

**COMMITTING TO PROMISES BY GUILT:  
A GENERALIZED APPROACH**

**YUSUKE NARITA<sup>\*1</sup>**

*Graduate School of Economics, University of Tokyo and VCASI*

**ABSTRACT**

We propose a new dynamic theory of “guilt” based on dynamic psychological game theory. It is shown that this theory can explain a surprisingly wide range of experimental results signifying a mechanism of emergence of trust through communication. We also study theoretical properties of sequential equilibrium when guilt defined by this theory is present.

**KEYWORD:** Trust, communication, guilt, lying-aversion, psychological game theory

---

<sup>\*1</sup> Address: 7-3-1, Hongo, Bunkyo-Ku, Tokyo, Japan, 113-0033. E-mail: [allofsudden@gmail.com](mailto:allofsudden@gmail.com). This paper is based on my undergraduate thesis at the University of Tokyo. I am very grateful to Hitoshi Matsushima for encouragement and guidance. I also thank Masahiko Aoki, Katsuhito Iwai, Kazuya Kamiya, Mamoru Kaneko, Toshiji Kawagoe, Akihiko Matsui, Takuya Onoda, and Tomomi Tanaka for comments.

## 1. INTRODUCTION

This paper gives a consistent explanation of a wide range of experimental results signifying a mechanism of emergence of trust through communication. (See, for example, Charness and Dufwenberg (2006, 2008a, 2008b), Ellingsen and Magnus (2004), Ellingsen *et al.* (2009), Gneezy (2005), Hurkens and Kartik (2008), Lundquist *et al.* (2009), Sanchez-Pages and Vorsatz (2009), and Vanberg (2008).) For this purpose, a new dynamic theory of “guilt” based on dynamic psychological game theory (Geanakoplos *et al.* (1989) and Battigalli and Dufwenberg (2009)) is proposed. According to this theory, people feel some disutility (guilt) when they choose a strategy which would betray some other person’s expectation, *with that expectation being generated by his very own actions.*

## 2. RESULTS

We find that this theory can explain a surprisingly wide range of the experimental results above. Then it is shown that the framework on which this theory is built includes both of the two extant theories of guilt (“expectation-based” theory (See, for example, Battigalli and Dufwenberg (2007, 2009)) and “action-based” theory (See, for example, Ettinger and Philoppe (2009) and Kartik (2009)) as special cases, while this theory remains free from an experimental criticism of expectation-based theory (Ellingsen *et al.* (2009)). Moreover, when guilt defined by this theory is present, sequential equilibrium (Kreps and Wilson (1982) and Battigalli and Dufwenberg (2009)) has several desirable properties and have more realistic efficiency implication that cannot be derived by expectation-based theory.

## 3. CONCLUSION

Our bottom line is that the two extant theories, which have been regarded as mutually exclusive ones, should be thought not as substitutive but as complementary in the sense that others’ “expectations” can urge trustful, reciprocal and lying-averse behavior in an individual only when these expectations were generated by his own “actions”. This further suggests that, when analysing functions of beliefs in games, it is often important to explicitly specify not only whose beliefs they are but also who or what generates them, i.e., sources of beliefs.

## REFERENCES

Battigalli, Pierpaolo, and Dufwenberg, Martin. 2007. Guilt in Games. *American Economic Review*.

- Battigali, Pierpaolo, and Dufwenberg, Martin. 2009. Dynamic Psychological Games. *Journal of Economic Theory* (Forthcoming).
- Charness, Gary, and Dufwenberg, Martin. 2006. Promises and Partnership. *Econometrica*.
- Charness, Gary, and Dufwenberg, Martin. 2008a. Broken Promises: An Experiment. Mimeo.
- Charness, Gary, and Dufwenberg, Martin. 2008b. Contracts and Communication. Mimeo.
- Ellingsen, Tore, and Johannesson, Magnus. 2004. Promises, Threats, and Fairness. *The Economic Journal*.
- Ellingsen, Tore, Johannesson, Magnus, Tjotta, Sigve, and Torsvik, Gaute. 2009. Testing Guilt Aversion. *Games and Economic Behavior*.
- Ettinger, David., and Jehiel, Philippe. 2009. A Theory of Deception. *American Economic Journal: Microeconomics*.
- Geanakoplos, John, Pearce, David, and Stachetti, Ennio. 1989. Psychological Games and Sequential Rationality. *Games and Economic Behavior*.
- Gneezy, Uri. 2005. Deception: The Role of Consequences. *American Economic Review*.
- Hurkens, Sjaak, and Kartik, Navin. 2008. Would I Lie to You? On Social Preferences and Lying Aversion. *Experimental Economics*.
- Kartik, Navin. 2009. Strategic Communication with Lying Cost. *Review of Economic Studies*.
- Kreps, David and Wilson, Robert. 1982. Sequential Equilibria. *Econometrica*.
- Lundquist, Tobias, Ellingsen, Tore, and Johannesson, Magnus. 2009. The Aversion to Lying. *Journal of Economic Behavior and Organization* (Forthcoming).
- Sanchez-Pages, Santiago, and Vorsatz, Marc. 2009. Enjoy the Silence: An Experiment on Truth-Telling. *Experimental Economics*.
- Vanberg, Christoph. 2008. Why Do People Keep Their Promises? An Experimental Test of Two Explanations. *Econometrica*.