

Grimm, Veronika; Choo, Lawrence; Horvath, Gergely; Nitta, Kohei

Conference Paper

Whistleblowing and Diffusion of Responsibility: An Experimental Investigation

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Social Norms and Diffusion of Responsibility, No. G21-V2

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Grimm, Veronika; Choo, Lawrence; Horvath, Gergely; Nitta, Kohei (2016) : Whistleblowing and Diffusion of Responsibility: An Experimental Investigation, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Social Norms and Diffusion of Responsibility, No. G21-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/145781>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Whistleblowing and Diffusion of Responsibility: An Experimental Investigation

Abstract

Societies today are increasingly reliant on whistleblowing to uncover unfair practices. For policy design it is therefore essential to better understand the impact of socio-psychological factors on whistleblowing propensity. We use an experimental setup to explore the role of “Diffusion of Responsibility” (DOR), which posits that individuals are less likely to whistleblow when others could similarly do so. We find that individuals do not shift responsibility to others if they expect an own monetary benefit from the consequences of whistleblowing, even if they could gain even more by freeriding on others who take action. In contrast, DOR affects whistleblowing propensities severely if whistleblowing is motivated solely by altruistic concerns. Our results highlight the fragility of purely altruistic behaviour and suggest that whistleblowing policies should ensure that potential whistleblowers perceive a gain to their in-group, but need not address the free riding problem among the insiders.

Introduction. Whistleblowing is often characterised by *Insiders'* disclosure of private information pertaining to unfair (e.g., immoral, illegal, illegitimate) practices that affect *Outsiders'* welfare. To encourage Insiders to whistleblow, policy makers often focus on mechanisms that afford protections (e.g., anonymity, anti-retaliation) and sometimes even rewards to whistleblowers. However, anecdotal observations from known whistleblowing cases suggest that unfair practices often persist for extended periods, evidential to numerous Insiders who remained silent. Are whistleblowers unique individuals or could the reluctance of even *altruistic* Insiders to whistleblow be explained by some broader social phenomenon?

The "Diffusion of responsibility" (DOR) or bystander effect (1) is one such explanation that has sometimes been advanced. It posits that each Insider becomes less willing to whistleblow, when aware of others who could similarly do so. We should, however, highlight that though DOR is well-documented in emergency situations [e.g., (2)], when there is little ambiguity about the victim's plight, applications to non-emergency cases such as whistleblowing is less obvious, especially given that Outsiders are often unaware of their own plight. Indeed, whilst some whistleblowing studies document tentative evidence in line with the DOR predictions [e.g., (3)], most others find no or contradictory evidence [e.g., (4,5)]. This has led some [e.g., (6)] to suggest that DOR might not be applicable in whistleblowing cases as Insiders often have ample time to strategically consider their decisions.

Nevertheless, most whistleblowing studies utilise hypothetical surveys where unfair practices are framed and emphasised. In addition to potential concerns with the experimenter demand effects, the nature of surveys implies that whistleblowing beneficiaries (e.g., who gains, who loses) are often ambiguous. This study revisits the DOR debate in an experimental game, where Insiders face *real* monetary consequences to their whistleblowing decisions (7), to study the impact of DOR on whistleblowing propensities. To do so, we introduce two games, the OneWB and TwoWB, both of which are simple extensions to the asymmetric information Ultimatum Game [e.g., (8)]. Both games are similar in design and structure but only differ on the number of Insiders who could whistleblow. Furthermore, permutations within each game allows us to identify cases where whistleblowing is motivated by both strategic self-interests and altruistic concerns, and cases where only altruistic concerns could motivate whistleblowers. This demarcation of whistleblowing motivations is useful given that it helps us better understand the impact of DOR (if any) in the broader social context. Whistleblowing on corporate tax evasion when taxes fund public projects

that are enjoyed by everyone, is a familiar example for the former cases. Doctors whistleblowing when hospital funds are used to enhance their pay packages instead of patient care, is an example for the latter cases.

