

# Communication and Guilt in a Prisoner's Dilemma

Topi Miettinen and Sigrid Suetens

Journal of Conflict Resolution: vol.52, No.6, 945-960 (2008)

## 1 導入

### 1.1 Communication

- 社会的ジレンマを扱った研究において、事前のコミュニケーションが協調的な振る舞いを促進させることが明らかになっている (Ledyard 1995; Sally 1995; Ostrom 2006; Bicchieri and Lev-On 2007)
  - 社会心理学者、行動経済学者、社会学者はこの効果の背後にある人の心理を明らかにしたいと考えている
- 研究者たちはコミュニケーションがもたらす効果について一貫した見解を示している
  - コミュニケーションはどのように振舞うべきかについて協定を結ぶことを可能にする
    - \* 例えば、良い意思疎通（コミュニケーション）は約束を生み出し、協調しようという気持ちにさせる (Orbell, van de Kragt, and Dawes 1988; Kerr and Kaufman-Gilliland 1994)
    - \* コミュニケーションは規律を生み出し、お互いに共通した信念を植えつける (Bicchieri and Lev-On 2007)
- 本研究では事前のコミュニケーションと罪悪感の関係に着目する
  - 実験参加者は見知らぬ人とお金を賭けて1度きりの囚人のジレンマゲームを行う
  - ゲームを行う前にコミュニケーションの機会を設ける（相手とコミュニケーションを取るか否かを選択出来る）
  - 選択肢を決定したところで罪悪感の評定を2度行う
    1. 相手が選んだ選択肢をフィードバックされる前
    2. フィードバックされた後

### 1.2 Guilt

- Guilt(罪意識)は社会心理学者にとって仲裁感情と見なされている
  - 罪意識の重大な役割とは犯罪を起こさないという気持ちを引き起こし、公共の規律を守ろうとさせることである (Baumeister, Stillwell, and Heatherton 1994)
- 囚人のジレンマ研究において、コミュニケーションによる協調の約束から生まれる「規律」に焦点を当てるのは自然である
- 近年、社会心理学や経済学の分野において、罪意識の重要性に関する多くの研究がなされている
  - 人間は約束を交わすことにより、相手と何を選択するのかの情報を共有し、その約束を破って相手を失望させることを拒む (Charness and Dufwenberg 2006)

- 社会的ジレンマゲームにおいて、相手を裏切って罪意識を持った人はそれ以降、より協調を選択するようになり (Ketelaar and Au 2003)、相手の裏切りに対して報復をしなくなる (Hopfensitz and Reuben 2007)
  - 本研究では「こちらだけ裏切り (=自分が裏切り、相手は協調)」の罪悪感に着目する
  - 結果
    - 裏切った人は相手とコミュニケーションを取ることで協調し合うと約束した場合に罪悪感を感じる
    - ペナルティは裏切ることによる罪悪感を軽減させる

## 2 Design, Procedure, and Hypotheses

## 2.1 実験

- 参加者
    - 大学生 140 名
      - \* 様々な専門分野から募集
  - 環境
    - 見かけの上独立したコンピュータターミナル
  - 課題
    - 2人制, 1度きりの囚人のジレンマゲーム (利得表は Table 1<sup>1</sup>)
      - \* 選択肢は「Option 1(裏切りに対応)」「Option 2(協調に対応)」

**Table 1**  
**Payoffs in the Prisoner's Dilemma Game**

	Your Payoff	The Other's Payoff	
Player A	Player B		
		Option 2(C)	Option 1(D)
Option 2(C)	10	10	16
Option 1(D)	16	4	6

- 条件
    - 「こちらだけ裏切り」だった場合に相手に支払うペナルティ(Ta)
      1. ペナルティなし
      2. 0.20 ユーロのペナルティ
      3. 0.20 ユーロのペナルティ + 画面にペナルティの存在を表示
      4. 1 ユーロのペナルティ + 画面にペナルティの存在を表示

