Supplementary Materials for

Factive mindreading reflects the optimal use of limited cognitive resources

Proceedings of the Royal Society B

doi: 10.1098/rspb

Tadeg Quillien*, Max Taylor-Davies

*Corresponding author. Email: tadeg.quillien@gmail.com

Contents

1 Modeling framework

1.1 Generative model

Recall that the *actor* must choose to go toward one box, and the *observer*'s task is to predict the actor's behavior. In the spirit of recent computational approaches to mindreading (*1, 2, 3*), we model the actor as an approximately rational utility-maximizer. That is, the actor assigns expected utility to possible actions, and chooses an action y as a stochastic function of these expected utilities using a softmax choice rule:

$$P(Y = y) \propto \exp(\tau^{-1} U(y)) \tag{S1}$$

where $U(y)$ is the expected utility of y, and τ is a temperature parameter regulating the stochasticity of action selection. Throughout we use $\tau = 1/3$ and $r = 1$, where r is the reward for obtaining the item.[1] With these values, the actor chooses the box where it last saw the item about 90% of the time when $A = 1$ (the exact value depends on N).

We also assume that the actor holds a uniform prior over $Pr(S = s)$ (this matches the actual environment dynamics), but is not aware of the possibility that it could be deceived (effectively holding $Pr(D) = 0$).[2]

Suppose $A = 0$. Then the actor has no reason to update its prior, so it assigns identical expected utility to all actions. As a result the softmax choice rule outputs a uniform probability over actions. Suppose now that $A = 1$. Then the actor assigns expected utility r to $Y = b$, where r is the reward for obtaining the item, and 0 otherwise. So the softmax choice rule outputs a high probability of choosing b, and a low probability of choosing another box. For convenience in the $A = 1$ case we can write the probability of choosing box b as $Pr(Y = b) = 1 - \epsilon_N$, and the probability of choosing each other box as $Pr(Y = y_i) = \epsilon_N/(N - 1)$; where ϵ_N is a small value that can be derived from Equation S1.

[1] Exploratory analyses suggest that manipulating τ and r does not substantially alter results.

[2] Assuming that the actor has access to $Pr(D)$ would entail qualitatively similar conclusions, in particular even for very high values of $Pr(D)$ choosing $Y = b$ is still the best option, given that the switch that occurs when $D = 1$ is equally likely to put the item back in its initial location as any other box. The only difference would be that ϵ would now increase with $Pr(D)$, since an actor entertaining the possibility of deception would assign non-null expected utility to all possible actions. We chose to assume that the actor doesn't have access to $Pr(D)$ so that the effects of $Pr(D)$ on factivity that we document cannot be attributed to simple changes in the level of noise in the actor's decision.

Combining this result with the other characteristics of the task, we obtain the following generative model:

$$a \sim \text{Bernoulli}(\Pr(A)) \tag{S2}$$

$$d \sim \text{Bernoulli}(\Pr(D)) \tag{S3}$$

$$b \sim \text{Uniform}\{1, ..., N\} \tag{S4}$$

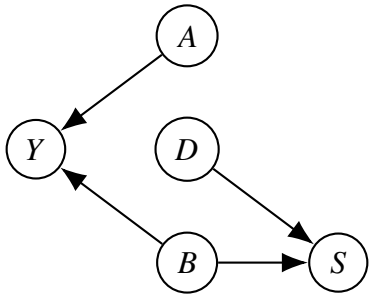
$$s \sim \begin{cases} b, & \text{if } d = 0 \\ \text{Uniform}\{1, ..., N\} & \text{if } d = 1 \end{cases} \tag{S5}$$

$$\Pr(Y = y) = \begin{cases} 1/N, & \text{if } a = 0 \\ 1 - \epsilon_N, & \text{if } a = 1, y = b \\ \epsilon_N/(N-1), & \text{if } a = 1, y \neq b \end{cases} \tag{S6}$$

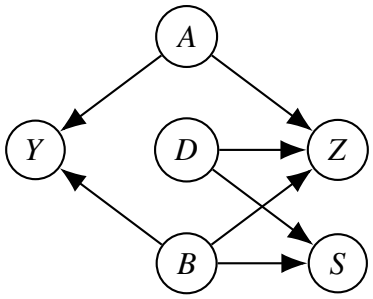
Note that s and d play no causal role in the actor's behavior. However, $s = b$ when $d = 0$, so s and d contain information about the actor's behavior which can be exploited by the observer.

1.2 Causal Bayesian Networks

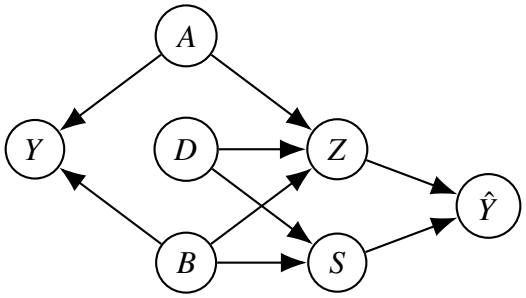
The information bottleneck framework requires working with joint probability distributions. Here the relevant joint distribution is the distribution induced by the Causal Bayesian Network defined by the task ((*4*), see also (*5*)). The generative model described above implies a Causal Bayesian Network with the following graph:



Combined with our assumption that the representation Z is a compressed encoding of A, B and D, we obtain a causal model characterized by the following graph:



Finally, the observer predicts the value of Y from Z and S; we denote this prediction $\hat{Y}$, and add it to our graph:



2 Information bottleneck

The information bottleneck framework was introduced in (*6*) as a way to formalize optimal lossy compression in a setting where some of the input data is more or less relevant to performing a task. We use a variant introduced by (*7*) that can incorporate side information (in our case the side information is S).

2.1 Relevant information-theoretic measures

The **Shannon entropy** of a variable V is defined as:

$$H(V) = -\sum_{v \in \mathcal{V}} \Pr(v) \log \Pr(v) \tag{S7}$$

Loosely speaking, the entropy of V measures our uncertainty over the value of V. In this paper we will use entropy to measure the maximum amount of information one can extract about a variable.

The **mutual information** between variables V and W is defined as:

$$I(V;W) = \sum_{v,w} \Pr(v,w) \log \frac{\Pr(v,w)}{\Pr(v)\Pr(w)} \tag{S8}$$

Intuitively $I(V;W)$ measures the distance between the joint distribution $Pr(V,W)$ and the product of the marginals $Pr(V)Pr(W)$, and can be seen as quantifying how much information each variable contains about the other. Note that mutual information is defined without assuming that V and W are single variables. For example the mutual information between W and the joint random variable (V,U) is given by:

$$I(W;V,U) = \sum_{w,v,u} \Pr(w,v,u) \log \frac{\Pr(w,v,u)}{\Pr(w)\Pr(v,u)} \tag{S9}$$

The **conditional mutual information** is defined as:

$$I(V;W|U) = I(W;V,U) - I(W;U) \tag{S10}$$

Finally, the **Kullback-Leibler divergence** of probability distribution Q from probability distribution P is defined as:

$$D_{\mathrm{KL}}(P||Q) = \sum_{v} P(v) \log \frac{P(v)}{Q(v)} \tag{S11}$$

For our purposes we can think of the KL divergence as a measure of the quality of Q as one's model of the distribution of V when the actual probability distribution is P.

2.2 Information bottleneck in our social prediction task

We have $Pr(\vec{X}) = Pr(A, D, B)$. The constrained optimization problem in the main text is equivalent (*6*) to the problem of minimizing the following functional, where β is a Lagrange multiplier:

$$\mathcal{F}_\beta[q(z|\vec{x})] = I(\vec{X}; Z) - \beta I(Z; Y|S) \tag{S12}$$

There is not a one-to-one mapping between values of C (upper bound on information extracted from $\vec{X}$) and β, so we compute optimal policies for various values of β and find the policy with a value of $I(\vec{X}; Z)$ closest to C. The parameter β regulates the trade-off between maximizing predictive performance and minimizing cognitive cost; high values of β result in more complex policies.

This problem can be solved using a variant of the Blahut-Arimoto algorithm, by iterating the following update equations until convergence (*8, 9, 6, 7*):

$$\Pr(z|x) \propto \Pr(z) \exp(-\beta \sum_s \Pr(x|s) D_{\mathrm{KL}}[\Pr(y|x, s)||\Pr(y|z, s)]) \tag{S13}$$

$$\Pr(z) = \Sigma_x \Pr(z|x) \Pr(x) \tag{S14}$$

$$\Pr(y|z, s) = \Sigma_x \Pr(y|z, s, x) \Pr(x|z, s) \tag{S15}$$

To compute the relevant conditional probabilities, we use the conditional independence relations given by the graph in the previous section. We then assume that the observer decodes Z in an optimal way by applying Bayes' rule (following e.g. (*10*)), so that $q(\hat{Y}|Z, S) = \Pr(Y|Z, S)$. Exploring resource constraints on this decoding process is a possible direction for future research.

To generate the observer predictions in our experiments, we compute $\Pr(\hat{Y}|\vec{X}, S)$ by marginalizing over Z:

$$\Pr(\hat{Y}|\vec{X}, S) = \sum_Z q(\hat{Y}|Z, S) q(Z|\vec{X}) \tag{S16}$$

Figure S1 illustrates the information bottleneck frontier (*6*), i.e. the relationship between predictive performance $I(Y; Z|S)$ and observer's channel capacity $I(\vec{X}; Z)$, across the resource-rational policies derived for the social ecology $\Pr(A) = .2$, $\Pr(D) = .05$, $N = 3$. Points above the curve are unattainable: increasing predictive performance would require increasing the complexity of the policy.

2.3 Deriving resource-rational policies

To minimize the chance that the iterative algorithm described above converge on local minima, we use reverse deterministic annealing to compute the optimal policies (*11*, *10*). We compute the optimal policy for a high value of β, and then slowly decrease the value of β, each time initializing the policy with the solution found for the previous value of β. The very first policy used to initialize the process is a deterministic bijection from $\vec{X}$ to Z, along with the corresponding optimal decoder.

We use a schedule of $\beta = 2^x$ with x varying from 6 to 0 in .1 increments. Throughout we keep the cardinality of Z at $|Z| = 2 \times 2 \times N$, which is the value that in principle allows a perfect deterministic mapping from $\vec{X}$ to Z. The reverse annealing process is performed independently for each combination of parameters ($Pr(A)$, $Pr(D)$, N).