Research Design. The TwoWB game involves two teams (Team-A and Team-B) of three players, each endowed with €1, and a common pot that can contain €30, €60, €90, €120 or €150, with equal probabilities. The pot's realised amount is only known to Team-A players. One Team-A player (the "Proposer") first proposes a distribution of the pot's amount between both teams. The three Team-B players (the "Outsiders") see only the amount offered to their team, whilst the two other Team-A players (the "Insiders") see both teams' amounts. Thereafter, each Insider must independently decide on whether to send an anonymous "Signal" at the cost of €0.50, informing Outsiders about the pot's realised amount. If no signal is sent, the two Insiders and the three Outsiders will each proceed to vote on their acceptance of the proposal. If a signal is sent by one or both Insiders, the Proposer gets the possibility to revise her distribution. Insiders and Outsiders are thereafter informed of the revised offer and vote on acceptance. After all the votes are casted, a majority rule is used to determine whether the Proposer's distribution is implemented. If not, the pot is not distributed. In addition to their endowment and associated signalling cost (if any), players' monetary payoffs are either 1/3 of their teams' share if implementation occurs or zero otherwise.

The OneWB game is similar to TwoWB with the exception that whilst both Insiders independently decide on their signalling decisions, the decision by only one random Insider will be used in determining whether a signal is sent. This random process ensures the anonymity of the signal sender and makes each Insider to take her signal decision as if she was the only one who could send a signal.

Whistleblowing and DOR. Whistleblowing in both games refers to the Insiders' decision to send a signal. This is however a dominated action given that the equilibrium predictions for each game and pot are for Proposers to give the minimum possible amount to Team-B, Insiders to never whistleblow and all players to always accept the Proposer's distribution. However, the equilibrium predictions are unlikely to hold as evidence from the standard Ultimatum game suggests that people often only accept the Proposer's offer if they believe themselves to be getting a "fair-deal", around 30 to 40% of the known pot (9).

To therefore see why Insiders might whistleblow, we will assume that Outsiders anchor their fair-deal expectations upon the €90 pot, the expected value of the unknown pot. Information asymmetry for "Small" (€30 or €60) pots is costly to both teams as Outsiders' expectations could easily exceed any offer to Team-B, however equitable. Whistleblowing here benefits all players by realigning Outsiders' expectations and supporting implementation. In contrast, whistleblowing for "Large" (€120 or €150) pots is only beneficial to Team-B, as it limits the Proposer's ability to exploit the information asymmetry, making unfair offers that are seemingly generous given the expected pot size. Furthermore, doing so is surely detrimental to Insiders as fairer distributions will inevitably reduce Team-A's payoff (10). The dichotomy of whistleblowing beneficiaries suggests whistleblowing willingness for Small pots could be motivated by both strategic self-interest and other-regarding altruistic concerns. In contrast, whistleblowing willingness for Large pots is inconsistent with strategic self-interest and can only be motivated by Insiders' altruistic concerns for the welfare of Outsiders.

Orthogonal to the above discussions, DOR predicts that for a fixed pot and an unfair offer to Team-B, Insiders in OneWB are more willing to whistleblow than those in TwoWB. This is plausible for Small pots, where selfish Insiders could strategically reduce their whistleblowing

willingness in an attempt to free-ride on the whistleblowing activity of their peer. Indeed, DOR in the Small pots can be understood in the same framework as the “volunteer’s dilemma” (11). For Large pots, where whistleblowing is motivated by altruistic concerns, the occurrence of DOR is less clear. Economic logic suggests that the loss due to reconsideration of the pot’s distribution by the Proposer should far exceed the whistleblowing cost. Thus, free riding is an implausible motive for lowering one’s willingness to whistleblow. This is an important feature of our design given that it allows to examine whether DOR might even permeate into pro-social altruistic behaviours.

Experimental Procedures. The experiment involved two treatments, OneWB and TwoWB, both of which are direct adaptations of the respective games described in the previous paragraphs. Each treatment involved 60 subjects, recruited from the student population at the University of Erlangen-Nuremberg. The treatments were one-shot, conducted through the z-Tree software (12) and employed neutral framing of both games. Participants received a €4 show-up payment in addition to their experimental earnings.