<sup>1</sup>単位はユーロ

**Table 2**  
**Treatments and Number of Participants**

	Penalty	# Observations
Treatment 1	No penalty	44
Treatment 2	Penalty 0.20	32
Treatment 3	Penalty 0.20 and reminder screen	32
Treatment 4	Penalty 1 and reminder screen	32

### • 手順

1. 教示
  - ペナルティに関する記述のみ，条件によって異なる
2. 教示の理解度チェック
3. 相手とコミュニケーションを取るか否かの選択
  - 参加者は以下の 2 つの中，いずれかを選択する
    - (a) 相手に「私はお互い協調したい」と伝える
    - (b) コミュニケーションを取らない
4. 「Option 1」か「Option 2」を選択
5. 予想のアンケート
  - (a) First-order-belief : 相手が何を選択すると思うか？
  - (b) Second-order-beliefs : 相手は自分が何を選択すると予想したと思うか？
6. 罪悪感の評定：1 回目
  - 7 点スケール (0:not at all ~ 6:very intensely)
  - 研究意図を悟られないため，恥ずかしさ，嬉しさなどの罪悪感以外についても評定
7. 相手の選択肢をフィードバック
8. 罪悪感の評定：2 回目

## 2.2 仮説

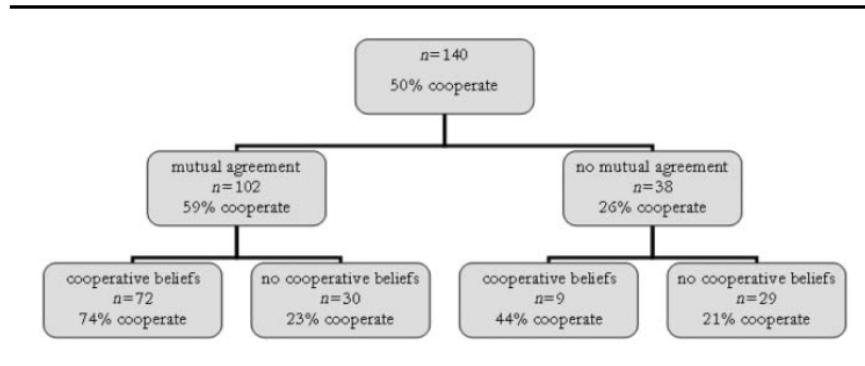
- 仮説 1：罪悪感は主に「こちらだけ裏切り」と分かった時に高くなる
  - 罪悪感を相手の選択肢をフィードバックする前と後の 2 度測定することで以下 3 つを検討出来る
    1. 自分が裏切った時の罪悪感
    2. 自分は裏切り，相手も裏切ったと知った時の罪悪感
    3. 自分は裏切り，相手は協調したと知った時の罪悪感
  - 裏切った人は「こちらだけ裏切り」の時にのみ罪悪感を高めるという先行研究の追試 (Bicchieri and Erte 2007; Lopez-Perez forthcoming; Miettinen 2006)
- 仮説 2：コミュニケーションを取り合って協調の約束をした時はコミュニケーションを取り合わなかった時よりも「こちらだけ裏切り」によって生み出される罪悪感が高くなる
  - お互い協調を促すコミュニケーションを行うと，約束を守ろうとするため協調を選択するだろう
  - その約束を破る行為は罪悪感を高めると考えられる
- 仮説 3：「こちらだけ裏切り」の時に科させるペナルティがあると，ペナルティが無い時に比べて「こちらだけ裏切り」による罪悪感は低くなる
  - 経済学の理論である crowding-out 効果<sup>2</sup>に由来
  - また，ペナルティは裏切りを誘発すると考えられる