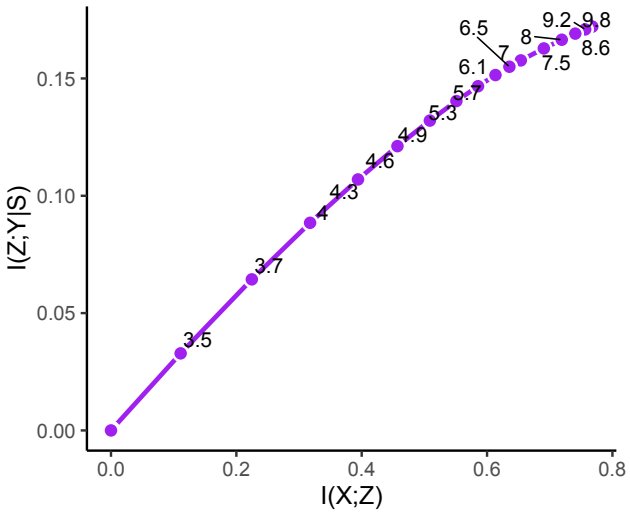


Figure S1: Information bottleneck frontier (here shown for $Pr(A) = .2$, $Pr(D) = .05$, $N = 3$). Each dot corresponds to a particular value of β.

3 Assessing the granularity of resource-rational representations

One key question is how many functionally different states of Z an observer represents. When deriving resource-rational encoders, we kept the number of possible states of Z as $|Z| = 2 \times 2 \times N$, which gives for example $|Z| = 8$ for $N = 2$. However, during optimization, different states of Z might end up playing an identical role, for example each functioning as a knowledge representation. We consider two states z_1 and z_2 to be functionally identical if $q(\vec{X}|z_1) = q(\vec{X}|z_2)$ and $q(\hat{Y}|z_1) = q(\hat{Y}|z_2)$, and we collapse all functionally identical states into the same state. Below we label these states with interpretable names like 'Ignorance' and 'Knowledge', but it should be kept in mind that the functional roles of these states emerged organically during optimization; they were not manually assigned at initialization.

We find that resource-rational policies typically have few states of Z: automatic policies have only a single state, while factive observers have two states (which we can interpret as an 'Ignorance' and a 'Knowledge' state). Observers with higher computational capacity typically have $2 + N$ different states (an Ignorance and a Knowledge state, as well as N different Belief states), see Figure S2. The sudden transition from knowledge to belief representation is most striking for high

values of N. For example for $N = 5$, increasing C causes a phase transition from $N = 2$ (Ignorance and Knowledge) to $N = 7$ (Ignorance and Knowledge plus five different Belief states).

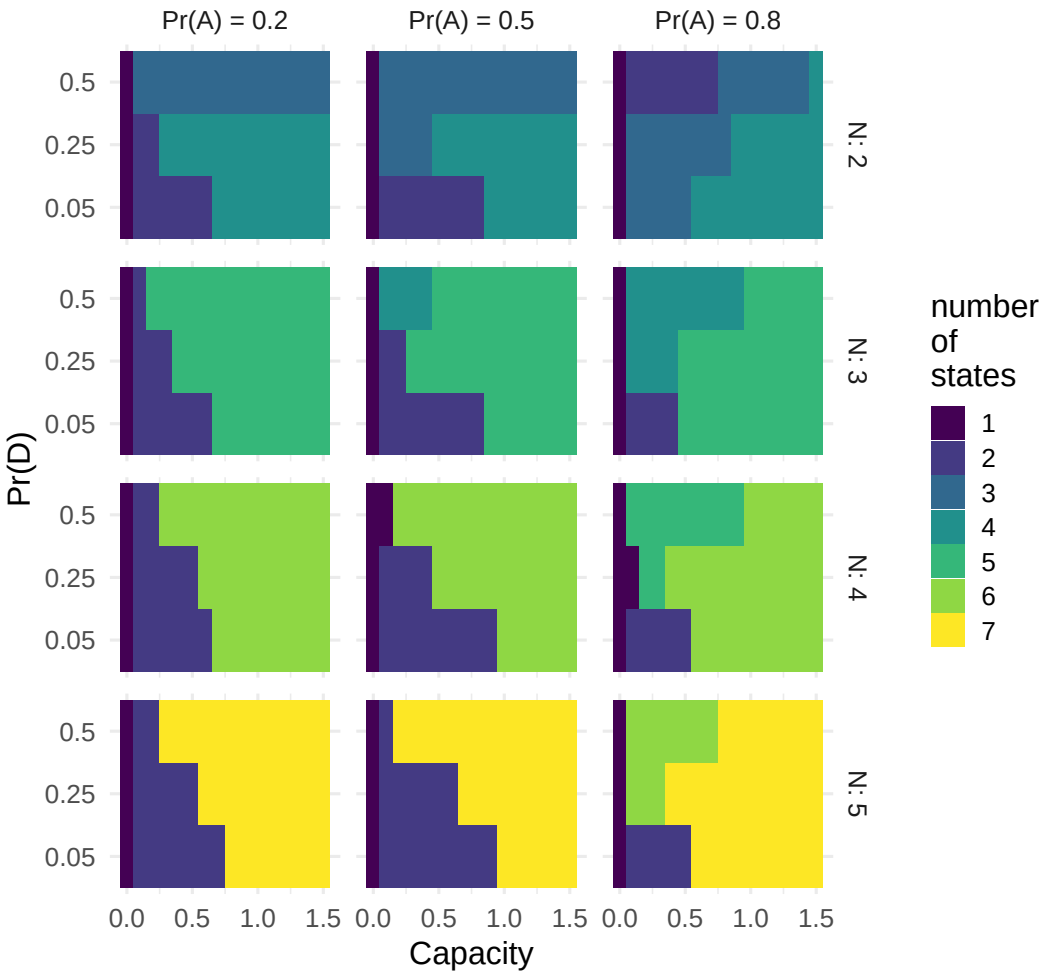


Figure S2: Number of functionally different states of Z in the resource-rational policy, as a function of $Pr(A)$, $Pr(D)$, N, and the observer's computational capacity C.

4 Visualizing the resource-rational mappings

We visualize two different resource-rational $\vec{X} \rightarrow Z \rightarrow \hat{Y}$ mappings discovered by our information bottleneck analysis (for observers with different channel capacity) in Figure S3. Focusing on a setting with $N = 2$ for better legibility, we find that Z can be in 4 possible states: an Ignorance, a Knowledge, and two Belief states (one for each box).

The left panel on Figure S3 shows the resource-rational mapping for an observer with limited resources, $I(\vec{X}; Z) = .5$. This observer engages in purely factive mindreading: we can see that it does not extract any information about B from the fact that the mapping from A and D to Z is similar across the two different values of B. The observer only distinguishes between two states: an Ignorance and a Knowledge state. Situations with $A = 1$, $D = 0$ are mapped to the Knowledge state, while all other situations (including false belief ones) are mapped to the Ignorance state. When Z is in the Knowledge state, the observer predicts that the actor will go toward box s (note that S is not shown in Figure S3 but when the actor is knowledgeable we have $s = b$). Note also that the mapping is not completely deterministic: for example even though the $A = 1$, $D = 0$ situations are mostly mapped to the Knowledge states, they are also sometimes mapped to the Ignorance state.

The right panel on Figure S3 shows the resource-rational mapping for an observer with enough cognitive resources ($I(\vec{X}; Z) = 1.5$) to achieve the theoretically maximal predictive performance. In addition to Ignorance and Knowledge, this observer can also represent the actor as Believing that the item is in box 1 (state B1) or box 2 (B2). In particular, when the actor has a deceived belief ($A = 1$, $D = 1$), the observer encodes either $Z = B1$ or $Z = B2$, depending on the value of B. Note however that these belief representations are mostly deployed in situations of deceived belief. When the actor knows the location of the item ($A = 1$, $D = 0$), the high-resource observer is most likely to encode the Knowledge state. In sum, even observers who achieve the theoretically maximal predictive performance do not always encode belief. We explore this finding in more depth in the next section.

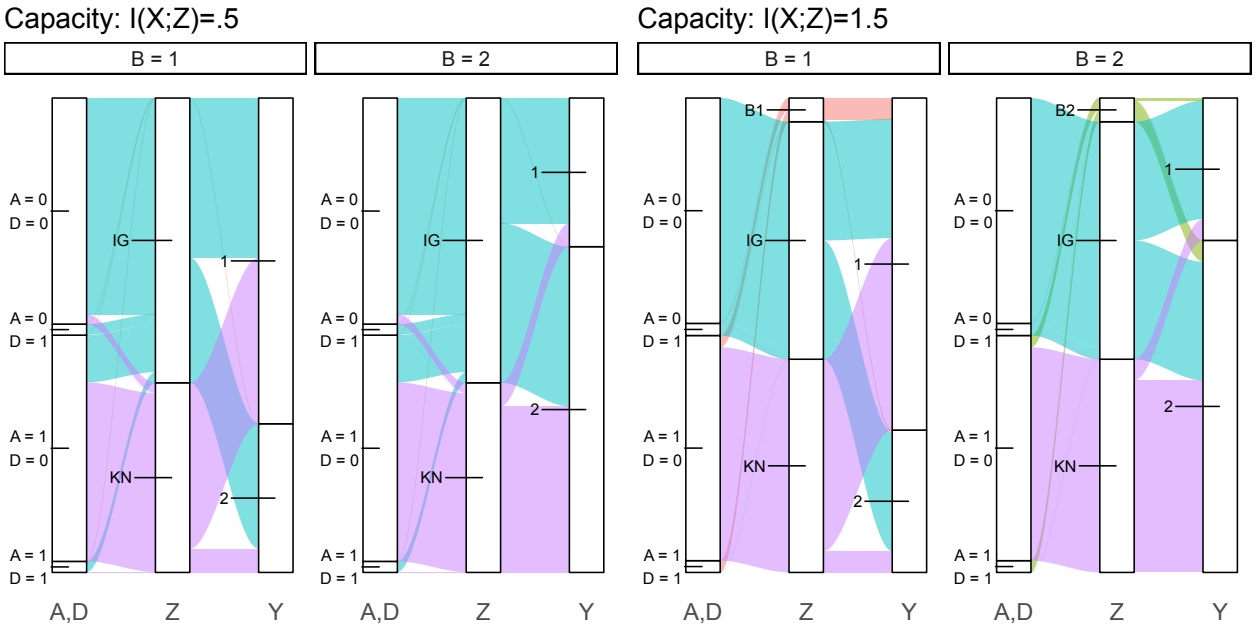


Figure S3: Visualization of the mapping from $\vec{X}$ to Z to $\hat{Y}$ for an observer with limited (left panel) and high cognitive resources (right panel). 'IG': ignorance, 'KN: knowledge', 'B': belief. The width of the flow between two states is proportional to the probability of going to the state on the right from the state on the left. For example an observer in a situation where $A = 0$, $D = 0$ is most likely to form the representation 'IG'. Parameters used were $N = 2$, $Pr(A) = .5$, $Pr(D) = .05$. Note that although we label the states as 'KN', 'IG', etc for better visualization, the functional role of these states emerged organically during optimization; we did not assign these functional roles manually during initialization.

