3, because the time pressure manipulations could have led participants to stop after computing PTS, without checking for the discriminatory value of an intervention.

Again, future work should address plausible mechanistic implementations of discriminatory and confirmatory strategies and their combinations. In particular, we think it might be worthwhile to study the time course of intervention decision-making with process measures like eye or mouse tracking to investigate, for example, when and to what degree participants compare hypotheses to one another, or search for root nodes in individual graphs.

6. Conclusions

Previous work has argued that people follow discriminatory principles of information search when planning causal interventions. In contrast, we found that peoples interventions were not well described by information gain alone but instead reflected the influence of a confirmatory positive testing strategy. Learners flexibly adapted their mix of strategies, performing more discriminatory tests when they produced large gains in accuracy and more confirmatory tests when time was short. These results suggest a more adaptive view of self-directed causal structure learning than has so far been considered.

Acknowledgments

This work was supported by Grant No. BCS-1255538 from the National Science Foundation and the Intelligence Advanced Research Projects Activity (IARPA) via Department of the Interior (DOI) contract D10PC20023 to TMG. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DOI, or the U.S. Government. Thanks to Gianni V. Galbiati for helping code the timer in Experiment 3.

Appendix A

Fig. 16 shows the hierarchical Bayesian model used to derive posterior predictions for the two individual models, IG and PTS, to create the triangle prediction plots in the main body of the text.

Appendix B

Table 4 shows other plausible discriminatory and non-discriminatory models that were fit to the data, along with a series of goodness-of-fit measures, which can be compared the original IG and PTS models in Table 2.

B.1. Alternative discriminatory models

Our aim in testing additional versions of IG was to determine if its inadequate fit was due to constraints imposed on its parameters. First, whereas our baseline version of IG assumed equal prior beliefs over the causal structures (common cause, common effect, chain, and one-link), in another version those priors were free parameters. Second, whereas baseline IG set the strengths of the causal links to the true (and instructed) value of .8, a third version allowed strength to vary freely. Table 4 reveals that both versions achieved better fits than baseline IG according to BIC, which corrects for their greater number of parameters. Importantly, however, neither outperformed the combined model (cf. Table 2). Thus, the conclusion that IG is an inadequate account of the present data cannot be attributed to constraints placed on its parameters. Note that the somewhat implausible parameter values produced by these additional fits also raise questions about whether they are faithfully capturing learners' underlying decision processes. For example, the near-zero priors for the common cause and one-link strikes us as unrealistic given that subjects were not told to favor any structure over another. Similarly, the fitted causal strength value of .4 is surprising given that subjects were told that the links' strength was .8 and received training in which they experienced that strength. An additional reason to doubt that strong structure priors influenced the data is that participants endorsed all structures with equal likelihood (50%) at the end of trials, so there was no bias at this stage.

We also tested an alternative discriminatory model, probability gain (Baron, 1985; Nelson, 2005), the predictions of which turned out to be almost identical to IG in the current task ($r = .99$ on the problems of Experiment 1) and thus offered a comparable fit to the data. The combination of PG and IG, shown in Table 5, also does not offer a better fit to the data than the original combined model.

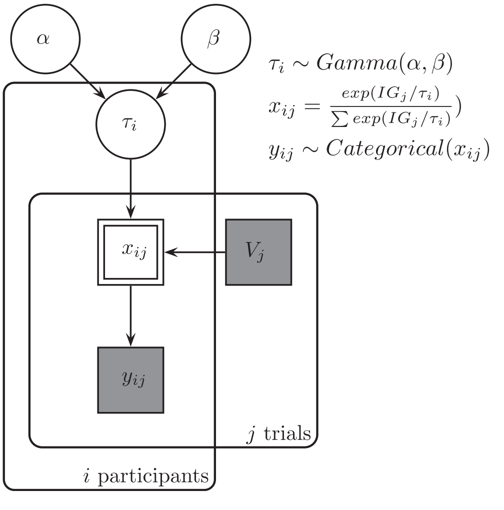

Fig. 16. Hierarchical Bayesian model of how one of the strategies (IG or PTS) generate intervention choices given a probabilistic choice procedure. Each trial, j, corresponds to one of 27 problem types for which each participant chose one intervention, y. V_j is a three-vector with the IG or PTS score of each node on problem j.

Table 4
Other models tested on the data of Experiment 1. *Probability Gain* maximizes the probability of making a correct structure choice after the next intervention. The *IG* & *Prior* model estimates parameter values for the prior beliefs over the four graph types: Common Cause (*CC*), Common Effect (*CE*), *Chain*, and One-Link (*OL*). The *IG* & *Strength* model estimates the probability, *p*, of one node to turn on its child nodes as a free parameter. *PTS (sum.)* uses the sum of relative centralities (see Eq. (6)) instead of the maximum. *Expected Changes* calculates the expected number of effects across both hypotheses, that is, this strategy amounts to the desire to make as many things "happen" as possible.

