AEA Papers and Proceedings 2025, 115: 643–647
https://doi.org/10.1257/pandp.20251109

Real and Assumed Information†

By Ned Augenblick, Matthew Backus, Andrew T. Little, and Don A. Moore*

One way to view the job of scientists is to translate information (or data) about the world into beliefs about how the world works. This translation might involve statistical procedures, writing formal theories, or other modes of analysis. Making progress in each case requires imposing assumptions: parametric assumptions in statistical analysis, tractable theoretical assumptions to create propositions, or conceptual boundaries to justify qualitative categorization. The need to make assumptions and form simplified models of the world is not limited to scientists. In the social sciences, we model people as translating information into beliefs that are relevant for the decisions they make—for example, which route to take to work, who to vote for, and what to eat for breakfast—but in reality this translation requires some mental framework (i.e., a model). In this paper we suggest a new way to explore how the use of simplified models—whether in science or more broadly—affects beliefs.

To understand the effect of using simplified models, we set out a model of models. The issue with modeling models is that models are complicated objects and typically situation specific.[1] As a result, it is challenging to create a tractable theoretical structure that allows meaningful and general predictive statements. To solve this challenge, we largely abstract away from discussing the *specific structure* of models and instead think about models as *plausible assumptions* about uncertain random variables.

In our model, a person observes data x and is asked to form a belief about a statistically related random variable ω. In realistic settings, there may be myriad plausible relationships between x and ω. For example, economists have a bevvy of models predicting inflation and unemployment, and there is no consensus on the correct way to use any set of data to predict the future values of these variables. We capture this by writing the joint distribution of ω and x as a function of a random variable $m \in M$.[2] Denoting each relationship $\Pr(\omega \mid x, m)$ as a "model," the person trying to determine $\Pr(\omega \mid x)$ without knowing m therefore faces "model uncertainty."

If the person appreciated all feasible models and could estimate the likelihood of each, there is a straightforward solution: average over the answers given by each model. However, there are a variety of reasons to question the feasibility of this approach given the structure of the problem and the constraints of the human mind (Simon 1955). Even in a relatively simple problem like forecasting a future outcome given past data, just indexing the space of possible trends and error terms is daunting. Second, developing reasonable priors over the likelihood of each relationship would seem to require a strangely profound amount of ex ante knowledge. Third, constructing the procedure to determine the likelihood and expected value under each relationship is analytically challenging, and implementing this procedure may be computationally infeasible under cognitive constraints.[3]

We posit that people deal with these challenges by making assumptions that constrain the model space. These assumptions simplify the problem by assuming away some potentially true

* Augenblick: University of California, Berkeley (email: augenblick@berkeley.edu); Backus: University of California, Berkeley, NBER, and CEPR (email: backus@berkeley.edu); Little: University of California, Berkeley (email: andrew.little@berkeley.edu); Moore: University of California, Berkeley (email: dm@berkeley.edu).

† Go to https://doi.org/10.1257/pandp.20251109 to visit the article page for additional materials and author disclosure statement(s).

[1] A burgeoning literature considers how learning gets distorted (even in the long run) when actors make incorrect assumptions or have a "misspecified" model (Schwartzstein 2014; Esponda and Pouzo 2016; Bohren and Hauser 2021).

[2] See also Esponda and Pouzo (2016) (and citations therein) for a theory that treats models in a similar fashion, though they focus on long-run beliefs when the true model is not in the set considered.

[3] See Little and Pepinsky (2021) for an example of how a Bayesian can learn about both a parameter of interest and the bias in the estimate, though even in this setting there are strong assumptions about the relationship between that bias and the data-generating process.

relationships, but they are "reasonable" in the sense that they only ever consider potentially true relationships.[4] Sometimes these assumptions are explicit (as in formal statistical work), whereas sometimes they are implicit or implied by a procedure that produces an answer (as might happen when a person faces a complex problem). We then assume that people infer correctly based on this constrained view of potential models; that is, their belief is correct, conditional on their assumptions, but potentially unconditionally incorrect because it constrains the set of models.

In other words, the person acts as if they truly observed some piece of data that provided information about M—this is her *assumed information*. Now, when we write $\Pr(\omega \mid m, x)$, we can see that the belief ω conditions on the "realization" of two random variables. One is the real information (x), and the other is the assumed information (m).

To demonstrate some of the interesting implications of this framework, we develop a pair of toy models. First, when "new information" (or "thinking") entails considering new models, this will typically *reduce* confidence. This contrasts with the standard model, in which giving people more information increases their confidence (decreases subjective uncertainty), at least on average. Second, in a standard model of persuasion, speakers form correct beliefs about how others will respond to new information. When speakers assume different models than their audience, they will misunderstand what kinds of information will be persuasive.

In Augenblick et al. (2025), we present a more detailed and general version of the model and empirical evidence on how ignoring model uncertainty leads to overprecision and disagreement among people with the same information.

