December 2, 13

Gilbert J. Chin,

Senior Editor, Science

Dear Dr. Chin,

Please find enclosed a manuscript titled “A Computational Model of Hyperbole and Nonliteral Language Understanding” for consideration in Science. My co-authors for this work are Jean Wu, Leon Bergen, and Noah Goodman.

Nonliteral language is ubiquitous in everyday communication. Ranging from metaphor, sarcasm, to hyperbole, people constantly produce utterances that are not meant to be interpreted literally. A successful listener needs to go beyond the literal meaning of an utterance in order to uncover the speaker’s intended meaning. People are incredibly adept at this; we can easily detect when an utterance is not meant to be taken literally and infer what it actually means. In addition, speakers do not always use language simply to describe the objective state of the world, but also to communicate their subjective opinions and emotions. These aspects of language understanding—figurative use and affective subtext—are critical areas of research. While they have been studied extensively across many fields, including psychology, linguistics, and literature, the work has been largely qualitative. On the other hand, there has been a great deal of work on applying formal models However, there are very few, if any, formal models that build upon core theories of communication to explain these rich phenomena in language understanding.

A significant step towards modeling more flexible and richer uses of language.

People are very good at reasoning about each other’s emotions (an ability we term *affective cognition*); we intuitively apply knowledge about how emotions “work” to understand people’s reactions to emotion-inducing events. Over recent decades, an enormous amount of research in affective science has described how such emotion attribution, as well as emotions more generally, operate. Critically, however, this work is largely qualitative, and there are few, if any, formal models that can track and predict real-world affective cognition. Conversely, in recent years computational models, especially those taking a Bayesian approach, have been used to successfully explain human inference in a number of domains. Such models have not yet to be applied to the critical domain of emotion.

Here we address this gap in psychological knowledge and make the first attempt to integrate affective cognition into computational models of reasoning---providing a working model of people’s “lay theories” of emotion. As a first step, we studied reasoning about emotional reactions to situations. We theorized that people consider only a low-dimensional summary of the situation; drawing on the extensive literature in affective science and behavioral economics, we proposed that this summary includes the amount won, the prediction error, as well as loss aversion. We further proposed that these form an integral part of people’s lay theories, facilitating complex inferences rather than a simple stimulus-response association.

To experimentally verify this, we designed a novel experimental paradigm where participants watch a character play a gamble (on a gameshow). We manipulated the situation parameters quantitatively, and measured participants’ attribution of emotion (“How did Bob feel after winning $X?”). A model built on participant responses verified that participants implicitly considered the low dimensional outcome summary predicted above. We then tested the relationship predicted by a Bayesian analysis for “reverse” inferences (“Given that Bob feels Y, what did Bob win?”). This model predicted participant responses from an independent sample with a high accuracy (of ~76%). This attests to the complexity and flexibility of intuitive lay theories to perform complicated computations.

In summary, this study (i) demonstrates that people’s lay theories of emotion implicitly take into account a prediction error signal from situation outcomes, (ii) further demonstrates that people can flexibly use their lay theories to do backward inferences, and (iii) opens up many avenues of future quantitative research on affective cognition.

This work has the potential to fundamentally shift the landscape of emotion science, by leveraging tools from computational modeling to provide a new window into the computations that support affective cognition. The progress provided by this work is multi-disciplinary in nature: derived from insights from affective science, decision theory, and behavioral economics. It further holds far-reaching implications across multiple disciplines. For instance, this could help us design and build emotional computers and robots that are more effective in interacting with people (i.e. *affective computing*). Our work could also inform the phenotyping and treatment of deficiencies in affective processes that underlie certain affective disorders (i.e. *computational psychiatry*), which is in line more broadly with the recent Research Domain Criteria (RDoC) initiative by the National Institute of Mental Health. Hence, we believe these findings will be of broad interest across the fields of psychology, cognitive science, computer science, and behavioral economics, as well as to the general public, and is ideally suited to the *Science* readership.

