# A unifying computational framework for teaching and active learning

## Scott Cheng-Hsin Yang[1], Wai Keen Vong[1], Yue Yu & Patrick Shafto
### Department of Mathematics & Computer Science, Rutgers University – Newark

Traditionally learning has been modeled as passively obtaining information or actively exploring the environment. Recent research has introduced models of learning from teachers that involve reasoning about why they have selected particular evidence. We introduce a computational framework that takes a critical step toward unifying active learning and teaching by recognizing that meta reasoning underlying reasoning about others can be applied to reasoning about oneself. The resulting Self-Teaching model captures much of the behavior of information-gain-based active learning with elements of hypothesis-testing-based active learning and can thus be considered as a formalization of active learning within the broader teaching framework. We present simulation experiments that characterize the behavior of the model within three simple and well-investigated learning problems. We conclude by discussing such theory-of-mind-based learning in the context of core cognition and cognitive development.

## 1 Introduction

We learn by interacting with the world, by choosing what to look at and what to do. Through these actions we accumulate knowledge about the world. We also learn from other people. Sometimes these people are knowledgeable about the world and demonstrate actions with the intent to foster learning, as is the case with informal notions of teaching (Shafto, Goodman, & Griffiths, 2014). Other times, people may not be fully knowledgeable, but may nevertheless select demonstrations that we observe (Shafto, Eaves, Navarro, & Perfors, 2012). Indeed, in some teaching contexts, teachers may be no more knowledgeable than ourselves, in which case their actions may be no better than uninformed choices of our own (Markant & Gureckis, 2014). To understand learning in realistic contexts, we must understand actions that are chosen for learn as in *active learning*, actions that are chosen for others and their implications for learners as in *teaching*, as well as integrating both active learning and teaching into a singular experience.

The literatures on active learning and teaching are not well-integrated despite having been the subject of extensive independent and joint research. For instance, there is a robust debate in the educational psychology literature regarding the relative merits of direct instruction and exploratory learning, which are the analogs of formal models of teaching and active learning (Bruner, 1961; Kirschner, Sweller, & Clark, 2006; Tobias & Duffy, 2009). Research has focused on identifying which of these is more effective for promoting

---

[1]Both authors contributed equally to this paper.

learning, with mixed evidence in favor of exploration (Steffe & Gale, 1995), direct instruction (Kirschner et al., 2006) and rough equivalence (Klahr & Nigam, 2004). There has also been extensive research into formalizing active learning and teaching as different kinds of computational models. Early computational models of active learning date back to early work by Bruner, Goodnow, and Austin (1956) and sustained recent attention in this area has resulted in a variety of computational models which converge around the idea that people choose examples that reduce some globally-defined notion of uncertainty over possible hypotheses (Nelson, 2005; Coenen, Nelson, & Gureckis, 2017; Crupi, Nelson, Meder, Cevolani, & Tentori, 2018). Similarly, recent modeling has formalized teaching and learning from teaching as a form of recursive reasoning between a teacher and a learner (Shafto et al., 2014; Shafto & Goodman, 2008; Shafto, Goodman, & Frank, 2012; Frank, 2014). These models characterize this form of social learning as a joint inference about the true hypothesis and the process by which the teacher selects data (Shafto et al., 2014; Shafto & Goodman, 2008). Although the problems of active learning and teaching are of common interest to many researchers, to our knowledge no work has attempted to synthesize both kinds of learning as a common formal framework in which we may understand both from a singular learning mechanism.

In this paper, we explore a unified approach to modeling active learning and teaching. The goal of both active learning and teaching is for the learner to infer the true hypothesis. The defining feature of active learning is that the learner can intervene the world however he or she likes to discover that hypothesis. However, because the learner does not know the true hypothesis, they learn by choosing interventions based on an intermediate objective. The literature on computational models of active learning reveals two relatively distinct sets of phenomena summarized by two kinds of objectives. One objective is that people select interventions with the goal of information gain or minimizing uncertainty (Nelson, 2005; Yang, Lengyel, & Wolpert, 2016; Crupi et al., 2018). The other objective is that people select interventions to test hypotheses with a bias toward confirmation (Wason, 1960; Klayman & Ha, 1987; Navarro & Perfors, 2011; Coenen et al., 2017; Bramley, Dayan, Griffiths, & Lagnado, 2017).

In contrast to computational models of active learning, the defining feature of formal models of teaching is that the learner receives demonstrations of an intervention from a knowledgeable teacher rather than choosing it themselves. Moreover, since demonstrations are chosen by a teacher who knows the true hypothesis, they will therefore choose demonstrations that they think will best help the learner infer the true hypothesis after observing the demonstrations. We explore the idea that models of teaching can be transformed to perform active learning by assuming that the correct hypothesis is unknown to the teacher. This has the virtue of preserving the core structural element of theory of mind but applying it reflexively when only a single learner is present. In this paper, we mathematically derive this form of *self-teaching* and show through simulations that the Self-Teaching model exhibits a range of active learning behaviors, favoring interventions that effectively reduces uncertainty in some cases and those that test for hypothesis in other cases. Thus, self-teaching not only unifies the phenomena of active learning and teaching; it may also offer insight into how the different phenomena of active learning can be explained.

The differences between Self-Teaching and standard models of active learning, such as Expected Information Gain (EIG), derive from different methods of computing which
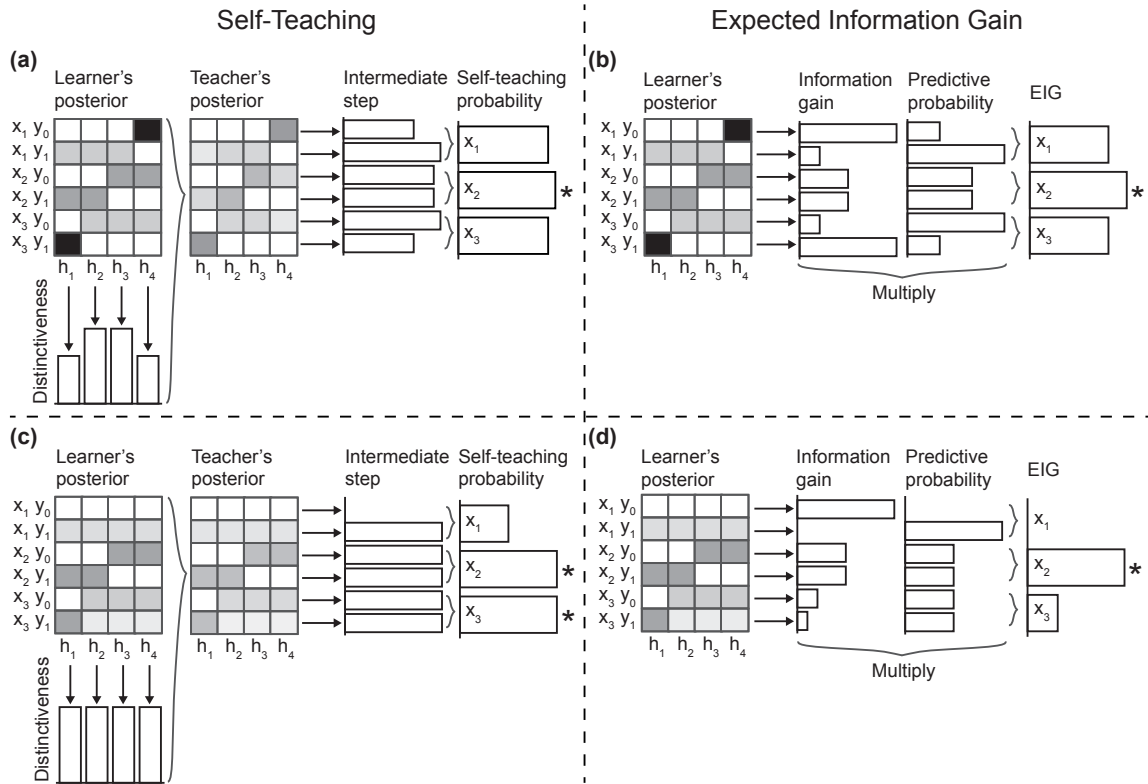
*Figure 1*. **An illustrative comparison between the Self-Teaching model and Expected Information Gain (EIG) model**. Panels (a) and (b) show a case where the two models begin with the same learner's posterior and yield the same most favorable intervention (the starred intervention). In contrast, panels (c) and (d) show a case where the two models differ in the most favorable choice: the Self-Teaching model favors the second and third interventions equally, whereas the EIG model predominantly favors the second one. Both models start with the learner's posterior, the probability that an agent infers a hypothesis $h$ to be the underlying truth after seeing an intervention $x$ and the corresponding observation $y$. The posterior probabilities are represented in gray scale (black = 1; white = 0). In the Self-Teaching model as shown in panels (a) and (c), a "distinctiveness" is computed for each hypothesis from the learner's posterior (downward arrows). A hypothesis is less distinctive if it is on average less likely to be inferred across interventions and observations (i.e., a small column sum). The learner's posterior is weighted by the distinctiveness vector to yield the teacher's posterior. An intermediate step that marginalizes the teacher's posterior over the hypotheses for each intervention-observation $(x, y)$ pair is shown. A further marginalization over the observations yields the Self-Teaching probabilities. In the EIG model in panels (b) and (d), the information gain and predictive probability of each $(x, y)$ pair are calculated from the learner's posterior. The information gain of an $(x, y)$ pair is high if the learner's posterior is concentrated across the hypotheses for that pair. The EIG of an intervention is calculated by weighting the information gains with the predictive probabilities and marginalizing out the observations.

