Why Do Oaths Work?

Sorravich Kingsuwankul
Chloe Tergiman
Marie Claire Villeval

Discussion by Ajalavat Viriyavipart

The Announcement Game

- **Game Structure:**
  - **Nature** draws product quality for a set of three cards, each has a probability 0.5 of displaying a star, which indicates a successful project.
  - **Advisor** observes and sends **unverifiable message** (0–3 stars) to investor.
  - **Investor** chooses whether to **invest** or **not**.
- **Payoffs**:
  - **If no investment**: both earn fixed payoff (30 for advisor and 100 for investor).
  - **If invest**:
    - **Advisor** gets commission (total earnings = 230).
    - **Investor's** payoff depends on a draw from 3 cards.
      - If a star is drawn, earns 300.
      - If a blank card is drawn, earns 30.

Treatments

- **Baseline:** No-Oath
- **Oath treatments**: "I swear upon my honor that during this experiment I will behave honestly and I will always tell the truth."
  - Public-Oath: common knowledge
  - Private-Oath: known only to advisor
  - Private-Oath^{++}: impossible for the experimenter to link subjects to their decisions
- **Only investor treatments**:
  - Necessary condition: use empirical distribution of the advisor in the Public-Oath treatment but doesn't know about the oath
  - Sufficient condition: know about oath and match with advisor in the Public-Oath treatment but choose distribution similar to the Private-Oath treatment

Main Results – Advisors

Table 1: Detectable and deniable lies in the No-Oath, Public-Oath, and Private-Oath treatments

Treatments	(1) No-Oath	(2) Public-Oath	(3) Private-Oath	(1-2)	(2-3) p-value	(1-3)
All A players						
% Player A who always tell the truth	9.2%	74.0%	39.7%	$p < 0.001$	$p < 0.001$	$p < 0.001$
% Untrue announcements	52.3%	10.8%	27.7%	$p < 0.001$	$p < 0.001$	$p < 0.001$
Relative frequency						
Detectable lies	25.2%	3.9%	9.6%	$p < 0.001$	$p = 0.002$	$p < 0.001$
Deniable lies	57.4%	13.8%	37.9%	$p < 0.001$	$p < 0.001$	$p = 0.001$
Absolute frequency						
Detectable lies	71.1%	11.7%	32.9%	$p < 0.001$	$p = 0.002$	$p < 0.001$
Deniable lies	86.8%	23.4%	56.2%	$p < 0.001$	$p < 0.001$	$p < 0.001$
Among A players who lie at least once						
% Untrue announcements	57.6%	41.4%	46.0%	$p = 0.010$	$p = 0.397$	$p = 0.004$
Relative frequency						
Detectable lies	27.8%	15.0%	15.9%	$p = 0.007$	$p = 0.593$	$p = 0.003$
Deniable lies	63.2%	53.2%	62.8%	$p = 0.302$	$p = 0.316$	$p = 0.884$
Absolute frequency						
Detectable lies	78.3%	45.0%	54.5%	$p = 0.004$	$p = 0.479$	$p = 0.008$
Deniable lies	95.7%	90.0%	93.2%	$p = 0.334$	$p = 0.660$	$p = 0.568$

Notes: Relative frequency statistics show the average frequency with which participants engage in any particular type of lie. Absolute frequency statistics show the proportion of A players who make any particular type of lie at least once. The relative frequency of detectable lies corresponds to how often A players who saw fewer than three stars announced three stars. The relative frequency of deniable lies corresponds to how frequently A players who saw fewer than two stars reported two stars or announced one star while seeing none. This table presents data both for all A players (top panel) and for the subset who make at least one lie (lower panel).

Suggestions

- Avoid requiring readers to read footnotes to decode key metrics.
  - "% of lies per opportunity (relative)"
  - "% of advisors who lied at least once (absolute)"
- Add number of subjects or observations
- More details of lying behavior (maybe in the appendix)

Observed \ Announced	0 Star	1 Star	2 Stars	3 Stars
0 Stars	✓	✅ Deniable	✅ Deniable	🔴 Detectable
1 Star		✓	✅ Deniable	🔴 Detectable
2 Stars			✓	🔴 Detectable
3 Stars				✓

Investor Behavior

- Investor decisions reflect advisor honesty across treatments.
- Investment rates given the same announced stars
  - Public-Oath > Private-Oath > No-Oath
  - Necessary condition ≈ Public-Oath
  - Sufficient condition ≈ Private-Oath

Concern

- Necessary and Sufficient treatments involve pre-selecting advisors based on behavior observed in earlier treatments.
- This introduces a **gray area** in experimental transparency:
- **Necessary condition**:
  - Investors are not fully informed about the oath or the matching procedure.
  - Typically seen as less problematic, but still not fully transparent.
- **Sufficient condition**:
  - Advisor behavior is real, but investor matching is engineered to fit the treatment logic.
  - May be viewed by some experimentalists as a form of deception.
- **Implication**: While this enables clear identification of mechanisms, it may conflict with norms against deception in experimental economics.

Suggestions & Open Questions

- Investigate individual heterogeneity (gender?)
- Would a voluntary oath affect behavior differently than a mandatory one?
- In real-world settings, commissions create strong monetary incentives.
  - If oaths are mandated, agents who are oath-sensitive may exit or underperform while agents who are oath-insensitive may lie more and succeed.
  - This creates a selection problem: oaths may not change average behavior, just who stays in the industry.