


# Honesty oaths for rule-following

Shaul Shalvi

 Check for updates

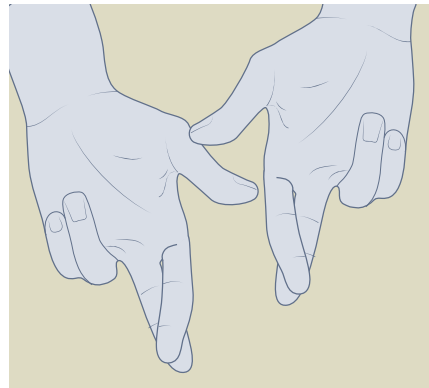
Honesty oaths are commonly used to promote ethical behaviour, but their effectiveness is not well understood. A mega-study involving thousands of people shows that taking an oath to be honest can reduce tax evasion in an online economic game.

Do honesty oaths lead people to obey rules? Organizations use honesty oaths to promote ethical behaviour among their members, but their effectiveness is unclear. For example, students often take an honesty oath before embarking on their studies but then use ChatGPT to write their assignments. Janis Zickfeld and 41 other authors investigate this question through a mega-study experiment<sup>1</sup>. The study reveals that being asked to take an honesty oath increases the likelihood of people following rules in an online economic game, as compared to those not asked to take an oath. However, the formulation of the oath has a crucial role: ten types of oaths statistically increased rule-following behaviour, but eleven others did not. These findings challenge the notion that taking an oath is merely symbolic and suggest that encouraging people to commit to honesty may be beneficial.

In January 2025, a new US president will be inaugurated. Similar to previous US presidents and other federal employees since 1884, this new president will take the Oath of Office<sup>2</sup> and will commit to “support and defend the Constitution of the United States against all enemies, foreign and domestic.” Medical doctors similarly take a version of the Hippocratic oath<sup>3</sup>, written about 2,500 years ago, and commit to adhering to the ethical guidelines of their profession. Business school students, traders and others also take oaths and commit to behaving ethically.

A plausible reason to follow one’s oath is fearing the penalties that are expected from violating it. Whereas often not seen as formal contracts, public officials who violate their oaths may be impeached, removed from office or disqualified from holding future office. Medical doctors may lose their licences and students can be expelled from their studies. Fear of punishment may motivate following one’s oath. However, until now, the extent to which oaths lead to rule-following behaviour without such fear has remained less well understood. Zickfeld and colleagues<sup>1</sup> propose that when no punishment exists for oath violations, the reason for following it is one’s desire to maintain a positive image. Testing this idea, they generated 21 different types of oaths and experimented how they influence people’s behaviour in an economic game in which violating the oath carried no sanction.

The authors assigned 21,506 participants from the UK and the USA to different treatments. Participants in the experimental treatments were asked to take an oath before engaging in an economic tax compliance game. Participants in the control condition engaged in the game without taking an oath first. In this game, participants generated income during four rounds of a sorting task in which they spent as little time possible on sorting eight numbers as either even or odd.



Participants earned money on the basis of their task performance. They were instructed to report these earnings as a 35% tax was to be deducted from the reported earnings, which determined participants’ final payment. The game mimics real life in that people earn an income and must report it to the tax authority. Assuming they are honest, the more people earn the higher the tax they need to pay. If people are dishonest and choose to evade taxes, they can declare earning less than they did and accordingly pay lower taxes than they were obliged to.

One of the interesting aspects of the study was the way that the oaths considered were generated. The authors first crowdsourced 98 suggestions by brainstorming together. They then screened and voted on the most promising oaths, which led them to settle on the most promising 20 oaths (those they believed were most likely to have a positive effect). They further implemented two benchmarks, a control treatment in which participant took no oath and a baseline oath treatment in which participants stated “I hereby declare that I will provide honest information in this study”. For every participant, the authors further varied whether the oath was presented before or after the sorting task, as well as whether committing to the oath was done by checking a box next to the oath or by retyping it.

Zickfeld and colleagues<sup>1</sup> found that a quarter of their participants (25.1%) underreported their earnings to some degree, and that more than a quarter of tax evaders (28.1%) chose to evade paying taxes altogether by reporting to have earned nothing. Across all interventions, being asked to take an oath led to a 3.9-percentage-points increase in tax compliance, as compared with the no-oath treatment. Ten interventions revealed a positive effect in statistically reducing evasion, even after correcting for multiple statistical comparisons. The most successful intervention alleviated the losses suffered from evasion by 47%. These effects were robust to the way that people committed to the oath (clicking a box versus retyping it) and the timing of the oath (before generating income versus before reporting the income).

Zickfeld and colleagues studied the effect of oaths on tax evasion in a game in which violating one’s oath has no negative consequences. Doing so is useful in terms of being able to attribute any positive effect of the intervention to people’s desire to maintain their

positive image<sup>4–9</sup>. However, outside of the laboratory, violating oaths can lead to punishment. Accordingly, the experimental setting used here limits our ability to generalize the findings to real-life settings. A potential path towards implementing any of the successful interventions identified into policy may require several key steps. First, it will require direct replication: testing whether successful interventions replicate in independent studies, ideally across various cultures. Second, it will need to be assessed for robustness: whether successful interventions also maintain their positive effect in settings in which violating one's oath may lead to punishment. Third, field evidence is needed: assessing whether successful interventions also work outside of the laboratory in field settings. Fourth and last, aggregation should be considered: assessing meta-analytically how solid the successful interventions are across settings.

Policymakers who seek to reduce tax evasion will no doubt find much of interest in the current work, granting them tools for shaping policy. However, in doing so, they should remember that good science is slow science. By carefully establishing the replicability and robustness of the current findings in real-life settings, we can design responsible evidence-based policy.

**Shaul Shalvi**  

Amsterdam School of Economics, University of Amsterdam, Amsterdam, The Netherlands.

✉ e-mail: [s.shalvi@uva.nl](mailto:s.shalvi@uva.nl)

Published online: 21 October 2024

### References

1. Zickfeld, J. H. et al. *Nat. Hum. Behav.* <https://doi.org/10.1038/s41562-024-02009-0> (2024).
2. United States Senate. About the Senate & the U.S. Constitution | Oath of Office. [senate.gov, https://www.senate.gov/about/origins-foundations/senate-and-constitution/oath-of-office.htm](https://www.senate.gov/about/origins-foundations/senate-and-constitution/oath-of-office.htm) (accessed 13 September 2024).
3. Miles, S. H. *The Hippocratic Oath and the Ethics of Medicine* (Oxford Univ. Press, 2005).
4. Mazar, N., Amir, O. & Ariely, D. *J. Mark. Res.* **45**, 633–644 (2008).
5. Shalvi, S., Dana, J., Handgraaf, M. J. J. & De Dreu, C. K. W. *Organ. Behav. Hum. Decis. Process.* **115**, 181–190 (2011).
6. Weisel, O. & Shalvi, S. *Proc. Natl Acad. Sci. USA* **112**, 10651–10656 (2015).
7. Leib, M., Köbis, N., Soraperra, I., Weisel, O. & Shalvi, S. *Psychol. Bull.* **147**, 1241–1268 (2021).
8. Gerlach, P., Teodorescu, K. & Hertwig, R. *Psychol. Bull.* **145**, 1–44 (2019).
9. Abeler, J., Nosenzo, D. & Raymond, C. *Econometrica* **87**, 1115–1153 (2019).

### Competing interests

The author declares no competing interests.