example to choose next (Figure 1).[2] In this example, we consider one case (Figure 1, top row) where the predictions of the Self-Teaching and EIG models exhibit the same ordering of preferences for which intervention to select, while in the other case (Figure 1, bottom row) is an instance where the two models have exhibit different preferences. Both models make use of the learner's posterior but in different ways. The Self-Teaching model takes the sums of the marginal distributions over hypotheses and observations to determine its response, while the EIG model calculates the concentration of posterior mass across different hypotheses as the information gain. A key component of the Self-Teaching model is the inverse of the marginal probabilities of the hypotheses, which we call *distinctiveness*. The distinctiveness of a hypothesis' indicates how unlikely a learner would infer that hypothesis if all interventions and observations are equally likely to occur. Equivalently, from a teacher's perspective, it indicates how difficult it would be to learn the hypothesis if the teacher were to uniformly randomly choose an intervention and observation. The Self-Teaching model has a tendency towards discovering hypotheses with high distinctiveness. From the simulations we see that the distinctiveness of each hypothesis determines whether the two models will have the same preference or not. When the hypotheses are all of the same level of distinctiveness as seen in the bottom row of Figure 1, then the behavior of the Self-Teaching model differs from EIG.

The rest of the paper will be devoted to formalizing and exploring active learning as a form of self-teaching. In Section 2, we introduce the mathematical models of Expected Information Gain and Self-Teaching. In Section 3.1, we describe the simulation of the learning tasks. In Section 3.2, we present the simulation results which characterize the behavior of the Self-Teaching model relative to active learning models. In Section 4, we revisit the potential implications of a unified model of teaching and active learning and describe open questions.

## 2 Theory

To bridge the gap between these two modes of learning, the focus of this theory section is to modify the teaching model (Shafto & Goodman, 2008; Shafto et al., 2014) in two stages to demonstrate how teaching may be reduced to a form of active learning. In the first stage, we modify the teaching model from providing both features and labels to just providing instruction on which feature to select (from the top part to the middle part of Figure 2(b)). Mathematically, this involves marginalization over the possible outcomes (labels), which are no longer under the teacher's control. In the second stage, we further modify the teaching model so that the teacher no longer knows the true hypothesis, and averages over their own uncertainty when teaching. Mathematically, this involves marginalization over the possible hypotheses. This kind of teaching is formally equivalent to a learner simulating a teacher to determine how to teach oneself, hence the term *self-teaching* (from the middle part to the bottom of Figure 2(b)).

Because the learner selects interventions for themselves in self-teaching, it is a form of active learning. Nevertheless, it is different from information-gain-based active learning in that the single architecture for meta-reasoning underlies learning for oneself through

---

[2]Figure 1(a) and (b) correspond to the boundary task in Figure 3(a), with additional details of the calculations in the Appendix).
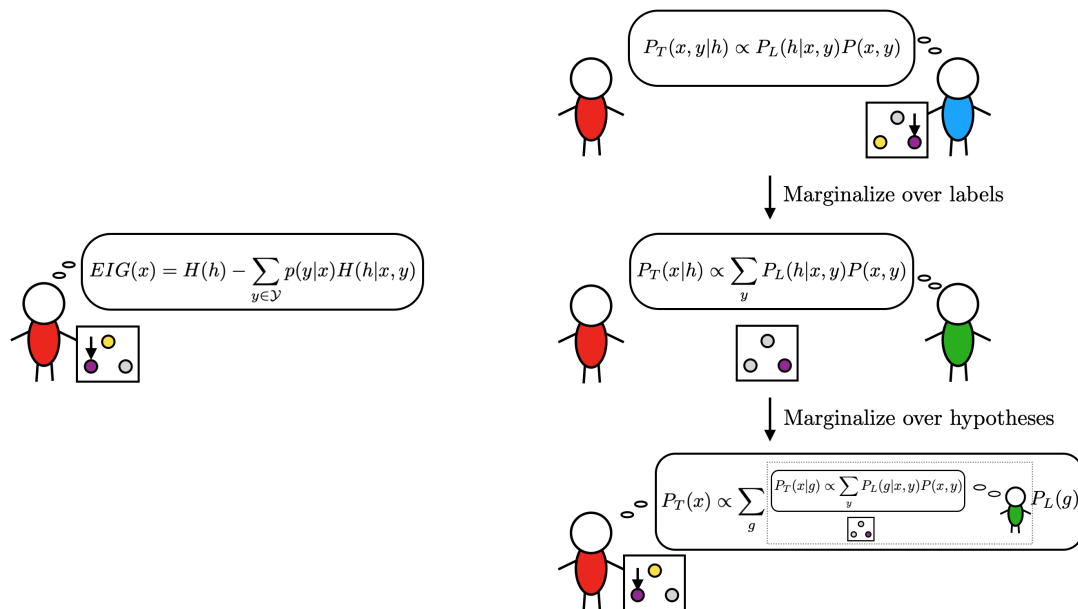
(a) Expected Information Gain (EIG) model     (b) From Teaching to Self-Teaching

*Figure 2*. **Comparison of active learning using either (a) the Expected Information Gain model or (b) the Self-Teaching model (and how it can be derived from teaching)**. In each subfigure, the learner is depicted on the left (in red) and teacher (if it exists), is shown on the right (in blue when performing "Teaching-by-Intervention", and in green when performing "Teaching-by-Instruction"), and the position of the object indicates how either person interacts with it. In the EIG model (left), the learner determines which feature $x$ to select on the object by calculating the expected information gain of choosing each feature. They perform a query on a particular feature (as shown by the downward arrow), and receive a set of observations (represented by the different colors on the circles in the box). On the right, we first illustrate the Teaching-by-Intervention model (top right), where the teacher demonstrates a feature-label pair $(x, y)$ to the learner. By marginalizing over the possible labels $y$, this produces the Teaching-by-Instruction model (middle right), where the teacher instructs the learner which feature $x$ to select (as depicted by the feature shown in purple). Finally, by marginalizing over the hypotheses $g$, we show how the Self-Teaching model is a form of active learning. Here, the learner simulates a self-teacher (using the Teaching-by-Instruction equations), and marginalizes over all hypotheses to determine which feature to select. In this case, the Self-Teacher ends up performing the same query as the EIG model.

active learning, and learning from others through teaching. From a formal perspective, this reflects a difference in the objective being optimized for. As will be discussed in the following subsection, the Self-Teaching model shares the goal of choosing interventions to help inferring the correct hypothesis with the teaching models, but having no access to the truth. It expresses this goal as a bias to confirm whether the true hypothesis is one that is difficult to learn if intervention-observation pairs are to occur uniformly at random (see Figure 1 and related text). This is in contrast with formalizations of active learning such as the Expected Information Gain model. The history of active learning models has been focused on performance. From this perspective, maximizing long-run reward—the posterior probability on the true hypothesis—is optimal, and uncertainty reduction is a well-accepted heuristic. Here, we take a drastically different perspective by presenting a framework where the selection of data and inference from data are *consistent* across teaching and active learning. In other words, the Self-Teaching model is not motivated by an optimization constraint on a performance-related quantity; rather, it is made to satisfy a consistency constraint of keeping the inference from data and selection of data the same in active learning and teaching. After introducing the EIG and Self-Teaching models more formally in this section, we highlight both similarities and differences in the predictions of the two models in Section 3.2.

## 2.1 Notation for inference

In this section, we formalize learning as Bayesian inference and the associated notation. The learner's goal is to infer, from amongst all possible hypotheses, the true hypothesis. We denote a hypothesis by $h$, the true hypothesis by $h^*$ and the space of all hypotheses by $\mathcal{H}$. Each hypothesis is characterized by a particular set of features and labels. We denote a feature and a label by $x$ and $y$, respectively, and the space of possible features and labels by $\mathcal{X}$ and $\mathcal{Y}$, respectively.

Initially, the learner has a prior over hypotheses $P_L(h)$, where the hypotheses refer to either concepts in the concept learning tasks or graphs in the causal learning task (Section 3.1). For all tasks, the prior is set to be uniform. After observing a feature, label pair $(x, y)$, the learner updates their beliefs in accordance with Bayes' rule: $P_L(h|x, y) \propto P(y|x, h)P_L(h)$. After the learner has observed a data point, the new learner's prior $P_L(h)$ is set to be the current learner's posterior $P_L(h|x, y)$, without loss of generality.

## 2.2 The Expected Information Gain (EIG) model

In the problem of active learning, a learner must determine how to efficiently intervene and sample from the environment for effective learning. One approach is to choose which pieces of information would be most helpful at reducing the uncertainty about the true hypothesis, known as *information gain* (MacKay, 1992; Nelson, 2005).[3] Using the expected information gain criterion, the basis for selecting which feature to intervene on is determined by how much information they would gain on average if they intervened on that particular

---

[3]While there exist other approaches for active learning that optimize other objectives (e.g., Nelson, 2005; Markant, Settles, & Gureckis, 2016; Bramley, Lagnado, & Speekenbrink, 2015), in this paper we primarily focus on expected information gain as a comparison to our self-teaching approach because of its wide-spread use.

feature. In order words, the learner selects the feature that maximizes the reduction in uncertainty about the true hypothesis.