Model	Log lik.	BIC	R^2	τ	$p(CC)$	$p(CE)$	$p(Chain)$	$p(OL)$	p
Probability Gain	−442	886	.61	.39	–	–	–	–	–
IG & Prior	−324	656	.78	.39	.00	.50	.47	.03	–
IG & Strength	−331	666	.75	.35	–	–	–	–	.40
PTS (sum.)	−490	989	.48	.43	–	–	–	–	–
Exp. Change	−448	904	.54	.39	–	–	–	–	–

Table 5
Combined versions of alternative non-discriminatory and discriminatory models. Compare these models to the fit of the combined model in Table 2 in the main text. All have lower log likelihoods, indicating they fit the data less well.

Model	Log lik.	BIC	R^2	τ	θ
PG & PTS	−243	502	.82	.22	.38
IG & PTS (sum.)	−236	489	.83	.24	.53
IG & Exp. Change	−243	501	.83	.25	.50

B.2. Alternative non-discriminatory models

We also tested two alternative versions of the non-discriminatory PTS model. According to *Expected Change*, the value of a node is the average, across graphs, of the number of expected effects, that is, the number of nodes that can be expected to be turned on after an intervention. As Table 4 shows, it does not fit as well as PTS. For example, for Problems 26 and 27 this model favors an intervention on the chain's root node (whose expected change is 1 [2 descendants in the chain, 0 in the one-link]) over the cause of the single link structure (.5; [0,1]). In fact however, learners were about evenly divided between the two, as predicted by baseline PTS.

By using a sum operator instead of a maximum in Eq. (6), the second variant of PTS favors interventions with the largest number of link activations over all graphs. For example, for Problem 3 this model favors the root of the common cause (2 activations in the common cause structure + 1 in the chain = 3) over the root of the chain (0 + 2 = 2). Learners instead favored them both equally, consistent with a maximum operator. Note that yet another alternative in which an average is used instead a sum yields the same predictions.

We also combined each of these models with IG to see if they complement IG better than baseline PTS. Table 5 reveals that they do not.

References

Ahn, W.-K., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology, 41*, 361–416.

Anderson, J. R. (1990). *The adaptive character of thought*. Psychology Press.

Austerweil, J., & Griffiths, T. (2008). A rational analysis of confirmation with deterministic hypotheses. In *Proceedings of the 30th annual conference of the cognitive science society* (pp. 1041–1046).

Austerweil, J., & Griffiths, T. (2011). Seeking confirmation is rational for deterministic hypotheses. *Cognitive Science, 35*, 499–526.

Baron, J. (1985). *Rationality and intelligence*. Cambridge, England: Cambridge University Press.

Bonawitz, E. B., Ferranti, D., Saxe, R., Gopnik, A., Meltzoff, A. N., Woodward, J., et al (2010). Just do it? Investigating the gap between prediction and action in toddlers causal inferences. *Cognition, 115*, 104–117.

Bramley, N. R., Lagnado, D. A., & Speekenbrink, M. (2014). Conservative forgetful scholars – How people learn causal structure through sequences of interventions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. http://dx.doi.org/10.1037/xlm0000061.

Bröder, A. (2000). Assessing the empirical validity of the take-the-best heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1332.

Bröder, A. (2003). Decision making with the "adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 611–625.

Chater, N., & Oaksford, M. (2008). *The probabilistic mind: Prospects for Bayesian cognitive science*. Oxford University Press.

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

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17*, 391–416.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31*, 187–276.

Gelman, A., Meng, X.-L., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica, 6*, 733–760.

Gigerenzer, G., Todd, P. M., & the ABC Research Group (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.

Ginzburg, I., & Sejnowski, T. J. (1996). Dynamics of rule induction by making queries: Transition between strategies. In *Proceedings of the 18th annual conference of the cognitive science society* (pp. 121–125).

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111*, 3.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology, 51*, 334–384.

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology, 73*, 407–420.

Gureckis, T. M., & Markant, D. (2009). Active learning strategies in a spatial concept learning game. In *Proceedings of the 31st annual conference of the cognitive science society* (pp. 3145–3150).

Gureckis, T. M., & Markant, D. (2012). A cognitive and computational perspective on self-directed learning. *Perspectives on Psychological Science, 7*, 464–481.

Hagmayer, Y., & Meder, B. (2012). Repeated causal decision making.

Jones, M., & Sugden, R. (2001). Positive confirmation bias in the acquisition of information. *Theory and Decision, 50*, 59–99.

Kim, N. S., Ahn, W.-k., et al (2002). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology-General, 131*, 451–476.

Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94*, 211–228.

Klayman, J., & Ha, Y.-W. (1989). Hypothesis testing in rule discovery: Strategy, structure, and content. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 596.