5 Do high-resource observers use a fully-metarepresentational strategy?

The algorithm we use to derive resource-rational encoders finds the encoder that has the lowest $I(\vec{X}; Z)$ for a given level of predictive performance $I(Y; Z|S)$. As such, even among policies that achieve the highest theoretical predictive performance $I(Y; Z|S)$, we find the encoder that does so while extracting the least amount of information about $\vec{X}$.

We can therefore ask what encoding policies observers with high levels of cognitive resources use. The idea is that even though these observers have enough computational resources to perfectly solve mindreading tasks, they still would like to minimize their use of cognitive resources.

Figure S4a shows the proportion of available information that high-resource observers extract about variables A, D and B, as a function of $Pr(A)$ and $Pr(D)$. Since the maximum amount of information one can extract about a variable is its entropy, we compute proportion of extracted information about a variable V as $I(V; Z)/H(V)$, where $H(V)$ is the variable's Shannon entropy.

We find that high-resource observers always extract all the available information about A. In contrast, they only extract partial information about D and B. Furthermore, the variable they prioritize depends on $Pr(D)$: observers tend to encode B to a greater extent when deception is a likely possibility (high $Pr(D)$). When $Pr(D)$ is low, observers extract very little information about B.

To gain more insight on the kind of information that high-resource observers extract, we can visualize the resource-rational mapping between $\vec{X}$ and Z directly. Figure S3 right panel gave one example, and Figure S4b gives an example for a higher value of $Pr(D)$. Resource-rational encoders map situations with $A = 0$ to the Ignorance state, and map all situations with $A = 1, D = 1$ to the corresponding belief state (e.g. to $Z = B1$ if $b = 1$). Situations in which the actor knows the location of the item ($A = 1$, $D = 0$) are mapped either to the Knowledge or to the corresponding Belief state (Figure S4b). As such, the overall amount of encoding of B depends on the encoder's propensity to map $A = 1$, $D = 0$ situations to a Belief or a Knowledge state.

The fact that the encoder sends $A = 1$, $D = 0$ to either a Belief or a Knowledge state can be understood intuitively in the following way. Suppose the observer observes that the actor has perceptual access ($A = 1$). The observer now has two possible options for predicting the actor's

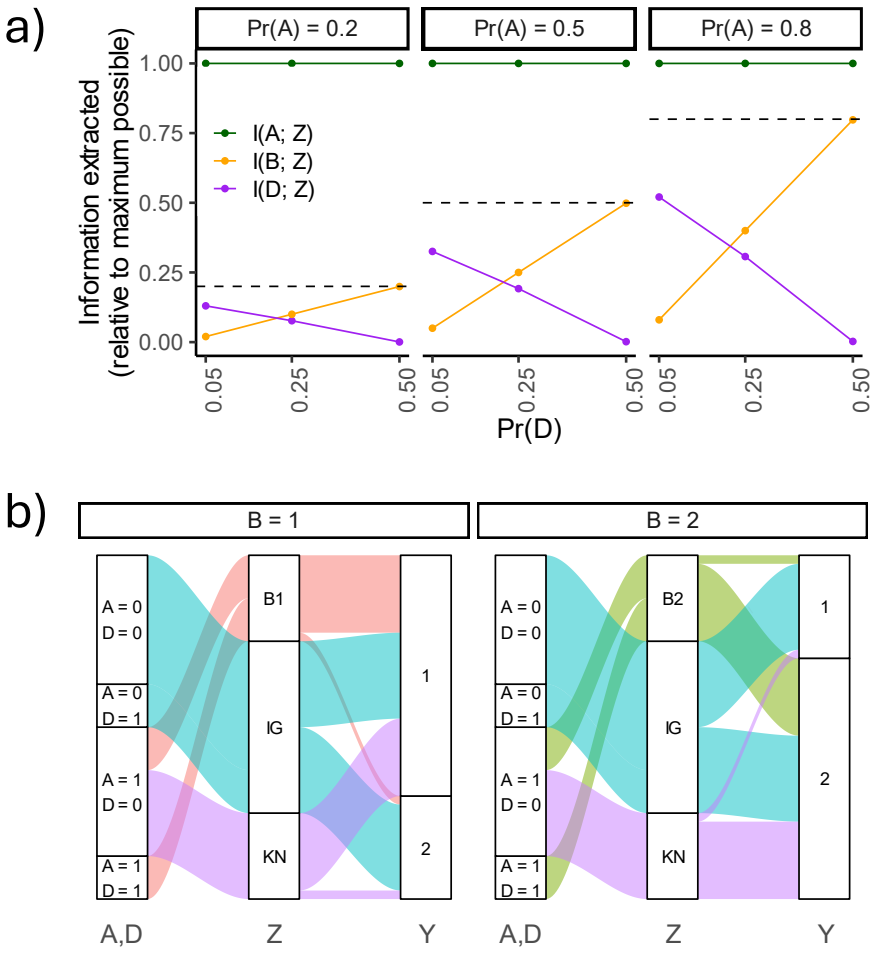


Figure S4: **a)** Proportion of available information extracted by high-resource observers about A, D and B, for $N = 2$. Proportion of information extracted is computed as (e.g.) $I(A, Z)/H(A)$. The dashed line represents $Pr(A)$. **b)** Visualization of the mapping from $\vec{X}$ to Z to $\hat{Y}$ for a high-resource observer with $C = 1.5$ for social ecology $N = 2$, $Pr(A) = .5$, $Pr(D) = .25$. 'IG': ignorance, 'KN: knowledge', 'B': belief. The width of the flow between two states is proportional to the probability of going to the state on the right from the state on the left. For example an observer in a situation where $A = 1$, $D = 0$ and $B = 1$ can form the representations 'KN' ('the actor knows the item's location'), and 'B1' ('the actor believes the item is in box 1') but is more likely to form 'KN'. Note that although we label the states as 'KN', 'IG', etc for better visualization, the functional role of these states emerged organically during optimization; we did not assign these functional roles manually during initialization.

behavior:

- Strategy 1: extract the value of D. If $D = 0$, represent the actor as having knowledge. If $D = 1$, extract the value of B.
- Strategy 2: extract the value of B without extracting the value of D.

Both strategies enable perfect predictive performance while extracting less than the full available information. Strategy 1 avoids paying the cost of extracting the value of B whenever $D = 0$. Strategy 2 always avoids paying the cost of extracting the value of D. The mapping shown in Figure S4b can be seen as a probabilistic mixture of the two strategies. The cost-efficiency of Strategy 1 in situations where $D = 0$ explains why propensity to encode D relative to B decreases with the value of $Pr(D)$, as observed in Figure S4a. Note also that the observer chooses to represent neither B nor D in situations with $A = 0$, saving considerable computational resources in these situations; this explains why the propensity to encode the value of these variables increases with $Pr(A)$, figure S4a.

In sum, even high-resource observers sometimes engage in factive mindreading, especially in environments where the possibility of deception is low (low $Pr(D)$). Factive mindreading manifests itself in two ways. First, observers represent the actor as being Ignorant in cases with $A = 0$, instead of explicitly representing the fact that the actor has a uniform distribution over possible beliefs. Second, observers often represent the actor as simply knowing the location of the item (instead of encoding the content of its belief) in situations where the actor has non-deceived perceptual access. These results show that implementing a fully meta-representational strategy is not a pre-requisite to achieve the maximal level of predictive performance. Instead, resource-rational observers can represent beliefs when it is needed, and represent knowledge / ignorance otherwise.

6 Using mindreading to learn about the world

In this series of tasks, we look at whether observers can solve the 'inverse' problem of predicting the location of an item from the actor's behavior. This series of tasks is motivated by a recent proposal that the proper evolutionary function of factive mindreading is social learning, and not social prediction ((*12*), but see (*13*)). We suggest that although factive mindreading is indeed helpful for social learning, this does not necessarily mean that it evolved primarily for that purpose. Specifically, we demonstrate that good social learning performance can also emerge simply as a byproduct of optimizing for social prediction. We show that the representations optimized for our *first* task (predicting behavior from the state of the world), can also be co-opted for predicting the state of the world from observation of another individual's behavior.

Formally, in this new task the observer must predict the true location of the item (S) in a situation where they know the actor's choice (Y), whether the actor had perceptual access (A), whether the actor was deceived (D), but don't know either the initial (B) or current location of the item (S). The goal is to compute $Pr(S|Y, A, D)$. By Bayes' rule:

$$Pr(S|Y, A, D) = \frac{Pr(Y|S, A, D)Pr(S|A, D)}{Pr(Y|A, D)} \tag{S17}$$

Because we marginalize over B (whose value we don't know), we have $Pr(S|A, D) = Pr(S) = \frac{1}{N}$ and $Pr(Y|A, D) = Pr(Y) = \frac{1}{N}$, such that $Pr(S|A, D) = Pr(Y|A, D)$.[3] So we have:

$$Pr(S|Y, A, D) = Pr(Y|S, A, D) \tag{S18}$$

Since resource-limited observers do not have access to the true generative model, they have to substitute $q(\hat{Y}|S, A, D)$:

$$Pr(S|Y, A, D) = q(\hat{Y}|S, A, D) \tag{S19}$$

where $q(\hat{Y}|S, A, D) = \sum_z q(\hat{Y}|S, Z)q(Z|A, D)$. We assume that the encoder $q(Z|A, D)$ and decoder $q(\hat{Y}|S, Z)$ are optimized for our main task of predicting the actor's behavior, i.e. we take them off-the-shelf from the policies derived in our first series of experiments. As above, we look at an automatic policy, a low- and a high-resource observer ($C \in \{0, .5, 1\}$), and use the same parameters ($N = 3$, $Pr(A) = .2$, $Pr(D) = .05$). See Section 12 for other points in parameter space.

[3] When we marginalize over B, information about A and D cannot change our estimate of Y or S.

In **Experiment 5**, the actor knows the location of the item ($A = 1$, $D = 0$, upper-left on Figure S5). We find that all three observers correctly infer that the item is located in the box that the actor is reaching toward, although this inference is stronger in observers with more cognitive resources.

In **Experiments 6 and 7**, the low- and high-resource observers correctly judge that the actor's behavior is not diagnostic about the item's location if the actor has information that is not up-to-date ($A = 1$, $D = 1$, upper right on Figure S5), or if the actor is ignorant ($A = 0$, lower-left). In contrast, in all conditions the automatic observer predicts that the item is in the box that the actor is reaching toward.

In sum, the low-resource observer (a factive observer) performs close to the ideal policy in all three experiments of our social learning task, despite the fact that its representations were optimized solely for the separate task of predicting behavior. These experiments also illustrate that factive observers in our model can behave adaptively in situations of *egocentric ignorance*, where another agent knows something that they don't (*14*).

Note that our analysis does not preclude the possibility that social learning has also been an evolutionary pressure for the evolution of factive mindreading. We explore this proposal in more detail in the next section.