In order to calculate the expected information gain, we first define the learner's prior entropy by the Shannon entropy over possible hypotheses as follows:

$$H(h) = \sum_{h \in \mathcal{H}} P_L(h) \log \frac{1}{P_L(h)}. \tag{1}$$

After a learner makes an intervention and observes a feature-label pair $(x, y)$, they update their beliefs via Bayes' rule to obtain a posterior distribution over hypotheses, $P_L(h|x, y)$. From this, we can calculate the learner's posterior entropy after observing a data point $(x, y)$:

$$H(h|x, y) = \sum_{h \in \mathcal{H}} P_L(h|x, y) \log \frac{1}{P_L(h|x, y)}. \tag{2}$$

We can then calculate the *information gain* of intervening on a particular feature $x \in \mathcal{X}$ produced a particular label $y \in \mathcal{Y}$, as follows:

$$IG(x, y) = H(h) - H(h|x, y). \tag{3}$$

However, since the set of possible labels that could be observed contains uncertainty, the *expected information gain* is calculated by weighting the posterior entropy by the probability of observing particular labels given a particular feature as follows:

$$EIG(x) = H(h) - \sum_{y \in \mathcal{Y}} P_L(y|x) H(h|x, y), \tag{4}$$

where $P_L(y|x) = \sum_{h \in \mathcal{H}} P(y|x, h) P_L(h)$ is the predictive distribution over labels. Once the learner has calculated the expected information gain using Equation (4), they can choose to select the feature with the highest expected information gain. However, research has shown that people do not always prefer to maximize information gain, and may instead elect to select features probabilistically in accordance to the expected information gain following a softmax choice rule (Sutton & Barto, 1998).

## 2.3 Teaching models

To unify the phenomena of teaching and active learning, we begin with the pedagogical model introduced by Shafto and Goodman (2008) and Shafto et al. (2014), which we call *Teaching-by-Intervention* here. This model outputs a teaching probability, $P_T(x, y|h)$, which is the probability that the teacher will demonstrate a feature-label pair $(x, y)$ to teach an underlying hypothesis $h$ to a learner. To move from Teaching-by-Intervention to our Self-Teaching model of active learning, we perform a number of marginalizations over possibilities. Each of these marginalizations produces a different kind of teaching model: marginalizing out $y$ gives what we call the Teaching-by-Instruction model with output $P_T(x|h)$ that depends only on $x$ and $h$ (Yang & Shafto, 2017); further marginalizing out $h$ produces the *Self-Teaching* model with output $P_T(x)$, where the probability of selecting a feature is independent of any particular hypothesis. The objective of the Self-Teaching model is to favor $x$ that confirms whether the true hypothesis is distinctive—unlikely to be identified with random observation—and is different from objectives based on information gain.

**2.3.1 Teaching by Intervention.** The standard set-up in a teaching model involves both a teacher and a learner. We assume that the teacher knows the true underlying hypothesis and gives the learner $(x, y)$ pairs as an intervention demonstration to help the learner infer the true hypothesis. The learner's and teacher's reasoning under this kind of pedagogical setting can be formalized as follows (Shafto & Goodman, 2008; Shafto et al., 2014):

$$P_L(h|x, y) \propto P_T(x, y|h) \, P_L(h) \,, \tag{5a}$$

$$P_T(x, y|h) \propto P_L(h|x, y) \, P_T(x, y) \,. \tag{5b}$$

The term $P_T(x, y)$ is set to be uniform because the teacher is indifferent to teaching any particular data points *a priori*. To solve this coupled system of equations, one typically begins by computing the learner's posterior with the observation likelihood: $P_{L_0}(h|x, y) \propto P(y|x, h) \, P_L(h)$. This posterior is then fed into Equation (5b) to obtain an estimate $P_{T_0}(x, y|h)$. Then, this estimate $P_{T_0}(x, y|h)$ is fed into Equation (5a) to obtain an updated estimate $P_{L_1}(h|x, y)$, which again is fed into Equation (5b) to obtain an updated estimate of $P_{T_1}(x, y|h)$. This recursion proceeds until the distributions $P_{L_n}(h|x, y)$ and $P_{T_n}(x, y|h)$ have converged. What this coupled process describes is that the teacher's objective is to select $(x, y)$ so that the learner can correctly infer $h$, and that the learner knows that the teacher is providing demonstrations to them for this purpose.

Once this process has converged, and since the teacher knows the true hypothesis $h^*$, they can then provide a demonstration $(x, y)$ to the learner according to $P_T(x, y|h^*)$, which is the distribution of demonstrations conditioned on the true hypothesis $h^*$. Similarly, as the learner knows that the teacher is attempting to provide demonstrations $(x, y)$ based on the knowledge of $h^*$, the learner can identify which of the hypotheses is the true hypothesis $h^*$ by examining the specific demonstration provided by the teacher. This recursive and cooperative inference involving a knowledgeable teacher makes the learning the true hypothesis more efficient than active learning (Yang & Shafto, 2017).

**2.3.2 Teaching by Instruction.** The next step to bring us closer to Self-Teaching model, in which the learner actively selects examples, is to consider the Teaching-by-Instruction model. In this model, rather than demonstrate to the learner the data point $(x, y)$, the teacher simply instructs the learner about which feature to intervene upon. We introduce this model for purely pedagogical purposes here, although it could be of independent interest in other contexts. Formally, we modify Equations (5b) and (6b) to

$$P_L(h|x, y) \propto P(y|x, h) \, P_T(x|h) \, P_L(h) \,, \tag{6a}$$

$$P_T(x|h) = \sum_{y \in \mathcal{Y}} P_T(x, y|h) \propto \sum_{y \in \mathcal{Y}} P_L(h|x, y) \, P_T(x, y) \,, \tag{6b}$$

where we have rewritten the distribution $P_T(x, y|h)$ as $P(y|x, h) \, P_T(x|h)$ via the chain rule and dropped the subscript $T$ from $P(y|x, h)$ because the label is now provided by the world. The teacher's choice of $x$ now involves marginalization over possible labels $y$ in $P_T(x, y|h)$, resulting in $P_T(x|h)$. In the same manner as Equations (5a) and (5b), these equations are iterated until convergence, and then the teacher samples $x$ from the converged $P_T(x|h^*)$. Equations (5b) and (6b) differ only in the marginalization of the label, $y$.

**2.3.3 Knowledgeability.**  So far, we have assumed that the teacher has perfect knowledge of the true hypothesis. In order to have a model of self-teaching, and instead capture the learner's lack of knowledge, we must relax this assumption. To do this, we introduce the concept of *knowledgeability*, which is modeled by the factor $P(g|h)$, where $g \in \mathcal{H}$ is the probability the teacher *thinks* $g$ is the true hypothesis, assuming that $h$ is the true hypothesis. More concretely, $P(g_j|h_i) = a$ says that if the true hypothesis is $h_i$, then the teacher thinks the true hypothesis is $h_j$ with probability $a$.[4]

A teacher with perfect knowledge of the true hypothesis is described by the *teacher's knowledgeability* $\delta(g|h)$, where $\delta(g_j|h_i) = 1$ when $j = i$ and is 0 otherwise. This means that when the true hypothesis is $h_i$, the teacher thinks that $g_i$ is the true hypothesis with probability 1, and this is true for all hypotheses. Using this knowledgeability factor, we can write the Teaching-by-Instruction distribution of Equation (6b) in a more general form:

$$P_T(x|h) = \sum_{g \in \mathcal{H}} P_T(x|g)\,\delta(g|h)\,. \tag{7}$$

Since the teacher knows the true hypothesis, $\delta_T(g|h)$ is 1 only when $g = h$, and 0 otherwise. Therefore, the marginalization over $h$ with the delta-function simply replaces $g$ with $h$, which recovers Equation (6b) exactly. While the generalization seems trivial in case of Teaching-by-Instruction because it takes the form of a delta function, when the knowledgeability factor is less concentrated, it offers a way to introduce uncertainty about the teacher's belief about the true hypothesis into the teaching framework.

**2.3.4 Self-Teaching.**  Since there is no actual teacher in self-teaching, the learner does not have access to the true hypothesis. While the teacher's knowledgeability factor is concentrated on the true hypothesis, the *self-teacher's knowledgeability* must therefore reflect the learner's current belief. Thus, we set the self-teacher's knowledgeability factor to $\delta_{ST}(g|h) = P_L(g)$. This says that no matter what the (hypothetical) true hypothesis is, the self-teacher believes that the distribution over true hypothesis is the same as the learner's prior, rendering the self-teacher only as knowledgeable as the learner is. Following Equation (7), we add the self-teacher's knowledgeability factor into the Teaching-by-Instruction probability. This turns Equation (6) into the Self-Teaching equations:

$$P_L(h|x,y) = \frac{P(y|x,h)\,\cancel{P_T(x)}\,P_L(h)}{\sum_{h' \in \mathcal{H}} P(y|x,h')\,\cancel{P_T(x)}\,P_L(h')}, \tag{8a}$$

$$P_T(x) = \sum_{g \in \mathcal{H}} P_T(x|g)\,\delta_{ST}(g|h) = \sum_{g \in \mathcal{H}} P_T(x|g)\,P_L(g)$$

$$= \sum_{g \in \mathcal{H}} \sum_{y \in \mathcal{Y}} P_L(g|x,y)\,P_T(x,y)\,\frac{P_L(g)}{Z(g)}. \tag{8b}$$

In Equation (8b), the second equality has the same form as Equation (7) but includes the self-teacher's knowledgeability. To obtain the third equality, we expand the term $P_T(x|g)$ according to Equation (6b). The term, $Z(g) = \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} P_L(g|x,y)\,P_T(x,y)$, is the normalizing constant that comes from changing the proportionality in Equation (6b) to an equality in Equation (8b).