Kruschke, J. K. (2010). What to believe: Bayesian methods for data analysis. *Trends in Cognitive Sciences, 14*, 293–300.

Kruschke, J. (2013). Bayesian estimation supersedes the t-test. *Journal of Experimental Psychology: General, 142*, 573–603.

Lagnado, D. A., & Sloman, S. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 856.

Lagnado, D. A., & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 451.

Lee, M. D. (2011). How cognitive modeling can benefit from hierarchical bayesian models. *Journal of Mathematical Psychology, 55*, 1–7.

Lindley, D. V. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 986–1005.

Marewski, J. N., & Schooler, L. J. (2011). Cognitive niches: An ecological model of strategy selection. *Psychological Review, 118*, 393.

Markant, D., & Gureckis, T. M. (2012a). Does the utility of information influence sampling behavior? In *Proceedings of the 34th annual conference of the cognitive science society*. Austin, TX: Cognitive Science Society.

Markant, D., & Gureckis, T. M. (2012b). One piece at a time: Learning complex rules through self-directed sampling. In *Proceedings of the 34th annual conference of the cognitive science society*. Austin, TX: Cognitive Science Society.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Freeman.

Meder, B., Gerstenberg, T., Hagmayer, Y., & Waldmann, M. R. (2010). Observing and intervening: Rational and heuristic models of causal decision making. *The Open Psychology Journal, 3*, 119–135.

Meder, B., Hagmayer, Y., & Waldmann, M. R. (2008). Inferring interventional predictions from observational learning data. *Psychonomic Bulletin & Review, 15*, 75–80.

Murphy, K. P. (2001). Active learning of causal Bayes net structure. Technical report, Department of Computer Science, U.C. Berkeley.

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature, 434*, 387–391.

Navarro, D. J., & Perfors, A. F. (2011). Hypothesis generation, sparse categories, and the positive test strategy. *Psychological Review, 118*, 120.

Nelson, J. D. (2005). Finding useful questions: On Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review, 112*, 979.

Nelson, J. D., Divjak, B., Gudmundsdottir, G., Martignon, L. F., & Meder, B. (2014). Children's sequential information search is sensitive to environmental probabilities. *Cognition, 130*, 74–80.

Nelson, J. D., Tenenbaum, J. B., & Movellan, J. R. (2001). Active inference in concept learning. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd conference of the cognitive science society* (pp. 692–697). Cognitive Science Society Mahwah, NJ: Erlbaum.

Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing one-reason decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 53.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101*, 608.

Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: Retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 983.
Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
Pearl, J. (2000). *Causality: Models, reasoning and inference* (Vol. 29). Cambridge Univ. Press.
Rehder, B. (2014). Independence and dependence in human causal reasoning. *Cognitive Psychology, 72*, 54–107.
Rieskamp, J., & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica, 127*, 258–276.
Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General, 135*, 207.
Rottman, B. M. (2014). Information search in an autocorrelated causal learning environment. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th annual conference of the cognitive science society*. Austin, TX: Cognitive Science Society.
Scheibehenne, B., & Pachur, T. (2013). Hierarchical bayesian modeling: Does it improve parameter stability? In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual conference of the cognitive science society*. Austin, TX: Cognitive Science Society.
Schulz, L. E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science, 10*, 322–332.
Shafto, P., Goodman, N. D., & Griffiths, T. L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology, 71*, 55–89.
Simon, H. E. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*, 129–138.
Sloman, S. A., & Lagnado, D. A. (2005). Do we "do"? *Cognitive Science A Multidisciplinary Journal, 29*, 5–39.
Sloman, S. A., Love, B. C., & Ahn, W.-K. (1998). Feature centrality and conceptual coherence. *Cognitive Science, 22*, 189–228.
Sobel, D. M., & Kushnir, T. (2006). The importance of decision making in causal learning from interventions. *Memory & Cognition, 34*, 411–419.
Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and bayesian reasoning in preschoolers. *Cognitive Science, 28*, 303–333.
Steyvers, M., Tenenbaum, J. B., Wagenmakers, E.-J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science, 27*, 453–489.
Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1). Cambridge Univ. Press.
Tong, S., & Koller, D. (2001). Active learning for structure in bayesian networks. *International joint conference on artificial intelligence* (Vol. 17, pp. 863–869). Lawrence Erlbaum Associates.
von Helversen, B., & Rieskamp, J. (2008). The mapping model: A cognitive theory of quantitative estimation. *Journal of Experimental Psychology: General, 137*, 73.
Waldmann, M. R., & Hagmayer, Y. (2005). Seeing versus doing: Two modes of accessing causal knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 216.
Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Jjournal of Experimental Psychology, 12*, 129–140.
Wason, P. C. (1983). Realism and rationality in the selection task. *Thinking and Reasoning: Psychological Approaches*, 44–75.