We hope you find this work to merit further scrutiny by additional experts. If you do, there are many cognitive scientists and psychologists whose experience and expertise are suited to evaluating this manuscript, including:

• Lisa Feldman-Barrett (Northeastern University; an expert on emotion)

• Rebecca Saxe (MIT; an expert in social cognition and theory of mind)

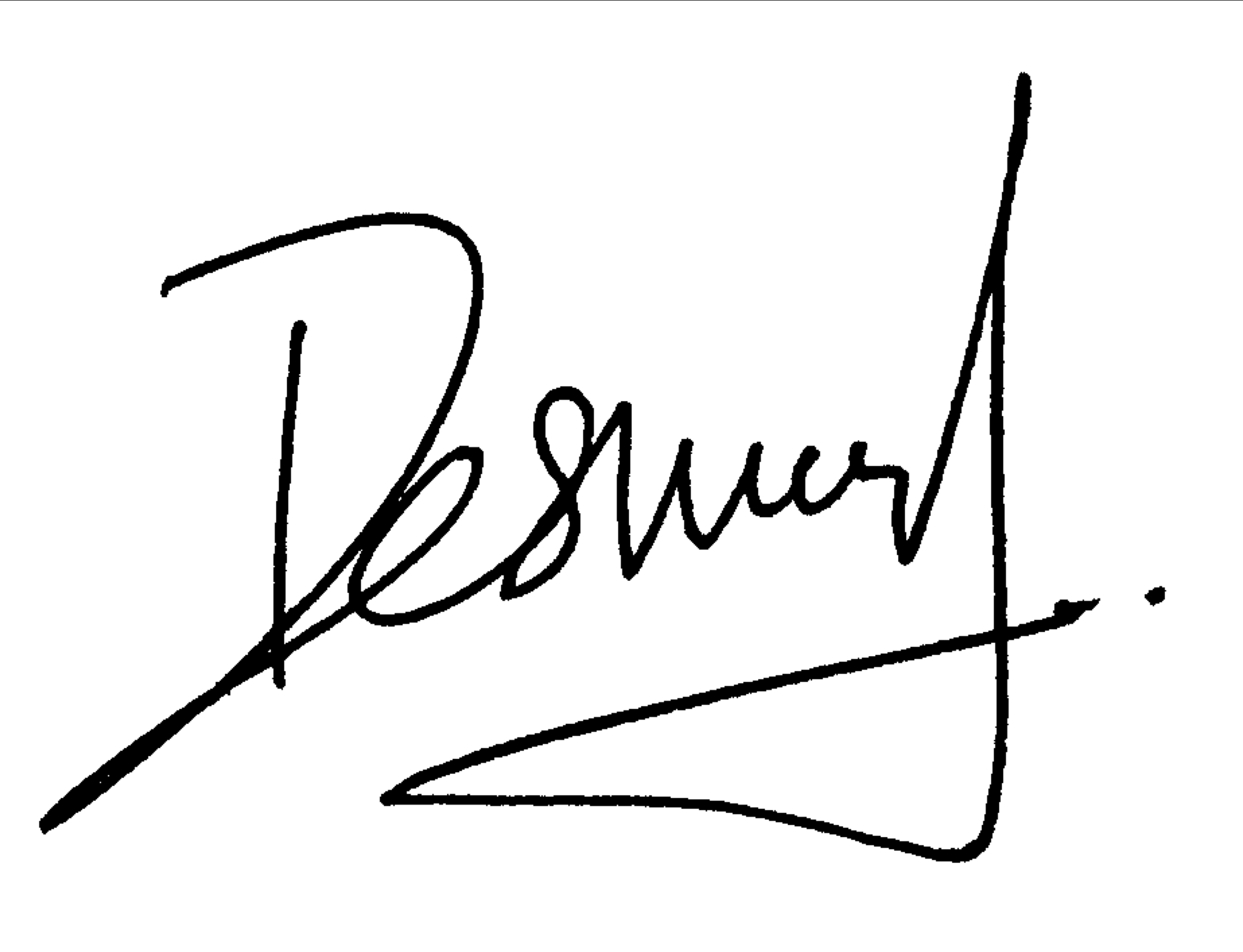
• Alison Gopnik (Berkeley; an expert in intuitive theories and cognition)

• Liz Phelps (NYU; an expert in emotional learning and decision making)

• Nick Chater (Warwick Business School; an expert in computational modeling of cognitive processes and decision making)

I will be the corresponding author and will assume responsibility for informing my co-authors of all progress through the review process. If you have any questions or require any clarifications about this work, please do not hesitate to contact us.

Sincerely,



Desmond C. Ong

Ph.D. Student,

Department of Psychology,

Stanford University