---

[4]This can be considered as a generalization of the knowledgeability factor introduced in Shafto, Eaves, et al. (2012).

Note that the self-teacher's knowledgeability, $P_L(g)$, does not depend on $h$. After marginalizing out $g$, the self-teacher's selection in Equation (8b) depends only on $x$ and not on $h$. Therefore, marginalization using the self-teacher's knowlegeability factor converts the term $P_T(x|h)$ from the Teacher-by-Instruction model that is dependent on $h$ into the term $P_T(x)$ that is independent of $h$. As in the teaching framework, we can compute an updated estimate of the Self-Teaching probability in Equation (8b) by feeding the result into the learner's inference in Equation (8a). However, because the teacher's selection does not depend on $h$, this factor cancels out in Equation (8a). The fact that the teacher's selection is independent of $h$ indicates that the learner gains no extra information about $h$ from reasoning about the self-teacher's selection, which is what one would expect. This is reflected by the fact that the self-teacher samples features to intervene according to Equation (8b), which does not depend on $h$ and hence cannot depend on the true underlying hypothesis, $h^*$. Computationally, this means that the recursive reasoning converges in one step, and only one pass through Equations (8a) and (8b) is needed to compute the probabilities in the Self-Teaching model.

Self-teaching makes predictions about which choices one should make to learn effectively, but differs in the objective and method to do so compared to the Expected Information Gain model. To see this, we first examine the term $Z(g)$ in Equation (8b). Since $P_T(x, y)$ is uniform, the term $Z(g)$ is the average probability that the learner will infer $g$ as the true hypothesis if all pairs of $(x, y)$ have equal probability to be observed. Thus, $Z(g)$ can be thought of as how "common" $g$ is, and its inverse, $Z(g)^{-1}$, can be thought of as how "distinct" $g$ is.[5] The $Z(g)$ term appears in a ratio with the prior, suggesting that it is the distinctiveness relative to the prior that matters. Lastly, two expectations over $h$ and $y$ are taken to yield the final intervention selection. Thus, the Self-Teaching model can be viewed as a hypothesis-testing model with the objective of confirming whether the underlying true hypothesis has high *expected relative distinctiveness*, that is, it selects $x$ so that on average the learner will more likely observe distinctive hypotheses (as measured by $Z(g)^{-1}$) relative to the prior.

## 2.4  Remarks on the models

The learner's inference equations in the three kinds of teaching models presented here — Equations (5a), (6a), and (8a) — show that different learning settings make use of different sources of knowledge. In the Teaching-by-Intervention model, in addition to the learner's prior, Equation (5a) includes the term $P_T(x, y|h)$, indicating that the teacher is the only source of information. In the Teaching-by-Instruction model, Equation (6a) includes the terms $P_T(x|h)$ and $P(y|x, h)$, indicating that information comes partly from the teacher and partly from interacting with the environment. In self-teaching, Equation (8a) only requires knowledge from $P(y|x, h)$, indicating that the environment is the only source of information. This is true of EIG and other active learning models, and demonstrates that self-teaching introduces no additional source of information for inference.

A similar observation can be made for the intervention selection equations. The selections of intervention according to Equations (5b) and (6b) in the Teaching-by-Intervention and Teaching-by-Instruction models respectively both involve the underlying true hypothe-

_____

[5]This concept of distinctiveness of a hypothesis is the same distinctiveness introduced in Figure 1.

sis $h^*$, whereas the selections according to Equation (4) and Equation (8b) do not. Thus, in a teaching setting, the learner benefits from the teacher's knowledge of the true hypothesis, but this does is not true in the case of self-teaching.

The Self-Teaching model and EIG model have at least three important differences. First, the EIG model, as well as other conventional active learning models, only reason about the world, whereas the Self-Teaching model involves a meta-cognitive component that reasons about one's own reasoning. Mathematically, this is expressed by the use of $P_T(x|g)$ and $P_L(g)$ in Equation (8b), indicating that the learner is simulating themselves (the $P_L(g)$ term) as a teacher (the $P_T(x|g)$ term) who reasons about themselves as a learner. In contrast, the EIG model is solely a function of the learner's prior and observation likelihood. Second, the reasoning behind the two approaches is quite different. The Self-Teaching model inherits reasoning about individual hypotheses from the teaching model (see the end of Section 2.3.1). This resembles a kind of hypothesis testing (see the end of Section 2.3.4), where interventions are chosen based on their expected relative distinctiveness with respect to individual hypotheses. In contrast, the EIG model is based on entropy, which is a quantity derived from averaging over hypotheses. Thus, the Self-Teaching model reasons more on the level of individual hypothesis, whereas the EIG model reasons more on the level of global changes in the beliefs about hypotheses. Third, mathematically, the Self-Teaching model involves only the rules of probability (sum, product, and Bayes' rules), whereas the EIG model requires the minus and log mathematical operators to formalize the idea of reduction and uncertainty. This difference suggests that the Self-Teaching model cannot be readily reduced to any of the uncertainty-reduction models (e.g., those covered by Crupi et al., 2018). Despite these differences, the two intervention selection behaviors of these two models can be quite similar. In the next section, we characterize the similarities and differences through simulation.

## 3 Simulations

### 3.1 Tasks

We illustrate the Self-Teaching model with three tasks: two variants of a concept learning task and a causal learning task. In each of these tasks, we compare and contrast the predictions of the EIG model to our Self-Teaching model.

**3.1.1 Concept Learning.** The first two tasks we consider are concept learning tasks, where the set of possible hypotheses are various line segments on an interval as illustrated in Figure 3(a) and Figure 3(b). The goal of the learner is to discover the true hypothesis by querying features from different positions along the line. We consider two variants of the concept learning task which we call the *boundary task* and the *line task*.

In the boundary task, the hypothesis space consists of all of the unique concepts that are defined by the location of their boundary (e.g. Markant & Gureckis, 2014). Intuitively, a unique hypothesis has been identified when the learner observes two neighboring positions with different labels. In the line task, the hypothesis space consists of all possible lines with $n$ features (Tenenbaum, 1999; Shafto & Goodman, 2008). While the set of concepts are similar in these two tasks, the value of selecting different features differs in either task due to the set of other hypotheses under consideration. For both concept learning tasks, the learner can only ask for one feature and observe one label at a time.

(a) Boundary task          (b) Line task          (c) Causal learning task
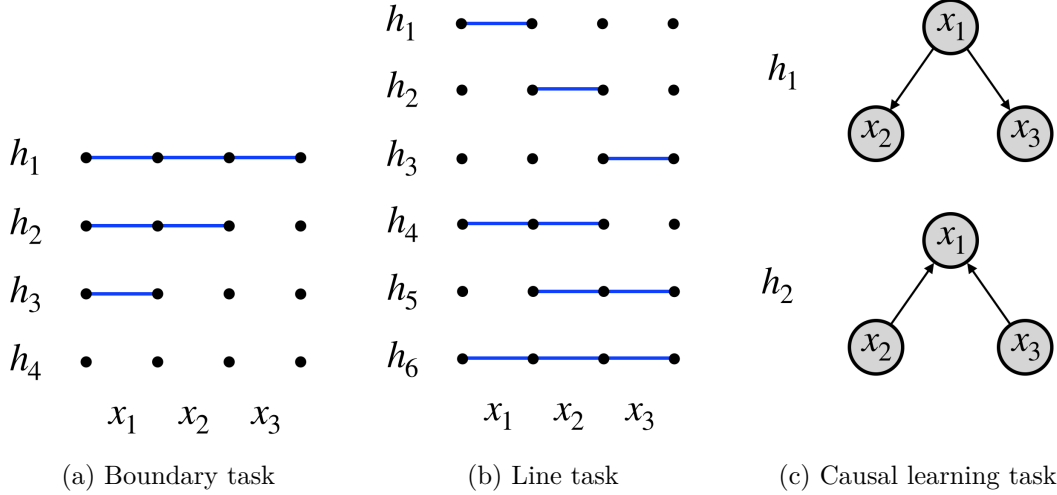
*Figure 3*. **Examples of the three tasks to compare the expected-information-gain and Self-Teaching models**. We explore two kinds of concept learning which we call the *boundary task* and the *line task*, which both involve concepts of varying line segments on a one-dimensional line. In both concept learning tasks, the set of features $x_i$ refer to the different positions along the line. In the boundary game, the set of hypotheses is restricted to lines where there is a single boundary to disambiguate between hypotheses. On the other hand, the line task includes all possible single line concepts for a given number of features. In the *causal learning task*, the hypotheses $h$ are two different causal graphs. The set of features $x_i$ in this task refer to which node the learner chooses to intervene on.