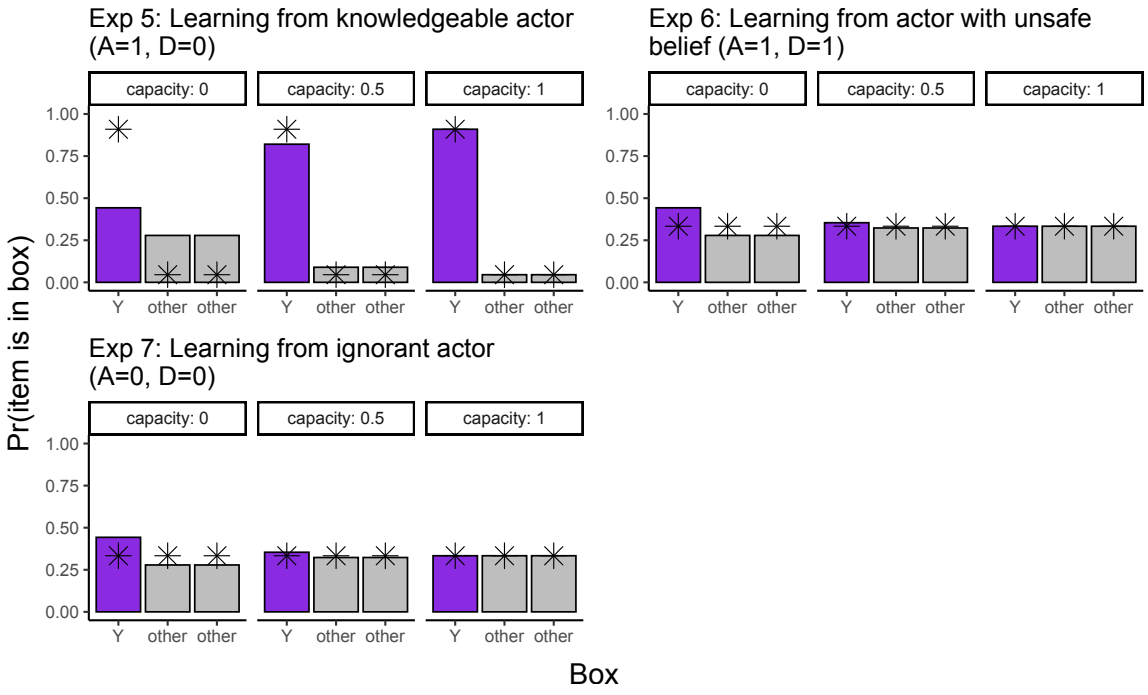


Figure S5: Observer predictions about the location of the item, given the behavior of another agent. Stars represent the ideal non resource-limited policy. Parameters used were $N = 3$, $Pr(A) = .2$, $Pr(D) = .05$.

7 Different training objective

Here we consider the possibility that factive mindreading can emerge from a different objective, namely that of learning about the state of the world by observing another agent's behavior (*12*). Our conclusion will be mixed. On the one hand, in our simple setting a factive mindreading strategy is indeed sufficient for that objective. On the other hand, there are also situations where representing the beliefs of others can help an observer learn about the world.

7.1 Factive mindreading is sufficient in our simple setting.

A mindreading policy optimized for learning about the state of the world S on the basis of A, D, Y can be co-opted for predicting behavior, but will not represent the value of B, since the observer does not have access to B in the social learning task.[4] Taking the true generative model $Pr(Y|S, A, D, B)$ and marginalizing over B, it is easy to show that the behavior prediction $Pr(Y|S, A, D)$ we obtain is similar to the factive policies we derived using the information bottleneck method in the main text, i.e. we have:

$$\Pr(Y = y|A, D, S) = \begin{cases} 1/N, & \text{if } A = 0 \vee D = 1 \\ 1 - \epsilon_N, & \text{if } A = 1, D = 0, y = s \\ \epsilon_N/(N-1), & \text{if } A = 1, D = 0, y \neq s \end{cases} \tag{S20}$$

The only difference between this policy and the factive policies studied in the main text is that the latter are noisier.

7.2 The value of belief representations for social learning.

Outside of our simple model, we can imagine situations where representing belief can help predicting the state of the world from another agent's behavior. Suppose for instance that Bob has a (possibly mistaken) belief that valuable items are painted red and non-valuable items are painted blue. A red item and a blue item are put into different boxes (in view of Bob, but outside of your

[4]The case where the observer has access to B in the social learning task is less interesting; with access to B, observing Y would not give the observer any additional information about S, which means the task would not really be a social learning task. This can be seen by the fact that B screens off Y from S in the causal graph.

view). You see Bob going toward one box. In order for you to infer the location of the valuable item, it can be valuable to know whether Bob has a false belief (if Bob is mistaken about the color of valuable items you should go toward one of the boxes he is not reaching for). In sum, factive mindreading is not optimal for learning about the world in this setting.

Note that factive mindreading might be optimal even in this more complicated setting if we re-introduce cognitive limitations on the observer. This would be a similar argument as the one we make in the main text, except that we would be explaining factive mindreading as a resource-rational strategy for social learning instead of social prediction.

We think that that it is plausible that factive mindreading arose because of evolutionary selection pressures for both resource-rational social prediction and resource-rational social learning. In the main text we choose to focus on social prediction because it is the setting where belief representation is most obviously valuable, and so it provides the more stringent test for our claim that resource limitations might favor factivity.

8 Representing ignorance

Researchers debate whether non-human primates have sophisticated representations of ignorance (*15*), or simply default to making 'no prediction' whenever the actor fails to witness a relevant event (*16, 17*). Factive mindreaders in our simulations implement the latter heuristic, see Figure 4 in main text. Meta-representational observers explicitly represent that the ignorant actor attributes equal probability to each possible item location, but factive mindreaders are simply agnostic about where the actor will go. At the same time, our work does not definitely establish that this simple way of representing ignorance is always the optimal strategy for resource-limited observers. More sophisticated representations might be adaptive in settings where ignorant actors behave non-randomly, and future modeling work should explore this possibility. Consider for example a 'partially ignorant' actor who doesn't know whether the reward is in box A or box B, but knows that it isn't in box C (*18*). A mindreader with a sophisticated grasp of ignorance will predict that the actor will look in either A or B, but a mindreader who defaults to 'no prediction' will predict that the actor might look in any box.

9 Decoding costs

For tractability, our model focuses on the information-processing costs involved in creating representation Z, and sets aside the cost of generating a posterior distribution over $\hat{Y}$ from Z (see also (*10*, *19*)). Here we report a preliminary exploration of how the cost of this inference might vary across factive and meta-representational policies.

Following (*20*), we operationalize the cost of inference as the KL divergence of the observer's prior from the posterior. In our case, this can be written as:

$$\text{cost} = D_{KL}(q(\hat{y}|z,s)||q(\hat{y}|s)) \tag{S21}$$

In words, the observer has a prior belief $q(\hat{y}|s)$ about the actor's action, conditional on the actual location s of the reward. The cost of updating this belief by consulting Z is the divergence between this prior and $q(\hat{y}|z,s)$.

To compute the expected inference cost, we compute the expectation of this value across different possible values of $\vec{x}$:

$$\text{expected cost} = \sum_{\vec{x}} D_{KL}(q(\hat{y}|z,s)||q(\hat{y}|s))p(\vec{x}) \tag{S22}$$

Figure S6 shows the expected inference cost for factive and meta-representational policies, for the social ecologies studied in the main text as well as the additional social ecologies studied in Figures S11 to S13.[5] We find that meta-representational policies incur a higher inference cost than factive policies.

For robustness, we also performed a similar analysis under a different operationalization, where we assume that the observer's prior is simply $q(\hat{y})$. This analysis effectively measures the cost of updating on the basis of S and Z jointly. Figure S7 shows that under this assumption meta-representational policies still incur a higher expected inference cost than factive policies.

[5]Since the factive policies that are discovered by our resource-rational optimization procedure are stochastic, here we consider an 'ideal' factive policy that makes Bayes-optimal inferences from representations of Knowledge and Ignorance. Tests with factive policies generated by the resource-rational model (not shown) yield qualitatively equivalent results. In each social ecology, the meta-representational policy was selected among the ones generated by the resource-rational model by choosing a high enough value of β so that the policy is effectively Bayes-optimal.

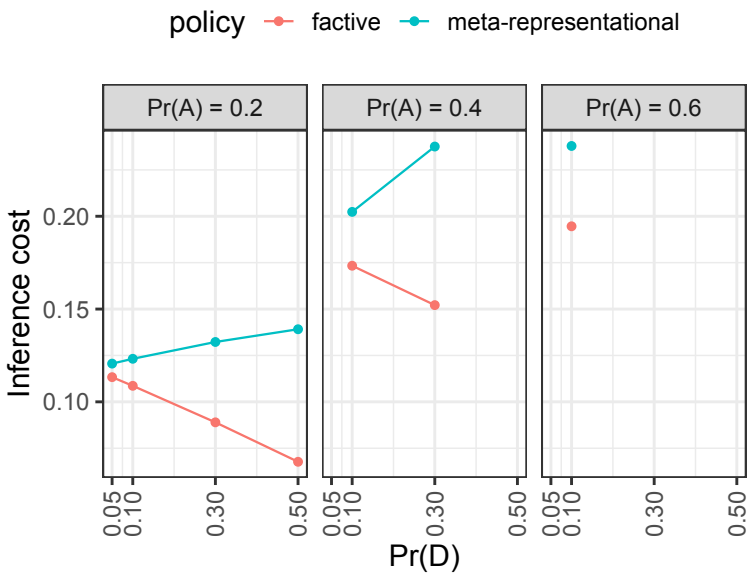


Figure S6: Expected inference cost for the factive and meta-representational policy, across selected social ecologies. Here we used $N = 3$.

In sum, although a fuller investigation is needed, this preliminary analysis suggests that considering decoding costs would not alter our conclusion that factive strategies can out-perform meta-representational policies for low-resource observers.

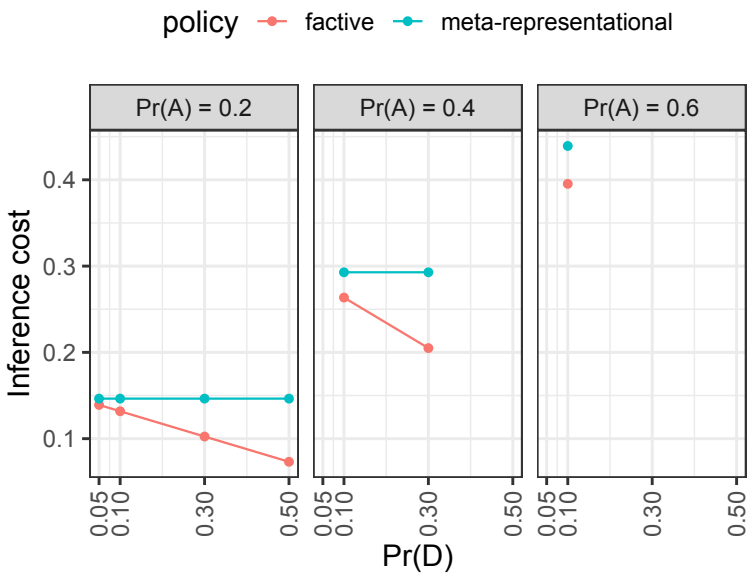


Figure S7: Expected inference cost for the factive and meta-representational policy, across selected social ecologies, for an alternative operationalization of inference cost. Here we used $N = 3$.