For efficient data collection and comparisons, we utilised the "Strategy Method" [e.g., (13)] in the experimental implementation, where subjects' contingent decisions were elicited in advance, (14). Specifically, subjects were first randomly assigned into groups of 6. Thereafter, Proposers submitted their "Distribution-Plans" for all pots. Outsiders and Insiders submitted their "Acceptance-Threshold", the minimum amount for their team that they would accept. However, Outsiders did so for the unknown pot whilst Insiders for each possible pot. Finally, Insiders also submitted their "Signal-Threshold" for each possible pot size, i.e. the minimum amount that the Proposer had to offer to team B in order to prevent the respective player from sending a signal (15). Thereafter, the group’s pot amount was randomly determined, Team-A subjects were informed of this amount and the corresponding contingent decisions were used to determine whether a signal was sent. In the absence of a signal, subjects' contingent decisions were used to determine whether implementation occurred. If a signal was sent, subjects had the possibility to revise their decisions (Distribution-Plans and Acceptance-Thresholds) for the known pot, which were thereafter used to determine implementation.

The above procedures resulted in 10 Proposers, 20 Insiders and 30 Outsiders per treatment. Here, subjects’ Acceptance-Thresholds can be interpreted as their *fair-deal expectations* in the game. Insiders’ Signal-Thresholds can be interpreted as their *whistleblowing willingness* for each pot.

Results. We find no significant between-treatment differences in Proposers' Distribution-Plans, Insiders’ Acceptance-Thresholds or Outsiders Acceptance-Thresholds (Mann-Whitney $p > .10$ for each comparison). This allows us to focus our discussions on whistleblowing decisions of Insiders. To do so, we define Insiders’ "relative whistleblowing willingness" (RWW) for a pot, as their Signal-Threshold for that pot normalised by the pot's size (16). Building on this, we define an Insider to be a “Potential Whistleblower” (PW) if $RWW > 0$. To this end, 65% and 60% of all Insiders in OneWB and TwoWB, respectively, were PW for all pots ($RWW > 0$ for all pots).

Insiders’ mean and median RWW are presented in Fig. 1, where we also report the Ratio of PW subjects. Given the above, OneWB should provide a benchmark description of whistleblowing willingness when Insiders know themselves to be the only ones who could whistleblow. On this matter, personal benefits of whistleblowing do not seem to have much influence on whistleblowing willingness. PW ratios ranged from 0.60 in the Small pots to 0.55 in the Large pots. Furthermore, RWW in the Large and Small pots do not seem very different. To see this more clearly, we computed for each Insider her average RWW over the Small (€30 and €60) and Large (€120 and €150) pots. Here, average RWW in the Small and Large pots are not significantly different (Wilcoxon Signrank $p = .135$). This implies that the lack of individual whistleblowing benefits or

the absences of strategic self-interest motivations in Large pots of OneWB did not lead to a significant reduction in whistleblowing willingness.

A treatment comparison reveals no significant between treatment differences in RWW for the €30, €60 and €90 pots (Mann-Whitney $p \geq 0.17$; ttest $p \geq 0.13$; in each comparison). However, RWW for the €120 and €150 pots are significantly higher in OneWB relative to TwoWB (Mann-Whitney $p < 0.01$; ttest $p \leq 0.01$; in each comparison). The evidence here suggests that DOR only influences whistleblowing propensities in instances where whistleblowing is motivated by altruistic concerns only (i.e. for Large pots).

To better understand the above findings, we turn our attention to the ratio of PW subjects in OneWB and TwoWB. The ratio of PW subjects decreases from 0.60 to 0.55 in OneWB, but from 0.55 to 0.20 in TwoWB. The ratios suggest that the decrease of the whistleblowing propensity for Large pots in TwoWB as compared to OneWB is driven by the extensive margin, i.e. the fact that a higher share of Insiders chooses to abstain from whistleblowing under any circumstances when aware of others who could similarly take action. Indeed, amongst the TwoWB who were PW subjects for Small pots, only 36% were also found to be PW subjects for Large pots. The corresponding proportion in OneWB was 83%. Fig. 2 provides further insights in this direction by showing mean and median RWW of PW subjects only. In fact, PW subjects show no difference in RWW across treatments for the €90, €120 and €150 pots (Mann-Whitney $p > 0.17$; ttest $p > .33$). This perhaps suggests that a minority of subjects is “immune” to the impact of DOR.