We use the following notation for describing the concept learning tasks. First, each hypothesis $h$ in the concept learning task consists of a set of features $\mathcal{X} = \{x_i | x_i = i, i = 1, \ldots, n\}$ and labels $\{y_i | y_i \in \{0, 1\}, i \in \{1, \ldots, n\}\}$, where each feature, $x_i$, refers to a position along the line segment, and each label, $y_i$, refers to whether the line segment for a given hypothesis exists at that position or not. Thus, each hypothesis can be described as a set of data points $h = \{(x_i, y_i)\}_{i=1}^n$. For this tasks, the likelihood $P(y|x, h)$ is given by weak sampling, and is equal to 1 when $x = x_i, y = y_i$ and 0 when $x = x_i, y \neq y_i$ for a given $h$.

**3.1.2 Causal Learning.** The third task we explore is a causal learning problem. In this task, the learner is presented with a causal system where the connections between the nodes are not shown, and the learner must determine which of two possible causal graphs is the true causal graph. Each causal graph consists of three nodes and the nodes are connected in different configurations (see Figure 3(c) for one example). The learner can choose to intervene on one of the three nodes, which will turn that particular node on, and will also cause other nodes that are connected downstream of the intervened node to turn on with high probability.[6] Depending on the true underlying graph, intervening on

---

[6]For our simulations, the probability that an active parent node would turn on its direct descendants was set to be 0.8, and the background rate of a node turning on spontaneously was set to be 0.0. These settings

different nodes produces different patterns of nodes turning on, which can help the learner determine which of the two graphs is the true hypothesis. We consider 27 possible pairs of causal graphs, matching the set used in Coenen et al. (2015), to compare the predictions of the EIG and Self-Teaching models for this task.

The notation for the causal learning tasks is similar to the concept learning tasks, albeit with a few minor differences. The hypotheses $h$ in the causal learning tasks refer to the different graphs, and each pair of graphs form a hypothesis space $\mathcal{H}$. The features in this task, $x$, refer to the three nodes the learner can intervene on, thus $x \in \{x_1, x_2, x_3\}$. The corresponding label, $y$, refer to the state of the causal graph after intervening on feature $x_i$. Thus, if the learner intervened on $x_1$, and this caused both $x_2$ and $x_3$ to turn on, then the corresponding label would be $y = \{\texttt{intervened\_on}, \texttt{observed\_on}, \texttt{observed\_on}\}$. The set of possible labels for the causal learning task are the four possible sets of observations over two of the nodes when intervening on the third node, leading to twelve possible values of $y$. For this task, the likelihood $P(y|x, h)$ is the likelihood of observing a particular combination of states for all three nodes, as given by the joint probability over the directed graphical model when intervening on $x$.

## 3.2 Results

In this section, we explore via simulation how the predictions of our proposed Self-Teaching model fares when compared to EIG model.[7] In particular, we explore similarities and differences between these models in their choice of particular features to intervene upon.[8]

For the boundary task, we ran simulations using both three features and eight features as shown in Figure 4. The results show that the EIG model and Self-Teaching model both favor selecting the central feature, rather than features at the ends. This is true both in the three feature case (Figure 4(a)) where the choice favored by either model is the central feature, and the eight feature case (Figure 4(b)), where both models show a preference for the two middle features. [9]

To further explore the behavior of both models in the boundary task, we conducted an additional set of simulations in the three feature case, considering how the EIG model and Self-Teaching model would determine which feature to select after observing a single data point. The plots showing the four different possibilities are shown in Figure 5.[10] For the four sets of observations, in each case the results show that the EIG model and Self-Teaching model both favor the same sets of features.

---

were chosen to match the settings used in the simulations reported by Coenen, Rehder, and Gureckis (2015).

[7]For all of the simulations, the EIG values were normalized so that they sum up to 1.

[8]The source code to run the simulations and produce the Figures in the Results section can be found at https://github.com/CoDaS-Lab/self-teaching.

[9]In the active learning literature, the expected information gains usually pass through a softmax decision function to yield the probabilities of intervention selection. Thus, the ordering of the preferences, rather than the values of the information gain, is the more meaningful quantity. The same is true with Self-Teaching.

[10]We do not show the two sets of data points which end up discovering the true hypothesis from a single observation. After the learner observes $x = 1, y = 1$, or $x = 3, y = 0$, the only consistent hypothesis remaining is $h_1$ or $h_4$ respectively as depicted in Figure 3(a), and thus there is no additional benefit to exploring the predictions of either model after having observed either of these two cases.
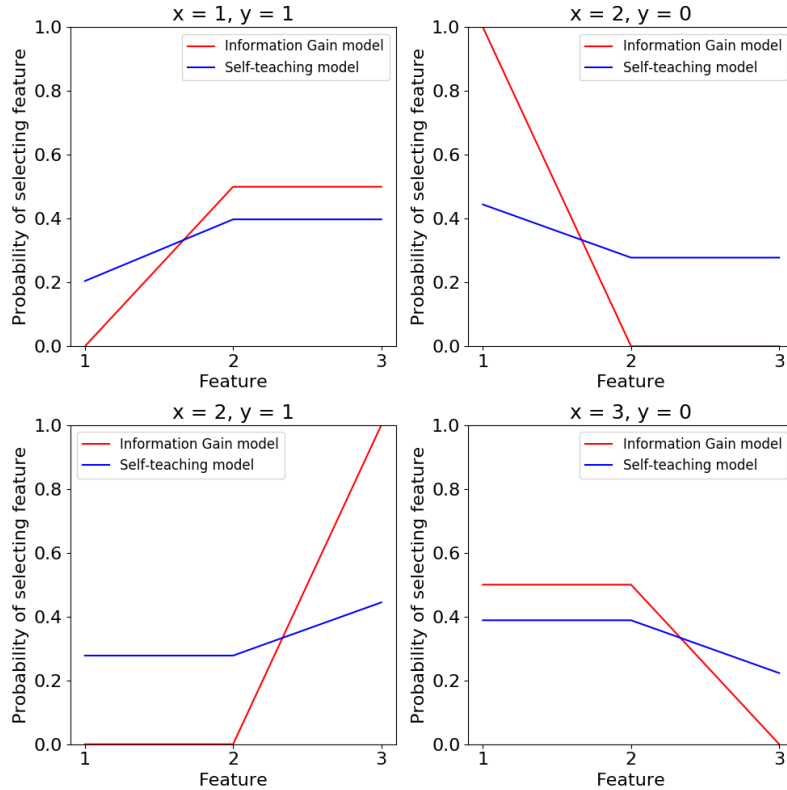
(a) Three features    (b) Eight features
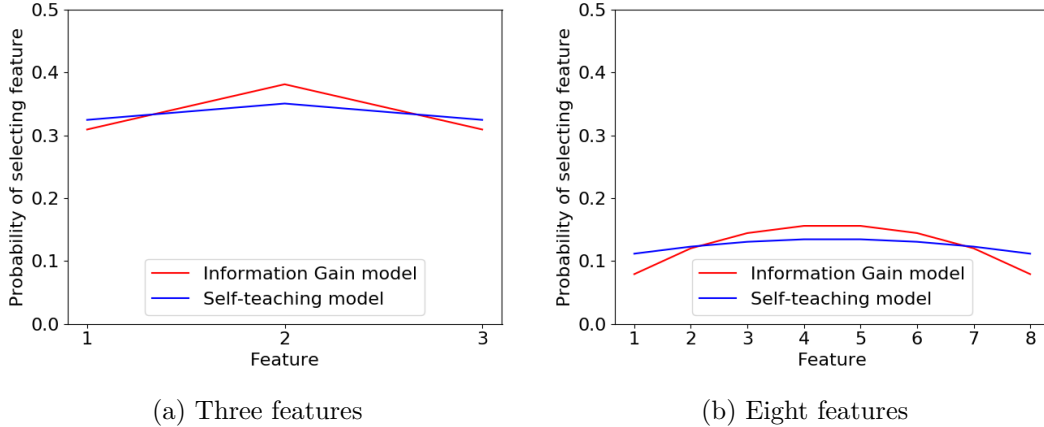
*Figure 4*. **Probability of selecting each feature for the boundary task with three and eight features**. The horizontal axes represent the feature position along the line for each concept. The vertical axis represents the probability of selecting each feature. The predictions of the Self-Teaching model are shown in blue, while the EIG model is in red. Both the EIG and Self-Teaching models favor the selection of features in the middle for the boundary task for both three and eight features.

To further support the conclusion that the Self-Teaching model accords with the behavior of the EIG model, we consider predictions on a related, but different concept learning task. Results for the line task are shown in Figure 6, where we considered hypothesis spaces that again consisted of either three or eight features. The results from these two simulations show different, yet interesting patterns of predictions that are qualitatively different from the boundary task. First, in the three feature case as shown in Figure 6(a), both the Self-Teaching and EIG models show a preference for the features on the ends, rather than in the center like the boundary task results in Figure 4. This suggests that Self-Teaching and EIG exhibit a similar sensitivity to the set of hypotheses and assign similar values to the features. Indeed, examination of the probabilities of selecting each feature in a line task with eight features in Figure 6(b) reveals interesting differences compared to the boundary task with eight features in Figure 4(b). Here, we find that the EIG model peaks in two points away from the center, suggesting selecting the features at those points would provide more information than the center. Similarly, we also see that Self-Teaching predicts the same qualitative behavior with two peaks.