10 Information extraction curves across social ecologies

Figure S8 shows how much information resource-rational observers extract (in our main task) about knowledge- and belief-relevant variables, as a function of their channel capacity $I(X;Z)$, across values of $Pr(A)$ and $Pr(D)$, for $N \in \{2,3,4,5\}$.

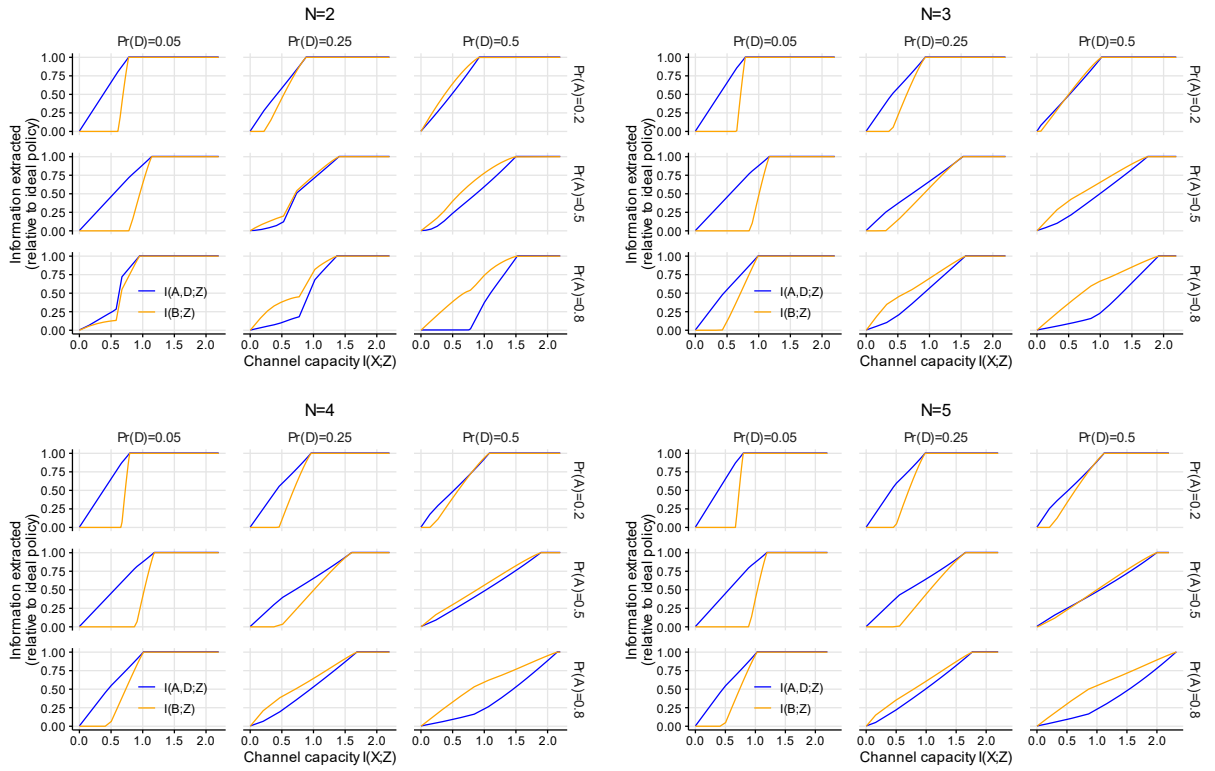


Figure S8: Amount of information extracted from knowledge-relevant variables A and D (blue), and belief-relevant variable B (orange), as a function of the observer's cognitive resources (channel capacity $I(X;Z)$), across different values of N, $Pr(A)$ and $Pr(D)$.

11 Information extracted about A, D and B as a function of information capacity

In the main text (and the previous section) we display the information that Z extracts about A and D jointly, to highlight the distinction between these 'knowledge-relevant' variables and B. In Figure S9, $I(A;Z)$ and $I(D;Z)$ are shown separately, allowing a more-fine grained appreciation of the dynamics.

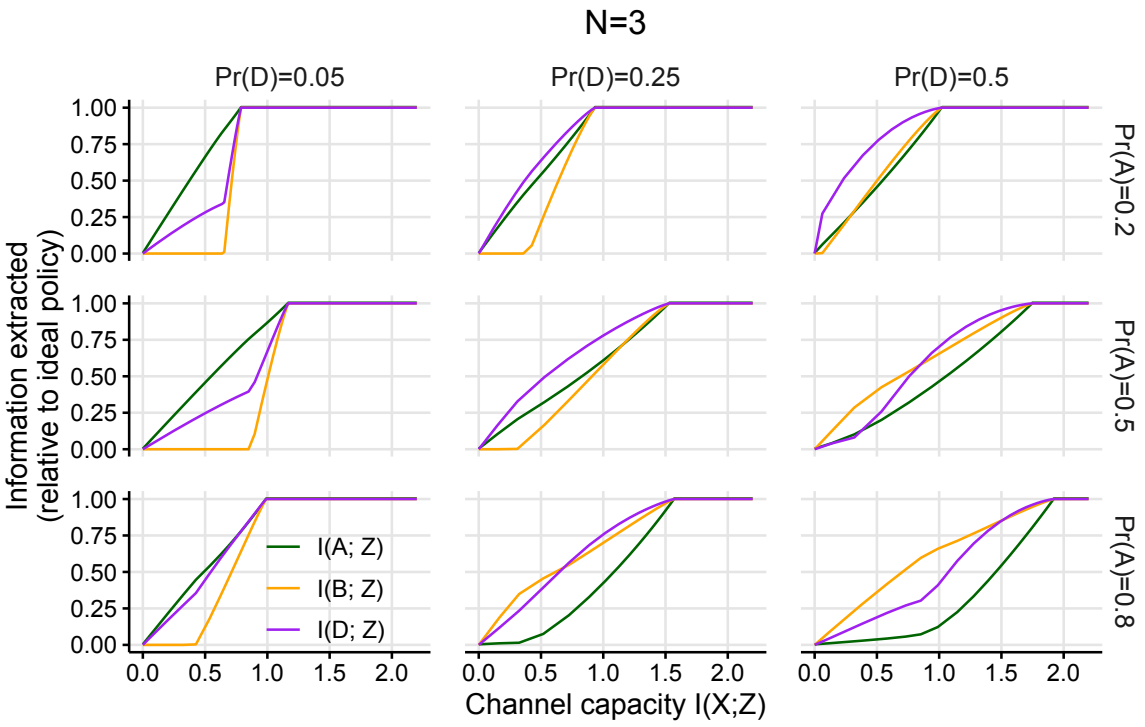


Figure S9: Amount of information extracted from variables A (green), D (purple), and B (orange), as a function of the observer's cognitive resources (channel capacity $I(X, Z)$), shown for $N = 3$, and different values of $Pr(A)$ and $Pr(D)$.

12 Supplementary Experiments

In the main text we report experiments conducted in social ecology $N = 3$, $Pr(A) = .2$, $Pr(D) = .05$. Here we conduct experiments in other points of parameter space.

For each selected social ecology we pick a set of four resource-rational observers varying in their channel capacity. We select the sets of channel capacities to provide a representative range of the types of resource-rational observers for a given social ecologies — for instance in settings that are favorable to factive mindreading we show a zero-capacity observer, a factive observer, an observer who imperfectly represents belief, and an observer that achieves the theoretically maximal predictive performance. As such, the set of channel capacities we select can differ slightly from one social ecology to the next.

Our first set of experiments (Figures S11 to S13) target areas of parameter space that we identified as being hospitable to factive representations, see Figure S10. We find similar results as in the main text: observers who extract information about A and D but not about B successfully pass Knowledge and Ignorance tasks, but fail False Belief and Gettier tasks. The one exception is the parameter combination $Pr(A) = .2$, $Pr(D) = .5$ (at the very edge of the space favorable to factivity), where factive mindreading is only resource-rational for an observer with very low capacity: this observer extracts very little information from A and D and therefore does not perform much better than the zero-capacity observer.

Our second set of experiments (Figures S14 and S15) target points of parameter space where purely factive mindreading is unsuccessful. We find that limited-capacity observers behave mostly as 'noisy' versions of the ideal policy. Note that for a low value of $Pr(D)$ limited-capacity observers perform better at the Knowledge than the False Belief task (Figure S14 left panel). Figure S10 shows the parameter combinations used in our first two sets of supplementary experiments.

Finally, our third set of experiments uses the same values of $Pr(A)$ and $Pr(D)$ as in the main text, but varies the value of N. We find that, as for $N = 3$, observers who extract information about A and D but not about B successfully pass Knowledge and Ignorance tasks, but fail False Belief and Gettier tasks.

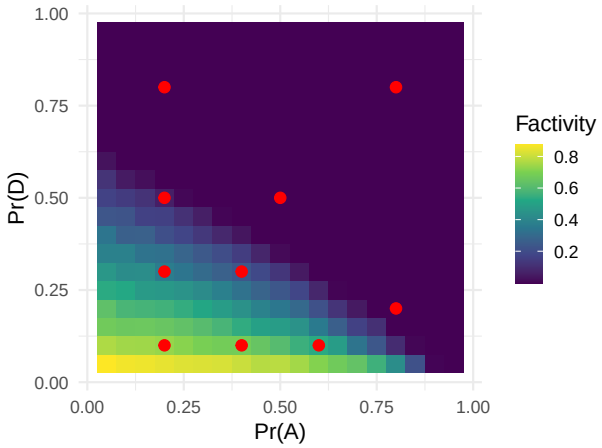


Figure S10: Location in parameter space of Supplementary experiments shown in Figures S11 to S15. For interpretation of the 'factivity' measure see Figure 3b in main text.

