STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

Desmond C. Ong
Stanford University
Department of Psychology
Jordan Hall, Building 420
Stanford, CA 94305

June 2, 2013

Inder M. Verma,
Editor-in-Chief, PNAS

Dear Dr. Verma,

Please find enclosed a manuscript entitled "Exploring Lay Theories of Emotion" for consideration in PNAS. My co-authors for this work are Jamil Zaki, and Noah Goodman (all at Stanford University).

We are very good at reasoning about other people's emotions; we easily use our observations and our intuitive theories of how emotions "work" to predict people's reactions. This is but one component of our ability to understand others' minds – their beliefs and goals, for instance – usually called our *Theory of Mind*, that allows us to interact with other people.

In recent years, there have been successes by many groups in using formal, computational models to study how we reason about others' beliefs and goals. These models traditionally have not tried to explain the role of emotions in the decision making process. At the same time, there has been enormous amounts of progress in affective science in describing how emotion "works", and the types of inferences that we can draw from others' emotions, yet there has not been an attempt to build a formal model that could, for example, predict future behavior.

This work is, to the best of our knowledge, the first attempt to address this gap in knowledge by integrating emotion into computational models of reasoning and decision making, and lays the foundation for what might be called a computational "Theory of Emotion." To begin to study lay theories of emotion, we first examine two requisite properties for emotion concepts that allow them to be used in a lay theory: they have to be *coherent* (emotion concepts are differentiable from non-emotion concepts in that they directly correspond to a prediction error signal), and *flexible* (consistent relationships between events and emotions that allow reasoning forward from events to emotions and backwards, from emotions to events). We then propose a simple model of emotional reactance to an outcome, and we test the model predictions in a simple "game-show" experiment, where participants watch a stick figure character plays a simple gamble. We then trained the model on forward attributions of emotions (i.e. "How did Bob feel after winning \$X?") and had it make backward inferences of emotions (i.e. "Given that Bob feels Y, what did Bob win?"); the model was able to predict participant responses from an independent sample with a very high accuracy (R^2 almost 80%).

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 research in developing more comprehensive models and examining more carefully how we make the types of inferences we do.

This work is multi-disciplinary in nature: for example, the theoretical framework for our model is derived from insights from affective science, decision theory, and behavioral economics. 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.

Because of our approach to using formal tools to study how people reason about emotion, we feel that **Walter Mischel (Columbia University)** would be an ideal choice as editor.

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)
- Tom Gilovich (Cornell University; an expert on judgment and decision making)
- Nancy Kanwisher (MIT; an expert in cognitive neuroscience)

I will be the corresponding author and will assume responsibility for informing my co-author 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
dco@stanford.edu