So far, our results show that the qualitative behavior of Self-Teaching matches the EIG model in two different concept learning tasks. However, is this always the case? Or are there situations where the predictions of the Self-Teaching model differ from the EIG model?

To explore this possibility, we also compared the EIG and Self-Teaching models in a causal learning setting. We examined the predictions of both models across the 27 problems used in Coenen et al. (2015), as they provide a challenging test-bed of active learning problems within the domain of causal learning. In particular, neither of the two models (EIG and Positive-Test Strategy) considered in the paper alone explains the human data particularly well, and we were interested in how the predictions of the Self-Teaching model
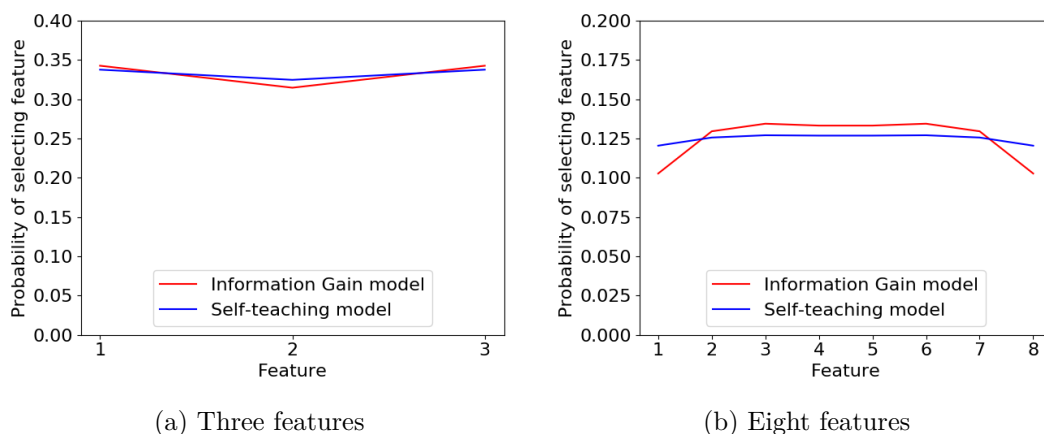
*Figure 5*. **Probability of selecting each feature for the boundary task with three features after observing a single data point**. The title of each figure is the first data point observed by the model. Two conditions are omitted, as observing either of these cases would lead the learner to infer the true hypothesis after a single step. After a single observation, both the Self-Teaching model (blue) and the EIG model (red) show the same qualitative behavior in determining which feature to select next.

would compare to the EIG and Positive-Test Strategy models in this domain.

In this task, participants were presented with an unknown causal system (by showing three nodes, but not the particular edges connecting the three nodes), and were asked to figure out how the causal system worked. Participants were presented with two possible hypotheses to explain the causal system (by showing two different sets of connecting edges between the nodes). They were asked to choose one of the three nodes to intervene on and observe how the system operated to determine which of the two possible hypotheses was the true underlying causal system. For each problem, the proportion that each node was selected across all participants was noted by Coenen et al. (2015).

The behavioral data collected from Coenen et al. (2015) was then compared to two different active learning models: the EIG model which we have described earlier, and a Positive-Test Strategy model (PTS), which is reminiscent of confirmation bias strategies for active learning. In this particular task, the behavior of the PTS model favors nodes which turn on as many other nodes as possible, by calculating the following:

(a) Three features                    (b) Eight features

*Figure 6*. **Probability of selecting each feature in the line task with varying numbers of features features.** The horizontal axes represent different features for each concept. The vertical axis represents the probability of selecting the feature. Here, the qualitative pattern of predictions differs from the boundary game for both the three feature case and the eight feature case. In the three feature case, both the Self-Teaching model (blue) and the EIG model (red) have a preference for choosing the features on the ends, rather than the central feature like the boundary task. Additionally, for the eight feature case, instead of a single peak at the center, both the Self-Teaching model (blue) and the EIG model (red) have a preference at two particular peaks away from the center.

$$\text{PTS}_x = \max_x \left[ \frac{\text{DescendantLinks}_{x,h}}{\text{TotalLinks}_h} \right]. \tag{9}$$

Overall, the results from Coenen et al. (2015) showed that while both models captured participant behavior well in some of the problems, neither of the two active learning models alone were able to capture participant's responses across all of the of problems.

We computed the predictions for the Self-Teaching model (along with the EIG and PTS models) across the set of 27 causal learning problems, and results for these simulations are shown in Figure 7.[11] For each problem, the two causal graphs under consideration for each problem are shown in the top right of each plot and the predictions from each of the three models are plotted in different colors and symbols. We say that the behavior of the Self-Teaching model differs from the EIG model if the predictions diverge in a notice-able manner, and there are two kinds of situations where this occurs. First, is when the predictions of the EIG model favors selecting a single feature with high probability such that the point lies on one of the vertices (Problems 8, 13 and 25). In these instances, the Self-Teaching model favors selecting the same feature as predicted by the EIG model (as both points lie within the same sector). The second case is when Self-Teaching favors se-lecting a different feature than the EIG model as evidenced by the predictions from the two

---

[11]We requested the data from the authors but it was not made available, thus our focus is to quantita-tively compare Self-Teaching to the Expected Information Gain and Positive-Test Strategy models of active learning.
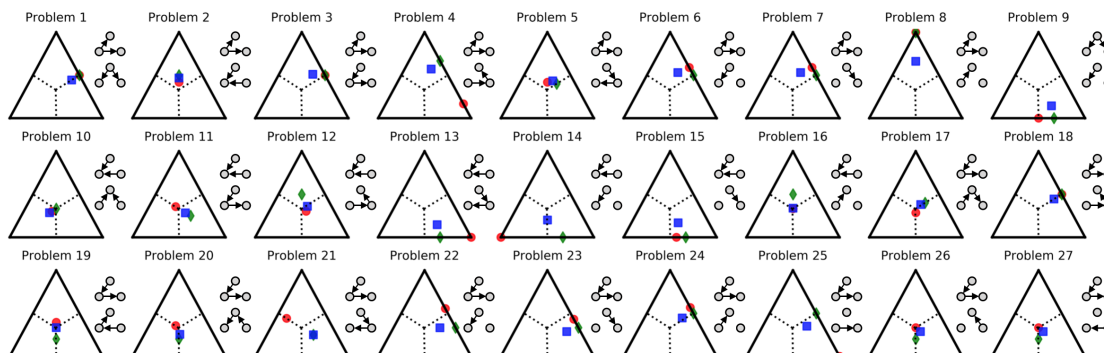
*Figure 7*. **The predictions of the Self-Teaching model (blue square), EIG model (red circle) and Positive-Test Strategy model (green diamond) for the different causal learning problems.** Each of the diagrams represents a probability simplex over the three nodes to intervene upon, along with the two hypotheses to be considered in the top right of each diagram. A point in the center of the simplex indicates indifference towards selecting any of the three nodes, while points that are closer to each corner represent a stronger preference for selecting one node over the others. The dotted lines indicate the boundaries for favoring one feature over the other two, partitioning the simplex into three distinct sectors. Our results show that for most of the 27 problems, the predictions of the Self-Teaching model for which node to intervene matches the EIG model (see text for more details). However, in a number of cases, the Self-Teaching model prefers intervening on a different node. We compared this to another model of active learning (Positive-Test Strategy) and found that this made similar predictions in these non-matching cases, suggesting that self-teaching mimics the behavior of both models under different circumstances.

models lying in two different sectors (Problems 4, 14, 21 and 22). In the latter cases where Self-Teaching prefers a different feature than the EIG model, the predictions of the PTS model also prefer the same feature that Self-Teaching does (except for Problem 14). Apart from these few cases, we consider the predictions of the EIG model and the Self-Teaching model to be matching, and find that there is a strong amount of similarity between these two active learning models.

## 4 Discussion

People can learn from both active exploration and from other people. Yet, research on active learning and teaching has largely been conducted independently, and to our knowledge there exist no unified framework of these basic cognitive phenomena. These models differ in terms of both the learning objective and learning architecture, making unification an interesting and challenging problem. We have presented an integration by formalizing active learning as *self-teaching.* In our approach, active learning by self-teaching determines which features to select by imagining teaching oneself and averaging across all possible outcomes and hypotheses. Mathematically, this corresponds to adding two simple marginal-

ization steps to a previously proposed teaching model, with the resulting model being a formalization of active learning as self-teaching.

We compare our Self-Teaching model to a representative model from the active learning literature, Expected Information Gain (EIG). Mathematically, the models yield quite different selection functions (see Sections 2.2, 2.3.4, and 2.4). Indeed, the differences are substantial enough that direct mathematical comparison is non-trivial. Inspection of the Self-Teaching model, however, shows that it seeks data that, in expectation, yield changes in beliefs about specific, distinctive hypotheses relative to random sampling. This expected relative distinctiveness differs from expected information gain in focusing on 1) specific hypotheses rather than the entire distribution and 2) hypotheses that are relatively distinctive in the sense that one would be unlikely to learn about through by random sampling rather than the divergence between prior to posterior.