I. Examples

We illustrate the implications of this relatively abstract approach with two concrete examples sharing a common information structure.

A. *Information Structure*

Suppose that people learn about a random state of the world $\omega \in \{0,1\}$ from two pieces of information, $x_1 \in \{0,1\}$ and $x_2 \in \{0,1\}$. Each value of the state occurs with probability $1/2$. There are two possible models of the world $m \in \{1,2\}$, which are also each true with probability $1/2$. The model dictates the data-generating process for x_i; if model i is true, x_i is informative about ω. Formally, $\Pr(x_m = 1) = \pi > 1/2$, and so $\Pr(\omega = 1 \mid x_m = 1) = \pi$ and $\Pr(\omega = 1 \mid x_m = 0) = 1 - \pi$. When model i is true, the "other" piece of information x_{-i} is uninformative, taking on each value with probability $1/2$ independent from all other variables.

This could correspond to a scenario where x_1 and x_2 are news reports from different sources and $m = 1$ means 1 is a reliable source, while $m = 2$ means 2 is a reliable source.

B. *Information and Thinking Can Decrease Confidence*

Take a person who, at the outset, observes one piece of information; without loss of generality, let this be x_1. Rather than thinking through both potential models, the person assumes that one of them is true and picks each model with probability $1/2$.[5] His belief at this stage is

$$\Pr(\omega = 1 \mid x_1, m) = \begin{cases} \pi, & m = 1, x_1 = 1 \\ 1 - \pi, & m = 1, x_1 = 0 \\ 1/2, & m = 2 \end{cases}.$$

[4] Unlike many models of misspecified learning (e.g., Esponda and Pouzo 2016), people in our model observe exogenously generated data, and if those data fully reveal the model m, they will only consider that model.

[5] Note that this implies that models are selected according to their true distribution. In Augenblick et al. (2025), we show that this assumption leads to several tidy results about how not appreciating model uncertainty affects the variance and mean squared error of beliefs, and also discuss what happens under different assumptions about how models are selected.

There are two new pieces of "information" that we could give such an person. First, we reveal the "real information" x_2. Second, we could get him to consider the other model, or "loosen his assumption."

Upon receiving the second piece of real information, his belief is unchanged if the model he assumed was $m = 1$. If he assumed $m = 2$, the belief will change to π or $1 - \pi$ depending on the revelation of x_2. Since the variance of a binary random variable is maximized when each possibility occurs with probability $1/2$, the variance of the belief will either stay the same or decrease, and it will decrease on average. More generally, the law of total variance (LTV) implies that learning more information decreases variance (increases confidence) on average (Augenblick and Rabin 2021).

Prodding the person to consider another model—and, hence, form beliefs about ω, knowing that both $m = 1$ and $m = 2$ are equally likely—can increase or decrease the variance of his belief. If he initially considered model $m = 1$ and, hence, his belief was π or $1 - \pi$, then his belief will move toward $1/2$ as he considers the possibility that his information is not the relevant fact, increasing the variance. If he initially only considered model 2, thinking about model 1 will move his belief from $1/2$ to the midpoint between $1/2$ and either π or $1 - \pi$, which implies a lower variance.

However, the increased variance effect dominates on average. Formally, let $v(x) = x(1-x)$ be the variance of a binary random variable with probability x. Then the average variance when assuming a model is

$$\frac{1}{2}v\left(\frac{1}{2}\right) + \frac{1}{2}v(\pi),$$

while the variance when considering both models is

$$v\left(\frac{1}{2}\left(\frac{1}{2} + \pi\right)\right) > \frac{1}{2}v\left(\frac{1}{2}\right) + \frac{1}{2}v(\pi),$$

where the inequality follows from the concavity of v.

The logic here is the LTV in reverse. The LTV says that when learning the true model, the variance must decrease on average. Conversely, by "unlearning," or in this case, "unassuming" the true model, variance increases on average. (See Augenblick et al. 2025 for a more precise statement of how ignoring model uncertainty affects the variance of beliefs.)

More generally, if we consider the possibility that giving people new information includes not just giving new data but giving a new lens through which to interpret data, doing so will typically render them less confident on average (i.e., with a higher variance). Similarly, if "thinking" means considering other ways to interpret data rather than accessing more facts in memory, it may decreasee confidence as well.

C. *Disagreement*, *Persuasion*, *Deliberation*

Now suppose that there is a second person involved who also starts with one model and one piece of information. We do not specify a complete communication or deliberation model, but have in mind a situation where information is verifiable and costly to share and where people want others to hold accurate beliefs,[6] so they want to share information that they think will move others' beliefs toward the truth (and listeners will "believe" what they hear).

There are four possibilities of how person 2's knowledge relates to that of person 1:

(i) They have the same model and information and, hence, will agree about the probability of the state. So there is no direct value to communication or deliberation, but if both were to learn x_2, they would update the same way.