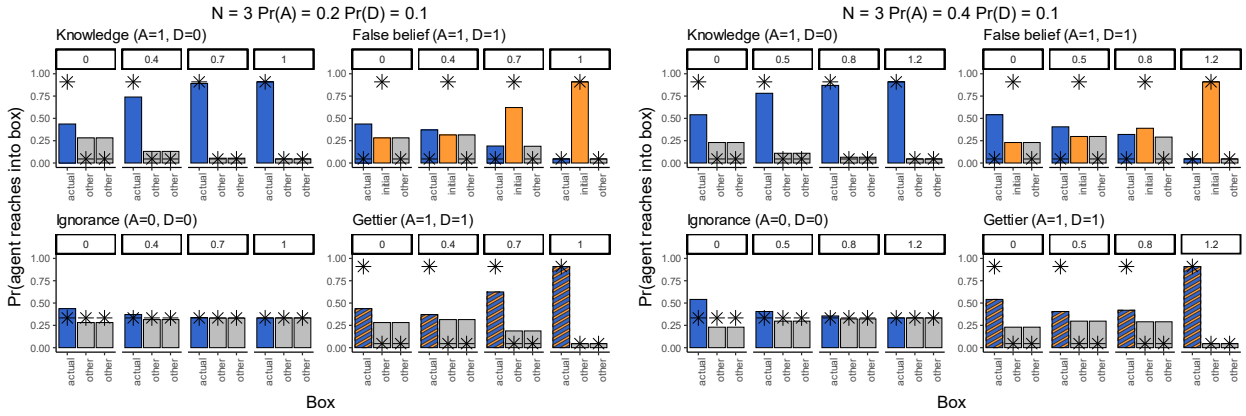


Figure S11: Predictions made by observers with different channel capacity in Supplementary experiments. Blue bars indicate the current location of the item, and orange bars (in cases where $D = 1$) indicate its initial location. Stars represent the ideal non resource-limited policy. Facet numbers represent channel capacity ($I(\vec{X}; Z)$) of observer.

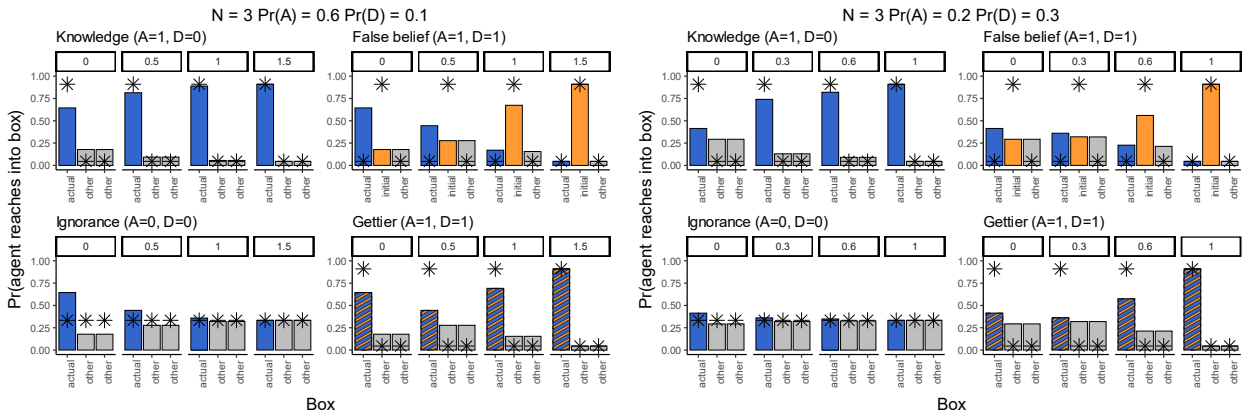


Figure S12: (See S11 for caption)

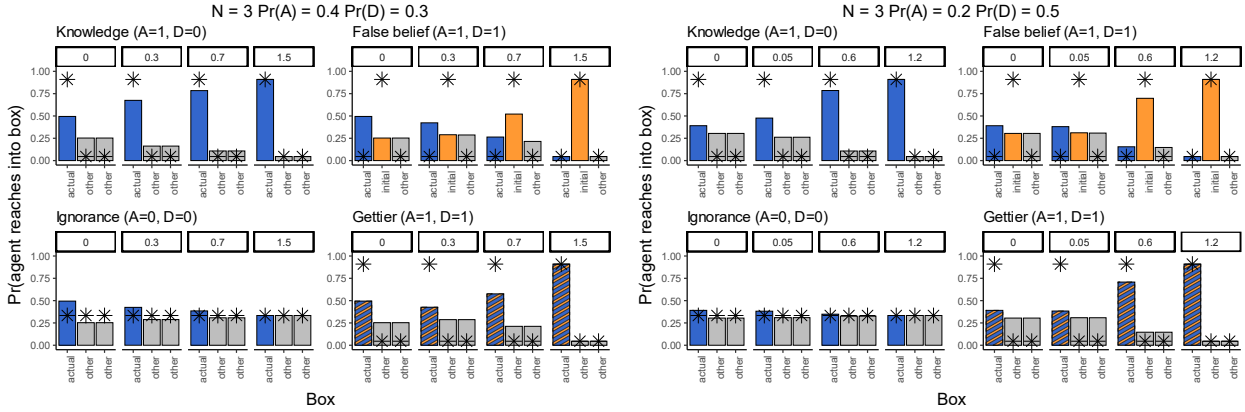


Figure S13: (See S11 for caption)

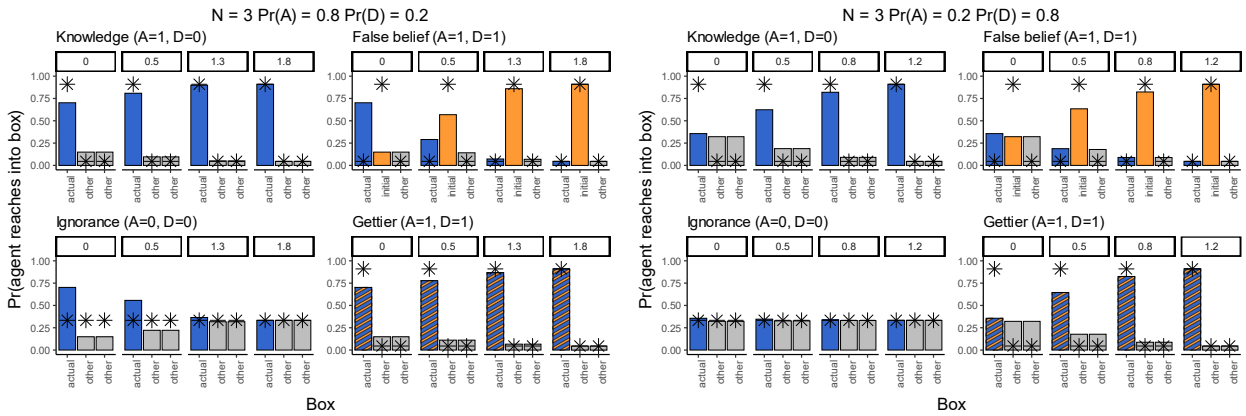


Figure S14: (See S11 for caption)

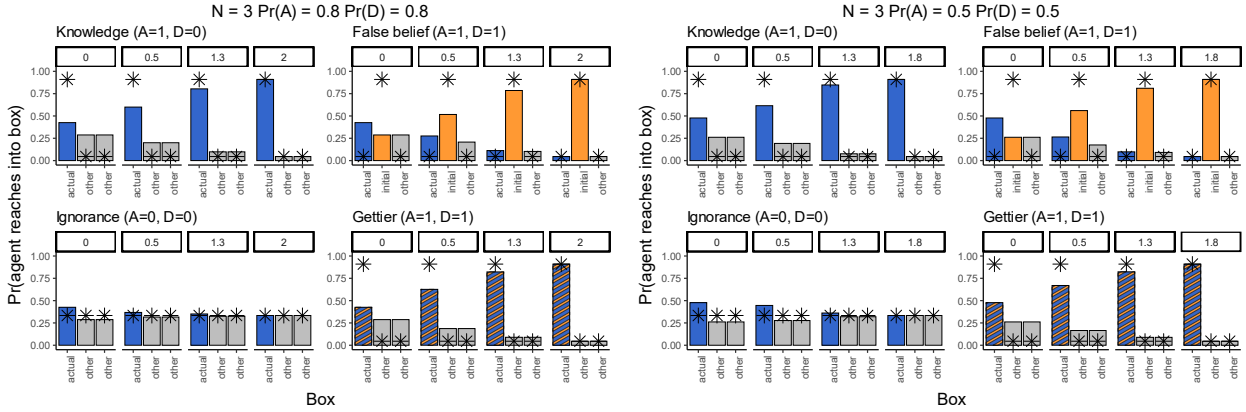


Figure S15: (See S11 for caption)

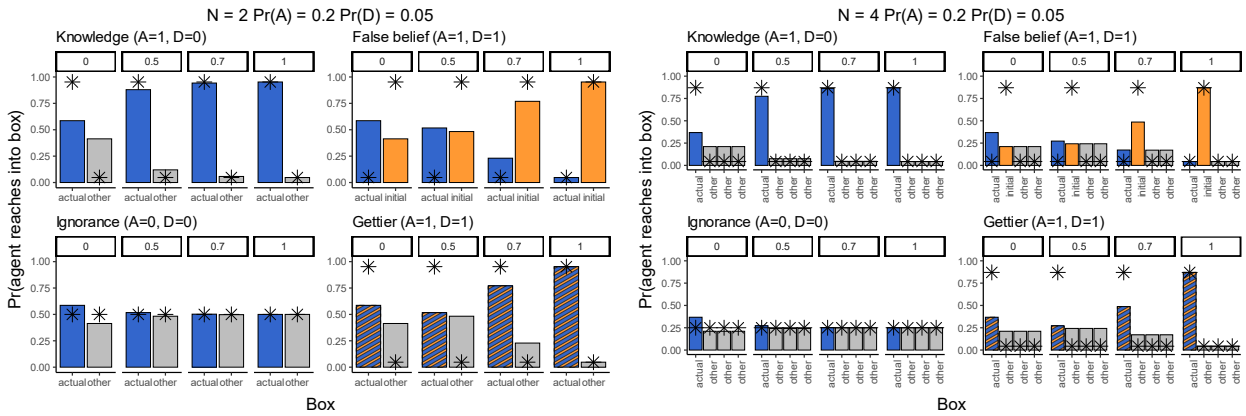


Figure S16: (See S11 for caption)

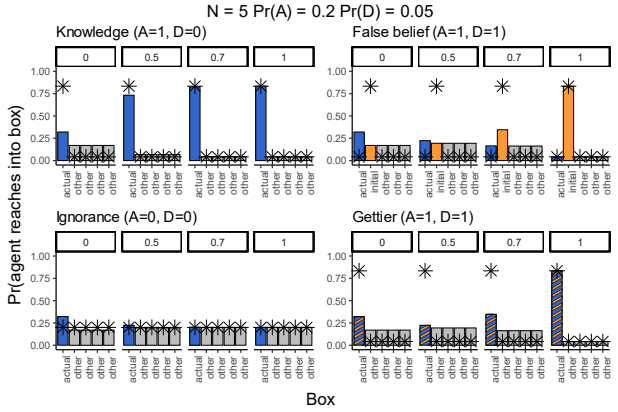


Figure S17: (See S11 for caption)

13 Supplementary Experiments (Social learning)

In the main text we report social learning experiments conducted in a setting with parameters $N = 3$, $Pr(A) = .2$, $Pr(D) = .05$. In Figures S18-S22 we report the results of experiments in other social ecologies.