A well-known challenge for existing models of active learning is that there exist relatively distinct sets of empirically observed phenomena. On one hand, there exists a robust literature that demonstrates that people's selection of interventions is best explained by models that minimize global uncertainty or maximize global information (Nelson, 2005; Crupi et al., 2018). Second, others have argued that learners minimize entropy with respect to manageable subsets of the full hypothesis space (Markant et al., 2016; Bramley et al., 2017). A third approach has characterized active learning as a form of hypothesis testing with a bias towards confirmation (Wason, 1960; Klayman & Ha, 1987; Navarro & Perfors, 2011). The EIG model can explain the first two approaches to active learning (Nelson, 2005; Markant & Gureckis, 2014), but not the latter on hypothesis testing. In contrast, the Self-Teaching model can qualitatively capture aspects of these different empirical phenomena through expected relative distinctiveness. Our results show close correspondence with the EIG model on the boundary and line tasks (Figures 4–6), while in the causal learning task it displays behavior that is close to a mixture of the EIG and positive-test strategy models (Figure 7). This suggests that self-teaching can capture different forms of active learning that previously required different mathematical formalizations, and in a manner that is consistent with formal theories of teaching.

Unifying active learning and teaching as meta-reasoning has potentially far-reaching implications. A model that unifies the two would allow interpolation between these two extremes by manipulating the teacher's knowledgeability. These interpolations allow direct comparisons between teaching from teachers of different knowledge levels, active learning, guided learning (Weisberg, Hirsh-Pasek, & Golinkoff, 2013), and collaborative learning between peers (Dillenbourg, 1999). More generally, unifying models of teaching and active learning may provide a first step towards unifying models of cognition about others (i.e., social cognition; (Shafto, Eaves, et al., 2012; Eaves & Shafto, 2012; Baker, Saxe, & Tenenbaum, 2011) and cognition about oneself (i.e., meta-cognition) in a single meta-reasoning architecture. Research in psychology and neural science has long speculated about a close link between social cognition and meta-cognition including the possibility of common developmental trajectory and neural substrates (Meltzoff, 2005). Work remains to explore how the current framework relates to more varieties of active learning models in the literature (e.g. Austerweil & Griffiths, 2011; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003) and how much of human behavior it captures. Overall, based on the simulation results, there is reason to be optimistic that the different models and model behaviors can be tied together

to yield an integrated account of human learning from people and the world.

## Acknowledgements

## References

Austerweil, J. L., & Griffiths, T. L. (2011). A rational model of the effects of distributional information on feature learning. *Cognitive Psychology*, *63*(4), 173–209.

Baker, C., Saxe, R., & Tenenbaum, J. (2011). Bayesian theory of mind: Modeling joint belief-desire attribution. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society.*

Bramley, N. R., Dayan, P., Griffiths, T. L., & Lagnado, D. A. (2017). Formalizing Neurath's ship: Approximate algorithms for online causal learning. *Psychological Review*, *124*(3), 301–338.

Bramley, N. R., Lagnado, D. A., & Speekenbrink, M. (2015). Conservative forgetful scholars: How people learn causal structure through sequences of interventions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 708–731.

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, *31*, 21–32.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). A study of thinking. *New York*.

Coenen, A., Nelson, J. D., & Gureckis, T. (2017). Asking the right questions about human inquiry. *PsyArXiv*. doi: 10.31234/osf.io/h457v

Coenen, A., Rehder, B., & Gureckis, T. M. (2015). Strategies to intervene on causal systems are adaptively selected. *Cognitive Psychology*, *79*, 102–133.

Crupi, V., Nelson, J. D., Meder, B., Cevolani, G., & Tentori, K. (2018). Generalized information theory meets human cognition: Introducing a unified framework to model uncertainty and information search. *Cognitive Science*, 1–47.

Dillenbourg, P. (1999). *Collaborative learning: Cognitive and computational approaches*. ERIC.

Eaves, B., & Shafto, P. (2012). Unifying pedagogical reasoning and epistemic trust. *Advances in Child Development and Behavior*, *43*, 295–319.

Frank, M. (2014). Modeling the dynamics of classroom education using teaching games. In *Proceedings of 36th Annual Conference of the Cognitive Science Society*.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, *15*(10), 661–667.

Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*(2), 211–228.

MacKay, D. J. (1992). Information-based objective functions for active data selection. *Neural Computation*, *4*(4), 590–604.

Markant, D. B., & Gureckis, T. M. (2014). Is it better to select or to receive? Learning via active and passive hypothesis testing. *Journal of Experimental Psychology*, *143*(1), 94–122.

Markant, D. B., Settles, B., & Gureckis, T. M. (2016). Self-directed learning favors local, rather than global, uncertainty. *Cognitive Science*, *40*(1), 100–120.

Meltzoff, A. N. (2005). Imitation and other minds: The "like me" hypothesis. *Perspectives on Imitation: From Neuroscience to Social Science*, *2*, 55–77.

Navarro, D. J., & Perfors, A. F. (2011). Hypothesis generation, sparse categories, and the positive test strategy. *Psychological Review*, *118*(1), 120–134.

Nelson, J. D. (2005). Finding useful questions: on bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, *112*(4), 979–999.

Shafto, P., Eaves, B., Navarro, D. J., & Perfors, A. (2012). Epistemic trust: Modeling children's reasoning about others' knowledge and intent. *Developmental Science*, *15*(3), 436–447.

Shafto, P., & Goodman, N. (2008). Teaching games: Statistical sampling assumptions for learning in pedagogical situations. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1632–1637).

Shafto, P., Goodman, N. D., & Frank, M. C. (2012). Learning from others: The consequences of psychological reasoning for human learning. *Perspectives on Psychological Science*, *7*(4), 341–351.

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.

Steffe, L. P., & Gale, J. E. (1995). *Constructivism in Education.* Lawrence Erlbaum Hillsdale.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E.-J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*(3), 453–489.

Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning* (Vol. 135). MIT press Cambridge.

Tenenbaum, J. B. (1999). Bayesian modeling of human concept learning. In *Advances in Neural Information Processing Systems* (pp. 59–65). MIT Press.

Tobias, S., & Duffy, T. M. (2009). *Constructivist instruction: Success or failure?* Routledge.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, *12*(3), 129–140.

Weisberg, D. S., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Guided play: Where curricular goals meet a playful pedagogy. *Mind, Brain, and Education*, *7*(2), 104–112.

Yang, S. C.-H., Lengyel, M., & Wolpert, D. M. (2016). Active sensing in the categorization of visual patterns. *Elife*, *5*, 1–22.

Yang, S. C.-H., & Shafto, P. (2017). Teaching versus active learning: A computational analysis of conditions that affect learning. In *Proceedings of the 39th Annual Conference of the Cognitive Science Society.*

## Appendix
### A worked example

Below, we show the calculations for both the Self-Teaching model and Expected Information Gain model for the probability of selecting each feature for the boundary task with three features as shown in Figure 3(a). In this task, there are four hypotheses and the **learner's prior** is set to be uniform for each hypothesis: $P(h) = 1/4$. The **observation likelihood**, $P(y|x, h)$, is shown in Table A1(a).

Both the expected-information-gain and self-teaching models require calculating the **learner's posterior** probabilities, which is given by $P_L(h|x, y) = \frac{P_L(y|x,h)P_L(h)}{\sum_{h' \in \mathcal{H}} P_L(y|x,h')P_L(h')}$. As a concrete example, consider the learner's posterior for $h_2$, after observing that $x = 1$ and $y = 1$:

$$P_L(h_2|x = 1, y = 1) = \frac{P_L(y = 1|x = 1, h_2)\, P_L(h_2)}{\sum_{h' \in \mathcal{H}} P_L(y = 1|x = 1, h')\, P_L(h')} = \frac{1 \cdot \frac{1}{3}}{0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} = \frac{1}{3}.$$

The learner's posterior probabilities for each hypothesis for every possible observation is recorded in Table A1(b).