<sup>2</sup>政府が資金需要をまかなうために大量の国債を発行すると，それによって市中の金利が上昇するため，民間の資金需要が抑制されること

### 3 Results

#### 3.1 Cooperation, Communication, and Beliefs

**Figure 1**  
**Overview of Cooperation Rates**



- Figure 1 は全体的な協調割合を示している
  - 1段目：全参加者中，50% が協調
  - 2段目：相手と「お互い協調しよう」というコミュニケーションを取り合った人と取り合ってない人を分けている
    - \* mutual agreement は 59% 協調しているが，no mutual agreement は 26% しか協調していない（有意差あり， $\chi^2$  検定  $p = .001$ ）
  - 3段目：予想のアンケートで「相手が何を選択すると思うか？」と「相手は自分が何を選択すると予想したと思うか？」で両方とも「Option 2(協調)」と答えた人<sup>3</sup>とそれ以外の回答をした人を分けている
    - \* cooperative beliefs の割合は mutual agreement の 102 名中 72 名に対して，no mutual agreement の 38 名中たったの 9 名（有意差あり， $\chi^2$  検定  $p = .000$ ）
    - \* 各条件の協調度合いの比較
      - (M,C)<sup>4</sup> : 74% > (M,N) : 23%       $\chi^2$  検定  $p = .000$
      - (M,C) : 74% > (N,C) : 44%       $\chi^2$  検定  $p = .000$
      - (M,N) : 23%      (N,N) : 21%

**Table 3**  
**Effects of Communication on Beliefs**

	Self-Communicate	Other Communicates		
	Estimate ( $p$ value)	Estimate ( $p$ value)	Log L	Pseudo R <sup>2</sup>
First-order beliefs	0.17 (.153)	0.55 (.000)***	-7.52	.13
Second-order beliefs	0.26 (.054)*	0.21 (.128)	-85.48	.04

Note: Marginal effects are reported. Standard errors are robust for potential within-pair dependency. The number of observations is 140.

\*significant at 10% level; \*\*\*significant at 1% level (two-tailed tests).

- Table 3 はコミュニケーションと予想アンケートとの重回帰分析を行った結果である

<sup>3</sup>すなわち，自分も相手も協調すると信じあっていると思っている人

<sup>4</sup>(M,C) : mutual agreement & cooperative beliefs , ( , ) の左をコミュニケーション，右に予想アンケート

- Self-Communicate : 自分が相手にコミュニケーションを要求したか否か
  - Other Communicates : 相手が自分にコミュニケーションを要求したか否か
  - First-order-beliefs : 相手の選択肢の予想
  - Second-order-beliefs : 相手が思っている自分の選択肢の予想
- Table 3 から分かること
    - 相手の選択肢の予測は相手がコミュニケーションを要求してきたかどうかに起因している
      - \* すなわち、相手がコミュニケーションを要求してきた場合、相手は協調したいと考えているはずと思い、相手は「協調」を選択すると考えるということ
    - 相手が思っている自分の選択肢の予想は自分がコミュニケーションを要求したかどうかに起因している
      - \* すなわち、自分がコミュニケーションを要求するということは相手に自分が協調したいと考えているはずだと思わせるため、相手は自分の行動を「協調」だと予測するだろうと考えるということ

**Table 4**  
**Cooperation and Communication Rates across Treatments**

Treatment	Penalty	Cooperation Rate	Communication Rate	# Observations
1	No	45% (20)	93% (41)	44
2	Yes	53% (17)	78% (25)	32
3	Yes	56% (18)	84% (27)	32
4	Yes	47% (15)	84% (27)	32
Average 2, 3, 4	Yes	52% (50)	82% (79)	96
Average all	—	50% (70)	86% (120)	140

- Table 4 は協調を選択した人の割合とコミュニケーションを選択した人の割合をペナルティの条件に分けて表したものである
  - 協調割合、コミュニケーション割合ともに条件で差がない ( $\chi^2$  検定 協調 :  $p = .585$  コミュニケーション :  $p = .119$ )
  - すなわち、ペナルティがあると裏切りを誘発するという先行研究 (Gneezy and Rustichini 2000b) に反する結果となった
    - \* しかし、ペナルティは行動(協調やコミュニケーション)に差を与えていないが、罪悪感には影響を与えているはずである