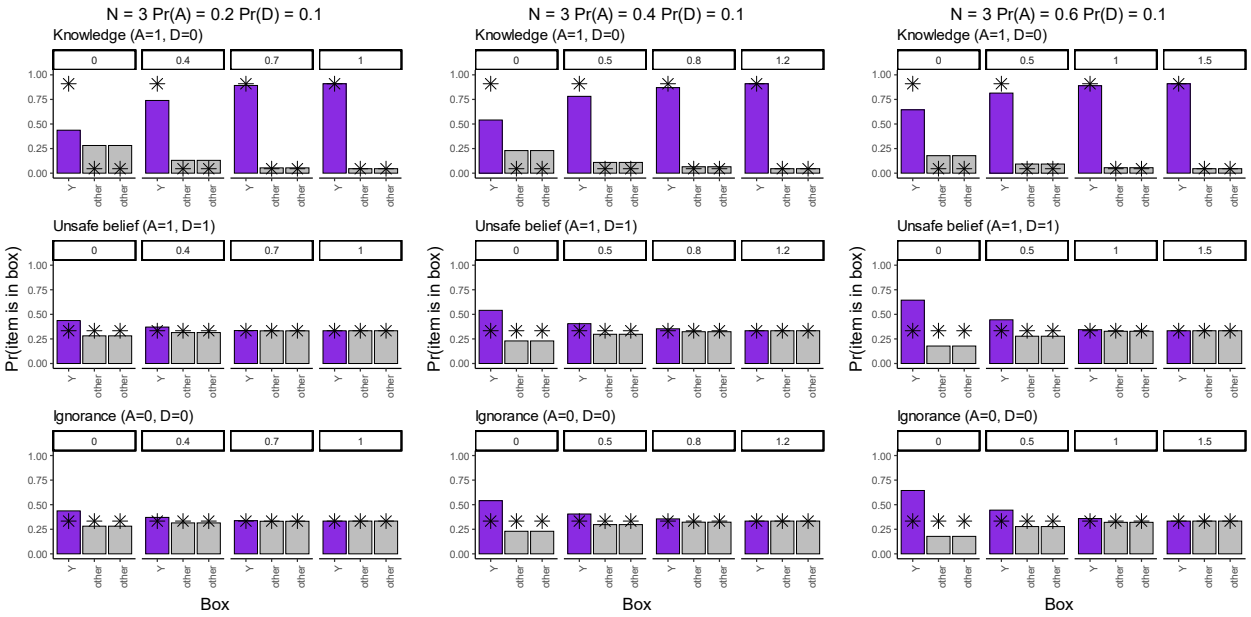


Figure S18: Observer predictions about the location of the item, given the behavior of another agent. Stars represent the ideal non resource-limited policy. Facet numbers represent channel capacity ($I(\vec{X}; Z)$) of observer.

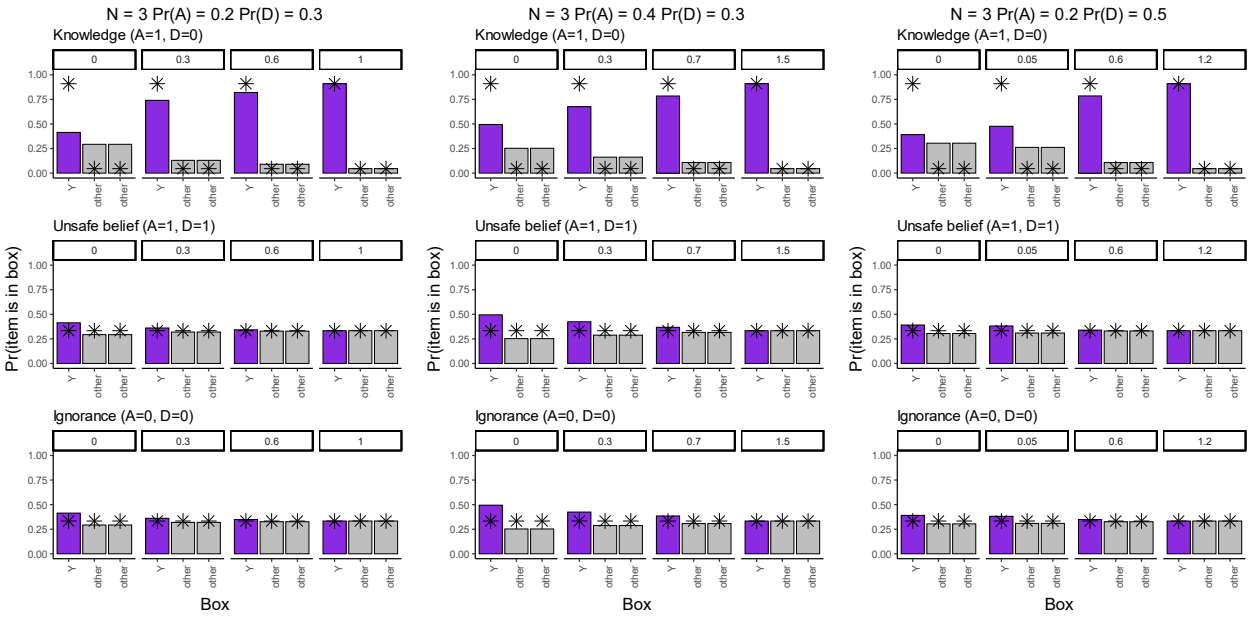


Figure S19: (See S18 for caption)

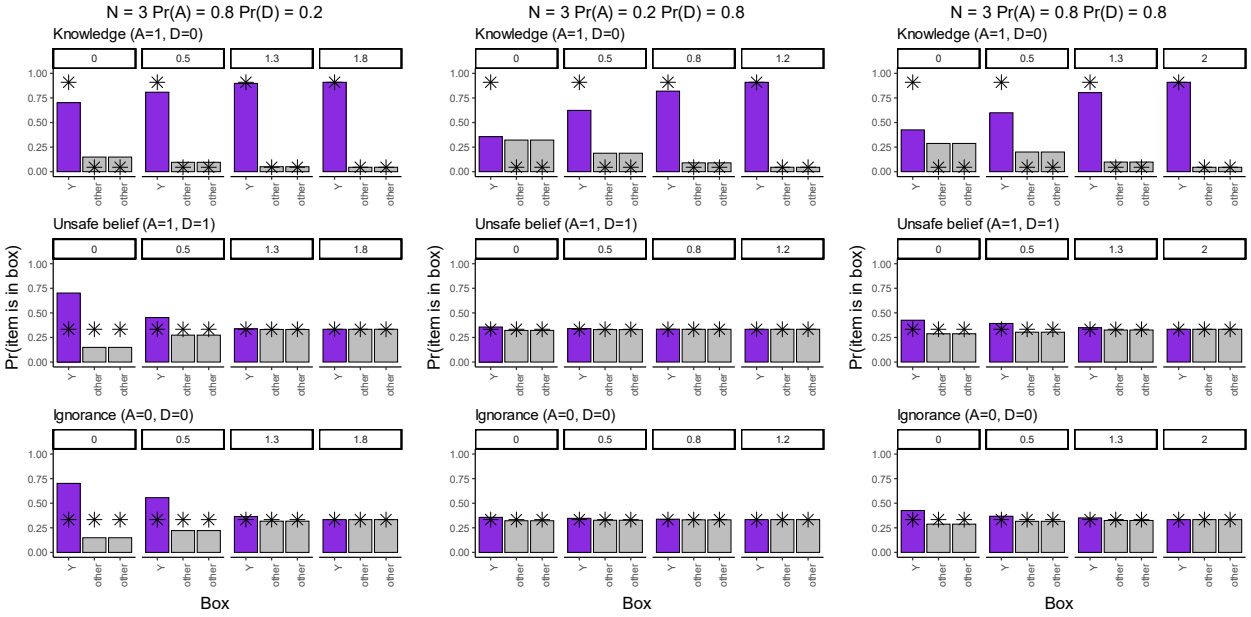


Figure S20: (See S18 for caption)

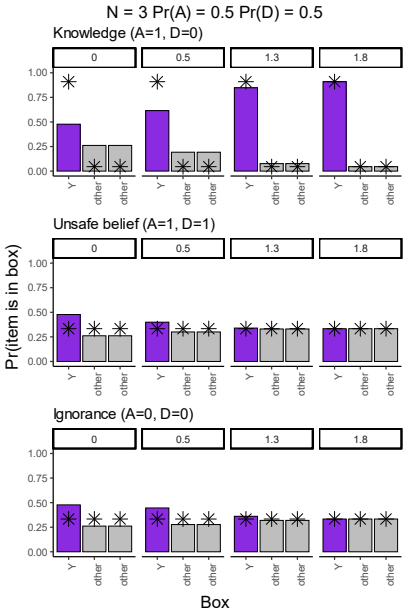


Figure S21: (See S18 for caption)

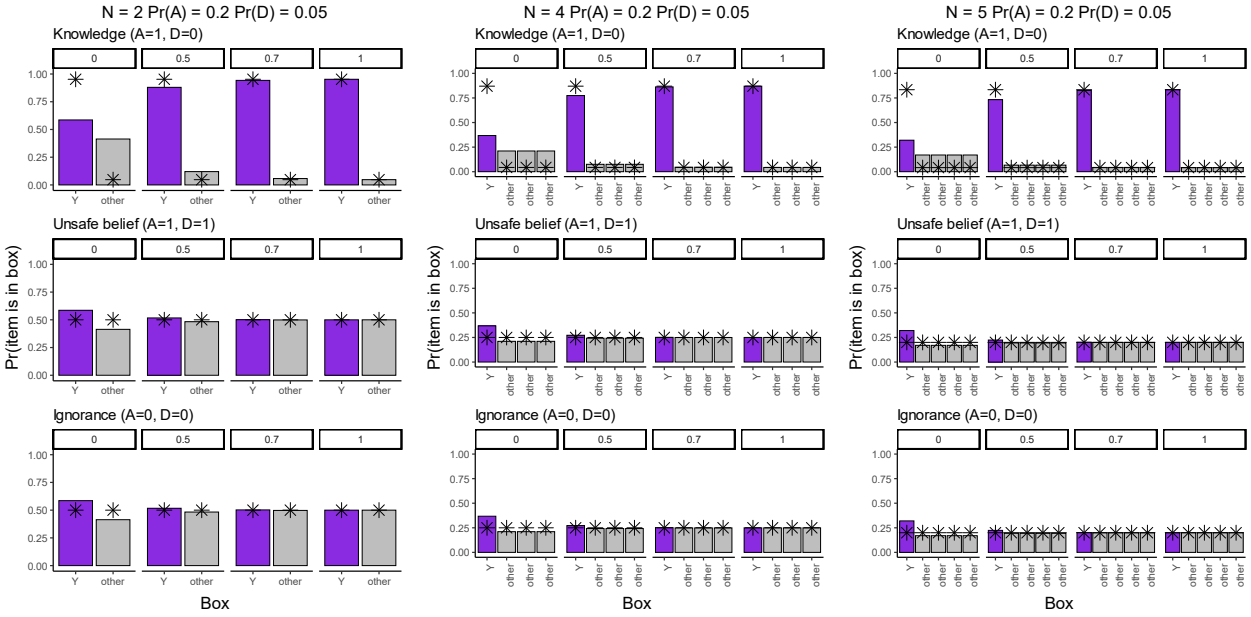


Figure S22: (See S18 for caption)

14 Control Simulations

In control simulations, we abandon the assumption that the observer has a pre-existing representation of the location of the item S. Instead the observer has to pay the information-theoretic cost of extracting information about S, in addition to A, D and B. That is, we are looking for the encoder:

$$q_C(z|\vec{x})^\star = \arg\max_q I(Y;Z) \quad \text{subject to } I(\vec{X};Z) \leq C \tag{S23}$$

where $Pr(\vec{X})$ now represents the joint probability distribution over A, D, B and S. Figure S23 shows the amount of information extracted by an observer as a function of the observer's information-processing capacity, across various values of $Pr(A)$ and $Pr(D)$, for $N = 3$. Figure S24 displays the results of experiments performed with selected resource-rational observers.