[6] Schwartzstein and Sunderam (2021) develop a related theory where a sender wants a receiver to adopt certain beliefs and proposes a model to interpret past data that leads to a favorable inference.

(ii) Person 2 has the same model m but different information. As a result, 1 and 2 hold different initial beliefs but agree about who has the useful information, and so the person with that information can share x_m. After doing so, they will hold the same belief about ω.

(iii) Person 2 has the same information but a different model. Here they will not agree in their belief about ω but may not understand why. If individuals recognize that they hold different models and can share them, this will lead to agreement. However, if they don't recognize that other models are possible, they may not know how to reconcile their disagreement.

(iv) The most interesting case is when person 2 has different information and a different model. If person 1 selects model 1 and person 2 selects model 2, they will hold the same belief if $x_1 = x_2$, but if they do not match, they will reach the opposite belief (e.g., π for 1, $1 - \pi$ for 2). If they want to move the other's belief toward their own, they may be tempted to share their information about x_i to get the other person to move their belief, but this will have no effect, because the other person already thinks that they possess the relevant information. If they share models but not information, they will partially converge as both will recognize that their information may not be relevant. However, they need to share both models and information to fully converge.

The broader point is that people who make different assumptions but are unaware of this will find it hard to predict how others will respond to new information. Combining with the previous analysis, if deliberation involves sharing information, it will tend to render all parties more confident, but if it involves sharing different models, it will tend to make them less confident.

II. Discussion

People in our theory assume a particular model and neglect alternative interpretations of the data. The human mind has limited computational capacity for understanding a world whose complexity is not similarly constrained. So people must at least sometimes neglect consideration of some possible models. But acknowledging this raises the question of what precisely it means to neglect alternative possible models. To what extent are people aware that they are neglecting alternative interpretations? Can people retrieve alternative models with effort? Or does the very acceptance of model assumptions foreclose the possibility of returning to reconsider alternative models (Vul et al. 2014; Gilbert 1991)?

Is it possible to simultaneously believe multiple models of the same set of observed facts? Some evidence suggests not. Visual perception entails conceptual models to account for visual data from our eyes. Some illusions, such as the Necker cube or the vase/face illusion, present an image that is amenable to multiple models. Human visual systems are generally unable to entertain two models simultaneously (Gregory 1997). We can see the vase or the face, but not both at the same time.

In this paper we explored some implications of a model in which people focus on one particular model, neglecting other possibilities. We develop this model more fully in our companion piece (Augenblick et al. 2025), exploring more predictions with an emphasis on overprecision disagreement and testing those predictions with experimental and observational data.

Scientific theories build models to account for empirical data. Those models must necessarily simplify the complexities of the world. Indeed, that is why they are useful. But, of course, because scientists only consider a subset of all possible models, they run the risk of being wrong and not realizing it. As a result, they will, on average, be too sure they have selected the correct model. Recognizing this possibility, we offer our theory with humility and in the hope that it might prove useful.

REFERENCES

Augenblick, Ned, and Matthew Rabin. 2021. "Belief Movement, Uncertainty Reduction, and Rational Updating." *Quarterly Journal of Economics* 136 (2): 933–85.

Augenblick, Ned, Matthew Backus, Andrew T. Little, and Don A. Moore. 2025. "Model Uncertainty and Overprecision: Theory and Evidence." Paper presented at 2025 ASSA Annual Meeting, San Francisco, January 5.

Bohren, J. Aislinn, and Daniel N. Hauser. 2021. "Learning with Heterogeneous Misspecified Models: Characterization and Robustness." *Econometrica* 89 (6): 3025–77.

Esponda, Ignacio, and Demian Pouzo. 2016. "Berk–Nash Equilibrium: A Framework for Modeling Agents with Misspecified Models." *Econometrica* 84 (3): 1093–130.

Gilbert, Daniel T. 1991. "How Mental Systems Believe." *American Psychologist* 46 (2): 107–19.

Gregory, Richard L. 1997. "Visual Illusions Classified." *Trends in Cognitive Sciences* 1 (5): 190–94.

Little, Andrew T., and Thomas B. Pepinsky. 2021. "Learning from Biased Research Designs." *Journal of Politics* 83 (2): 602–16.

Schwartzstein, Joshua. 2014. "Selective Attention and Learning." *Journal of the European Economic Association* 12 (6): 1423–52.

Schwartzstein, Joshua, and Adi Sunderam. 2021. "Using Models to Persuade." *American Economic Review* 111 (1): 276–323.

Simon, Herbert A. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics* 69 (1): 99–118.

Vul, Edward, Noah Goodman, Thomas L. Griffiths, and Joshua B. Tenenbaum. 2014. "One and Done? Optimal Decisions from Very Few Samples." *Cognitive Science* 38 (4): 599–637.