### 3.2 Guilt, Norms, and Crowding-Out

- この章では前述した以下の 3 つの仮説を検証する
  1. 罪悪感は主に「こちらだけ裏切り」と分かった時に高くなる
  2. コミュニケーションを取って相手と協調する約束を取り付けた時はコミュニケーションを取らなかった時よりも「こちらだけ裏切り」によって生み出される罪悪感が高くなる
  3. 「こちらだけ裏切り」の時に科させるペナルティがあると、ペナルティが無い時に比べて、「こちらだけ裏切り」による罪悪感は低くなる

Table 5  
Median Guilt Intensities

		After Feedback	
	Before Feedback	Other Defects	Other Cooperates
Participant defects			
No mutual agreement	0 <i>n</i> = 28	0 <i>n</i> = 20	1 <i>n</i> = 8
Mutual agreement	2 <i>n</i> = 42	0 <i>n</i> = 20	2 <i>n</i> = 22
Participant cooperates			
No mutual agreement	0 <i>n</i> = 10	0 <i>n</i> = 8	0 <i>n</i> = 2
Mutual agreement	0 <i>n</i> = 60	0 <i>n</i> = 22	0 <i>n</i> = 38

- Table 5 は「相手の選択肢」と「コミュニケーションの成立、未成立」と「相手の選択肢をフィードバックされる前と後」を分け、それぞれの罪悪感の数値<sup>5</sup>を示した表である（詳細な結果は Appendix B と Appendix C）
- 仮説 1(罪悪感は主に「こちらだけ裏切り」と分かった時に高くなる)
  - Table 5 の右上「After Feedback, Other Cooperate, Participant defects」の罪悪感の数値が「Other Defects」よりも高いことから、仮説 1 は確証
- 仮説 2(コミュニケーションを取り合って協調の約束をした時はコミュニケーションを取り合わなかつた時よりも「こちらだけ裏切り」によって生み出される罪悪感が高くなる)
  - Table 5 の 2 行目「Mutual agreement, Participant defects」が「No mutual agreement」よりも高いことから、仮説 2 は確証

Table 6  
Estimated Coefficients of Six Regressions of Guilt (Including *p* Values)

	(1) <i>n</i> = 70	(2) <i>n</i> = 30	(3) <i>n</i> = 15
Guilt before feedback			
Mutual agreement	0.78 (.006)**	—	—
Guilt after feedback			
Other cooperates	1.38 (.000)**	—	—
Mutual agreement	—	0.74 (.045)*	1.16 (.042)*
Penalty	—	-0.93 (.017)*	-0.71 (.116)

Note: (1) = given defection; (2) = given defection and other's cooperation; (3) = given defection and other's cooperation by surprise.

\*significant at the 5% level; \*\*significant at the 1% level (one-tailed tests); *p* values are in parentheses.

- Table 6 は 6 变量プロビット回帰分析 (six univariate ordered probit regressions) の結果である
  - 1 列目：裏切りを選択した人 (70 名)
  - 2 列目：自分は裏切り、相手は協調だった人 (30 名)
  - 3 列目：相手は裏切ると思い、裏切りを選んだが、相手は協調を選択してきた人 (15 名)
- Table 6 から分かること

<sup>5</sup>罪悪感の数値は中央値

- 仮説 1
  - \* 1列目の「Other cooperate」行より、有意な正の相関があることから仮説 1 は確証
- 仮説 2
  - \* 1列目及び2列目の「Mutual agreement」行より、有意な正の相関があることから、仮説 2 は確証
- 仮説 3(「こちらだけ裏切り」の時に科させるペナルティがあると、ペナルティが無い時に比べて、「こちらだけ裏切り」による罪悪感は低くなる)
  - \* 2列目の「Penalty」行より、有意な負の相関があることから、ペナルティが「こちらだけ裏切り」による罪悪感を減少させることができ、仮説 3 は確証