Both Figures show that factive mindreading does not arise in this setting. Instead, resource-limited observers behave as 'noisy' versions of the ideal policy: they make the same directional predictions, but put less probability mass on the dominant option, see Figure S24. These policies are for example capable of predicting where an actor with a false belief will look, but are less certain of that prediction than the optimal non-resource-constrained policy. In contrast, the factive observers in our main simulations are agnostic about the behavior of actors with a false belief.

Note that in the experiments shown in Figure S24, the low-capacity observer performs exactly as well in the False Belief and Gettier task as in the Knowledge task. In this sense the control simulations show that situations that require belief attribution are not inherently more difficult than knowledge attribution tasks in our setting.

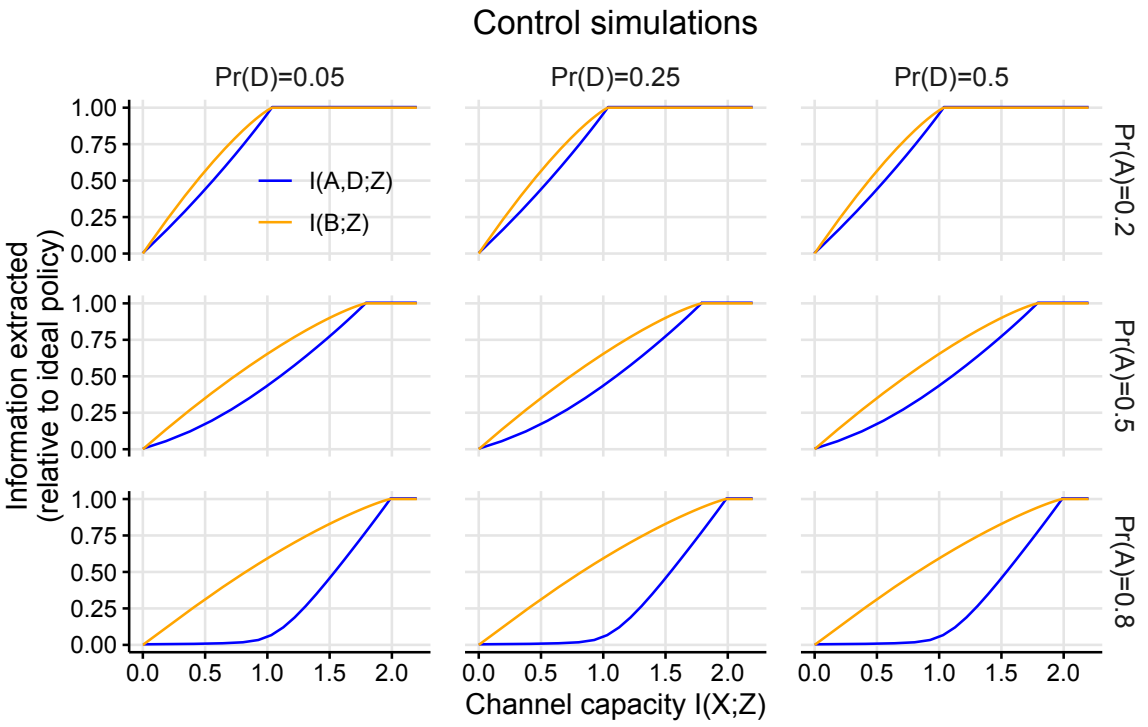


Figure S23: Amount of information extracted by resource-rational observers about knowledge-relevant variables *A* and *D* (blue) or belief-relevant variable *B* (orange), as a function of the observer's computational capacity, when observers do not have free access to *S*.

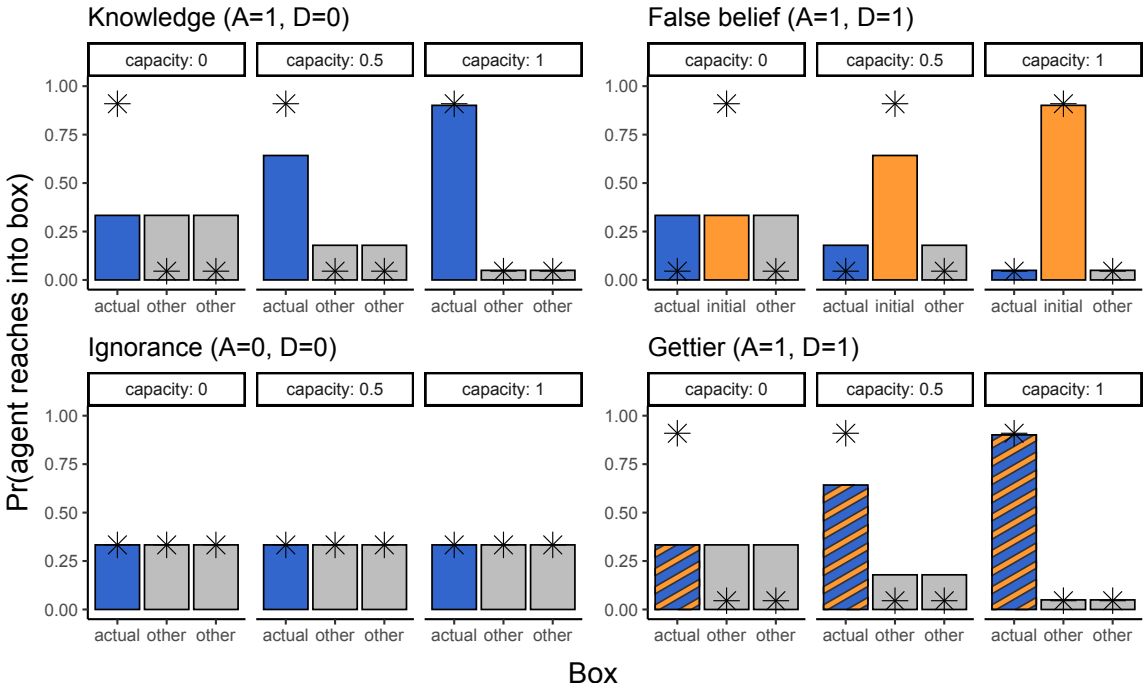


Figure S24: Model predictions in control simulations without side information. Stars represent the ideal non resource-limited policy. Parameters used were $N = 3$, $Pr(A) = .2$, $Pr(D) = .05$.

References

1. C. L. Baker, J. Jara-Ettinger, R. Saxe, J. B. Tenenbaum, Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour* **1** (4), 0064 (2017).

2. J. Jara-Ettinger, Theory of mind as inverse reinforcement learning. *Current Opinion in Behavioral Sciences* **29**, 105–110 (2019).

3. C. G. Lucas, *et al.*, The child as econometrician: A rational model of preference understanding in children. *PloS one* **9** (3), e92160 (2014).

4. J. Pearl, *Causality* (Cambridge university press) (2009).

5. N. Slonim, *The information bottleneck: Theory and applications*, Ph.D. thesis, Citeseer (2002).

6. N. Tishby, F. C. Pereira, W. Bialek, The information bottleneck method. *arXiv preprint physics/0004057* (1999).

7. D. Gondek, T. Hofmann, Conditional information bottleneck clustering, in *3rd ieee international conference on data mining, workshop on clustering large data sets* (2003), pp. 36–42.

8. S. Arimoto, An algorithm for computing the capacity of arbitrary discrete memoryless channels. *IEEE Transactions on Information Theory* **18** (1), 14–20 (1972).

9. R. Blahut, Computation of channel capacity and rate-distortion functions. *IEEE transactions on Information Theory* **18** (4), 460–473 (1972).

10. N. Zaslavsky, C. Kemp, T. Regier, N. Tishby, Efficient compression in color naming and its evolution. *Proceedings of the National Academy of Sciences* **115** (31), 7937–7942 (2018).

11. N. Slonim, N. Tishby, Agglomerative information bottleneck. *Advances in neural information processing systems* **12** (1999).

12. J. Phillips, *et al.*, Knowledge before belief. *Behavioral and Brain Sciences* **44**, e140 (2021).

13. D. Kampis, G. Csibra, Three cognitive mechanisms for knowledge tracking. *Behavioral and Brain Sciences* **44**, e157 (2021).

14. E. Westra, J. Nagel, Mindreading in conversation. *Cognition* **210**, 104618 (2021).

15. L. Townrow, C. Krupenye, Bonobos point more for ignorant than knowledgeable social partners. *Proceedings of the National Academy of Sciences* **122** (6) (2025).

16. A. Martin, L. R. Santos, What cognitive representations support primate theory of mind? *Trends in cognitive sciences* **20** (5), 375–382 (2016).

17. D. J. Horschler, L. R. Santos, E. L. MacLean, Do non-human primates really represent others' ignorance? A test of the awareness relations hypothesis. *Cognition* **190**, 72–80 (2019).

18. A. Royka, D. J. Horschler, W. Bargmann, L. R. Santos, Exploring the evolutionary roots of theory of mind: Primate errors on false belief tasks reveal representational limits. *Cognition* **270**, 106400 (2026).

19. M. Taylor-Davies, C. G. Lucas, Balancing utility and cognitive cost in social representation. *arXiv preprint arXiv:2310.04852* (2023).

20. J.-Q. Zhu, T. L. Griffiths, Computation-limited Bayesian updating: A resource-rational analysis of approximate Bayesian inference. *Psychological Review* (2025).