Table A1

**(a).** *The observation likelihood, $P(y|x,h)$. The last column of this table computes $\sum_{h'\in\mathcal{H}} P_L(y|x,h')$, which is useful for computing the posterior.* **(b).** *The learner's posterior probabilities, $P_L(h|x,y)$. Since the learner's prior is constant in h, this can be obtained by taking each element of the likelihood table and divide that by the sum of the corresponding row recorded in the last column. The last row of this table is the sum of each column, and the inverse of this sum is equivalent to the "distinctiveness" term.*

| $P(y\|x,h)$ | $h_1$ | $h_2$ | $h_3$ | $h_4$ | $\sum_h$ |
|---|---|---|---|---|---|
| $x=1, y=0$ | 0 | 0 | 0 | 1 | 1 |
| $x=1, y=1$ | 1 | 1 | 1 | 0 | 3 |
| $x=2, y=0$ | 0 | 0 | 1 | 1 | 2 |
| $x=2, y=1$ | 1 | 1 | 0 | 0 | 2 |
| $x=3, y=0$ | 0 | 1 | 1 | 1 | 3 |
| $x=3, y=1$ | 1 | 0 | 0 | 0 | 1 |

(a) *The observation likelihood*

| $P_L(h\|x,y)$ | $h_1$ | $h_2$ | $h_3$ | $h_4$ |
|---|---|---|---|---|
| $x=1, y=0$ | 0 | 0 | 0 | 1 |
| $x=1, y=1$ | $^1/_3$ | $^1/_3$ | $^1/_3$ | 0 |
| $x=2, y=0$ | 0 | 0 | $^1/_2$ | $^1/_2$ |
| $x=2, y=1$ | $^1/_2$ | $^1/_2$ | 0 | 0 |
| $x=3, y=0$ | 0 | $^1/_3$ | $^1/_3$ | $^1/_3$ |
| $x=3, y=1$ | 1 | 0 | 0 | 0 |
| $\sum_x \sum_y$ | $^{11}/_6$ | $^7/_6$ | $^7/_6$ | $^{11}/_6$ |

(b) *The learner's posterior*

### Self-teaching

To compute the Self-Teaching probabilities, we first compute the **Teaching-by-Intervention probabilities** according to Equation (5b): $P_T(x,y|h) = \frac{P_L(h|x,y)P_T(x,y)}{\sum_{x'\in\mathcal{X}}\sum_{y'\in\mathcal{Y}} P_L(h|x',y')P_T(x',y')}$. Since there are 6 combinations of $(x,y)$, we set the **teacher's selection prior** to be uniform, $P_T(x,y) = {}^1/_6$ for all $x,y$ pairs. Note that this $P_T(x,y)$ always cancels out, but we still show it for clarity. Now, we can use the posterior probabilities in Table A1(b) to compute the Teaching-by-Intervention probabilities. Continuing our example, suppose the true hypothesis is $h_2$, and the teacher wants to demonstrate the data point $x=1, y=1$:

$$P_T(x=1, y=1|h_2) = \frac{P_L(h_2|x=1, y=1)\, P_T(x=1, y=1)}{\sum_{x'\in\mathcal{X}}\sum_{y'\in\mathcal{Y}} P_L(h_2|x',y')\, P_T(x',y')}$$

$$= \frac{\frac{1}{3}\cdot\frac{1}{6}}{0\cdot\frac{1}{6} + \frac{1}{3}\cdot\frac{1}{6} + 0\cdot\frac{1}{6} + \frac{1}{2}\cdot\frac{1}{6} + \frac{1}{3}\cdot\frac{1}{6} + 0\cdot\frac{1}{6}} = \frac{^1/_3}{^7/_6} = \frac{2}{7}.$$

All the demonstration probabilities are recorded in Table A2(a).[12] Next, we compute the **Teaching-by-Instruction probabilities** according to Equation (6b) by marginalizing the teacher's demonstration distribution over $y$: $P_T(x|h) = \sum_{y\in\mathcal{Y}} P_T(x,y|h)$. Again, continuing our example and assuming $h_2$ is the true hypothesis, and the teacher wants to instruct the learner to select $x=1$:

$$P_T(x=1|h_2) = \sum_{y\in\mathcal{Y}} P_T(x=1, y|h_2) = \frac{2}{7} + 0 = \frac{2}{7}.$$

---

[12]These probabilities are not the probabilities at convergence but are the teaching probabilities in the first iteration, i.e., $P_{T_0}(x,y|h)$ as described after Equation (5).

Table A2

*Teaching and Self-teaching.  **(a).**  The teacher's intervention probabilities.  Since the teacher's selection prior is uniform, these can be computed by taking each posterior probabilities in Table A1(b) and divide that by the corresponding value in the last row of the same table.  **(b).**  The teacher's instruction probabilities, obtained from summing over the y of the intervention probabilities.  **(c).**  The Self-Teaching probabilities, obtained from summing over the hypotheses of the instruction probabilities then multiply by the learner's prior. $P_T(x=2)$ is the highest.*

| $P_T(x,y\|h)$ | $h_1$ | $h_2$ | $h_3$ | $h_4$ |
|---|---|---|---|---|
| $x=1, y=0$ | 0 | 0 | 0 | $6/11$ |
| $x=1, y=1$ | $2/11$ | $2/7$ | $2/7$ | 0 |
| $x=2, y=0$ | 0 | 0 | $3/7$ | $3/11$ |
| $x=2, y=1$ | $3/11$ | $3/7$ | 0 | 0 |
| $x=3, y=0$ | 0 | $2/7$ | $2/7$ | $2/11$ |
| $x=3, y=1$ | $6/11$ | 0 | 0 | 0 |

(a) *Teaching-by-Intervention*

| $P_T(x\|h)$ | $h_1$ | $h_2$ | $h_3$ | $h_4$ |
|---|---|---|---|---|
| $x=1$ | $2/11$ | $2/7$ | $2/7$ | $6/11$ |
| $x=2$ | $3/11$ | $3/7$ | $3/7$ | $3/11$ |
| $x=3$ | $6/11$ | $2/7$ | $2/7$ | $2/11$ |

(b) *Teaching-by-Instruction*

| | $P_T(x)$ |
|---|---|
| $x=1$ | $25/77$ |
| $x=2$ | $27/77$ |
| $x=3$ | $25/77$ |

(c) *Self-teaching*

All the Teaching-by-Instruction probabilities are recorded in Table A2(b).  Finally, we can compute the **Self-Teaching probabilities** according to Equation (8b) by marginalizing the Teaching-by-Instruction distribution over all possible hypotheses $g$: $P_T(x) = \sum_{g \in \mathcal{H}} P_T(x,|g)\,P_L(g)$.  Finally, using our worked example, we can show that the self-teacher will intervene $x = 1$ with probability given by:

$$P_T(x=1) = \sum_{g \in \mathcal{H}} P_T(x=1|g)\,P_L(g)$$
$$= \frac{2}{11} \cdot \frac{1}{4} + \frac{2}{7} \cdot \frac{1}{4} + \frac{2}{7} \cdot \frac{1}{4} + \frac{6}{11} \cdot \frac{1}{4}$$
$$= \frac{14 + 22 + 22 + 42}{77} \cdot \frac{1}{4} = \frac{25}{77}.$$

The Self-Teaching probabilities are recorded in Table A2(c), which matches the results for the self-teaching model shown in Figure 4(a).

**Expected information gain.**

To compute the expected information gain in Equation (4), we first compute the the **prior entropy** according to Equation (1):

$$H(h) = \sum_{h \in \mathcal{H}} P_L(h) \log \frac{1}{P_L(h)} = 4 \times \left( \frac{1}{4} \log \frac{1}{1/4} \right) = \log 4.$$

Table A3

*Calculating the predictive probability, posterior entropy and expected information gain*

| | $P_L(y\|x)$ | $H(h\|x,y)$ |
|---|---|---|
| $x=1, y=0$ | $\frac{1}{4}$ | $0$ |
| $x=1, y=1$ | $\frac{3}{4}$ | $\log 3$ |
| $x=2, y=0$ | $\frac{1}{2}$ | $\log 2$ |
| $x=2, y=1$ | $\frac{1}{2}$ | $\log 2$ |
| $x=3, y=0$ | $\frac{3}{4}$ | $\log 3$ |
| $x=3, y=1$ | $\frac{1}{4}$ | $0$ |

(a) *The predictive probability and posterior entropy for all observations*

| | $EIG(x)$ |
|---|---|
| $x=1$ | $\log 4 - \frac{3}{4}\log 3$ |
| $x=2$ | $\log 4 - \log 2$ |
| $x=3$ | $\log 4 - \frac{3}{4}\log 3$ |

(b) *The expected information gain for each intervention*

Then, we compute the **posterior entropy** according to Equation (2). For example, the posterior entropy after observing $x = 1, y = 1$ is calculated as follows:

$$H(h|x=1, y=1) = \sum_{h \in \mathcal{H}} P_L(h|x=1, y=1) \log \frac{1}{P_L(h|x=1, y=1)}$$

$$= 0 \log 0 + \frac{1}{3}\log 3 + \frac{1}{3}\log 3 + \frac{1}{3}\log 3 = \log 3.$$

Then, we need to compute the **predictive probability**, $P_L(y|x) = \sum_{h \in \mathcal{H}} P_L(y|x, h)\, P_L(h)$. Continuing our example, the predictive probability of $y = 1$, given $x = 1$ is given by:

$$P_L(y=1|x=1) = \sum_{h \in \mathcal{H}} P_L(y|x, h)\, P_L(h) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} = \frac{3}{4}.$$

The predictive probabilities and posterior entropy values are recorded in the first and second column of Table A3(a), respectively. Finally, using the prior entropy calculated and the values in Table A3(a), we can obtain the **expected information gain** according to Equation (4). Finally, we can calculate the expected information gain from intervening on $x = 1$ like so:

$$EIG(x=1) = H(h) - \sum_{y \in \mathcal{Y}} P(y|x=1)\, H(h|x=1, y)$$

$$= \log 4 - \left(\frac{3}{4}\log 3 + \frac{1}{4}0\right) = \log 4 - \frac{3}{4}\log 3.$$

All *EIG* values are recorded in Table A3(b).