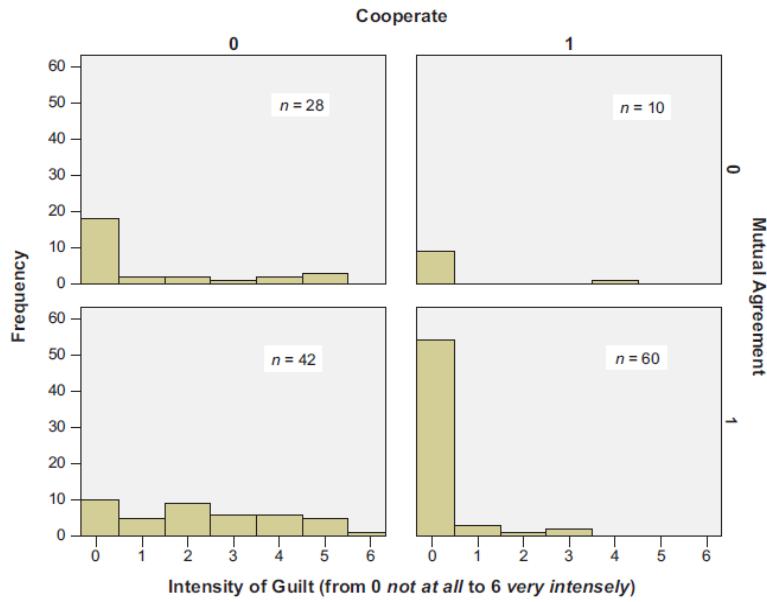
### 3.3 Guilt as a Counterfactual Emotion

- 経済学において、プレイヤーは罪悪感を避けようとする（協調する）と言われている（Dufwenberg 2002; Battigalli and Dufwenberg 2007）
- また、社会心理学において、罪悪感は反事実的（counterfactual）であると考えられている（Mandel and Dhami 2005）
  - Table 5 の相手に裏切られると全く罪悪感がないという結果は罪悪感の反事実的な一端を示している
    - \* 裏切りは罪悪感を生み出し、また罪悪感を避けようと協調しようとする
- 囚人のジレンマゲームの選択は罪悪感にのみ依存しているわけではない
  - しかし、コミュニケーションや選択肢によって罪悪感に差が生じたという結果は驚くべきことである
- 相手も裏切ると思い裏切ったが、実際は相手が協調してきたという「予期せぬこちらだけの裏切り」は罪悪感を引き起こすと考えられる
  - Table 6 の 3列目の「Mutual agreement」行に有意な正の相関がある
  - しかし、「Penalty」行には有意差は認められなかった

## 4 Conclusion

- 本研究は1度きりの囚人のジレンマゲームを用い、コミュニケーションを許可した状況における罪悪感に焦点を当てた
- そして、罪悪感と選択肢、及び罪悪感とコミュニケーションの有無との関係を検討した
  - 罪悪感を測ることで、コミュニケーションを取り合うと、約束を守ろうという意識が生まれるかどうかを明らかにしようとした
- 結果として
  1. 裏切ると罪悪感が高まり、特に「こちらだけ裏切り」の時に非常に高くなる
  2. 「こちらだけ裏切り」による罪悪感はコミュニケーションがあった時により高まる
    - 加えて、コミュニケーションによって「お互い協調」の約束が生まれ、その約束を破ることを避けるために協調を選択する割合が高くなつた
  3. 「こちらだけ裏切り」による罪悪感はペナルティがある時により低くなる
    - これはペナルティが自分に裏切られた相手に対するお詫びとなっており、その結果、相手を裏切った罪悪感が減少されたと解釈できる
- 本研究はコミュニケーションを伴う社会的ジレンマのプレイ中とプレイ後に感情（罪悪感）を測定するという今までに無い研究である
- また、人が取る選択パターンの背後にある根拠を明らかにするための指標として、感情を測定することの有用性を示した

## Appendix B



## Appendix C