Finally, note that RWW among the PW subjects significantly increases for Small pots. One possible explanation might be that PW subjects try to compensate for DOR in the TwoWB setup. Note that an increase in RWW might be a plausible reaction to a lower PW ratio, since other than in OneWB Insiders have the possibility to compensate for an abstention of their peer.

Discussion. Why would the treatment differences only be significant for Large pots? Our prior on this matter is as follows. A substantial share of Insiders in TwoWB realised the potential relationship between whistleblowing and implementation for Small pots but were uncertain as to whether the other Insider arrived at the same conclusion. Faced with the risk of non-implementation, each Insider therefore “took matters into their own hands”. For this reason, RWW in the Small pots of OneWB and TwoWB are not significantly different. By the same logic, most Insiders in both treatments realised that whistleblowing for Large pots is not necessary for implementation but instead serve to improve the welfare of Outsiders. The evidence from OneWB suggests, however, that a substantial share of Insiders feel responsibly. However, our results suggest clearly that most Insiders chose to delegate away the responsibility for mitigating the plight of Outsiders when aware of others who could whistleblow. As such, RWW in Large pots are significantly higher in OneWB. This suggests that DOR is less related to the phenomenon of free riding, but rather related to changes in the attribution of responsibility.

In summary, we provide differentiated evidence on the effect of DOR on whistleblowing propensity. In particular, we show that purely altruistic motivations are wiped out by Insiders’ awareness of other Insiders who could similarly take action, while we do not find evidence for DOR (in the sense of freeriding) if Insiders expect in-group benefits from whistleblowing. This highlights the fragility of altruistic behaviours in the whistleblowing context and provides some support for the conjecture that DOR is a particularly important phenomenon if solely altruistic motives could motivate whistleblowing.

References and Notes:

1. D. Latane, J.M. Darley, Bystander intervention in emergencies: Diffusion of responsibility, *Journal of Personality and Social Psychology*, **8**, 377–383, (1968).
2. P. Fischer et al, The Bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies, *Psychological Bulletin*, **137**, 517-537, (2011).
3. J. Gao, R. Greenberg, B.W.O. Wing, Whistleblowing intentions of lower-level employees: The effect of reporting channel, bystander, and wrongdoer power status, *Journal of Business Ethics*, **126**, 85-99, (2015).
4. S.N. Robinson, J.C. Robertson, M.B. Curtis, The effects of contextual and wrongdoing attributes on organizational employees' whistleblowing intentions following fraud, *Journal of Business Ethics*, **106**, 213-227, (2014).
5. M.P. Miceli, J.M. Near, Individual and situational correlates of whistle-blowing, *Personnel Psychology*, **41**, 267-281, (1988).
6. M.P. Miceli, J.P. Near, T.M. Dworkin, *Whistle-blowing in Organization*, (Taylor & Francis, NY, 2008).
7. Though we do not believe that pecuniary factors are the primary drivers of whistleblowing behaviours, they cannot be ignored.
8. M. Mitzkewitz, R.Nagel, Experimental results on ultimatum games with incomplete information, *International Journal of Game Theory*, **22**, 171-98, (1993).
9. C. Camerer, *Behavioral game theory: Experiments in strategic interaction*, (Princeton University Press, 2003).
10. We remain nonchalant as to the €90 pot which serve as our anchor for discussions.
11. A. Diekmann, Volunteer's Dilemma, *Journal of Conflict Resolution*, **29**, 605-610, (1985).
12. U. Fischbacher, z-Tree: Zurich toolbox for ready-made economic experiments, *Experimental economics*, **10**, 171-178, (2007).
13. J. Brandts, G. Charness, The strategy versus the direct-response method: a first survey of experimental comparisons, *Experimental Economics*, **14**, 375-398, (2011).
14. Instruction, screenshots and further calculations are available in the online supplement.
15. For example, a Signal-Threshold of 30 for the €90 pot implies that if the realized pot contains €90, the Insider only wants to whistleblow if the Proposer gives strictly less than €30 to team B.
16. For example, a RWW=0.50 for specific pot implies that the Insider only wishes to whistleblow if Proposers offers less than 50% of that pot to Team-B.

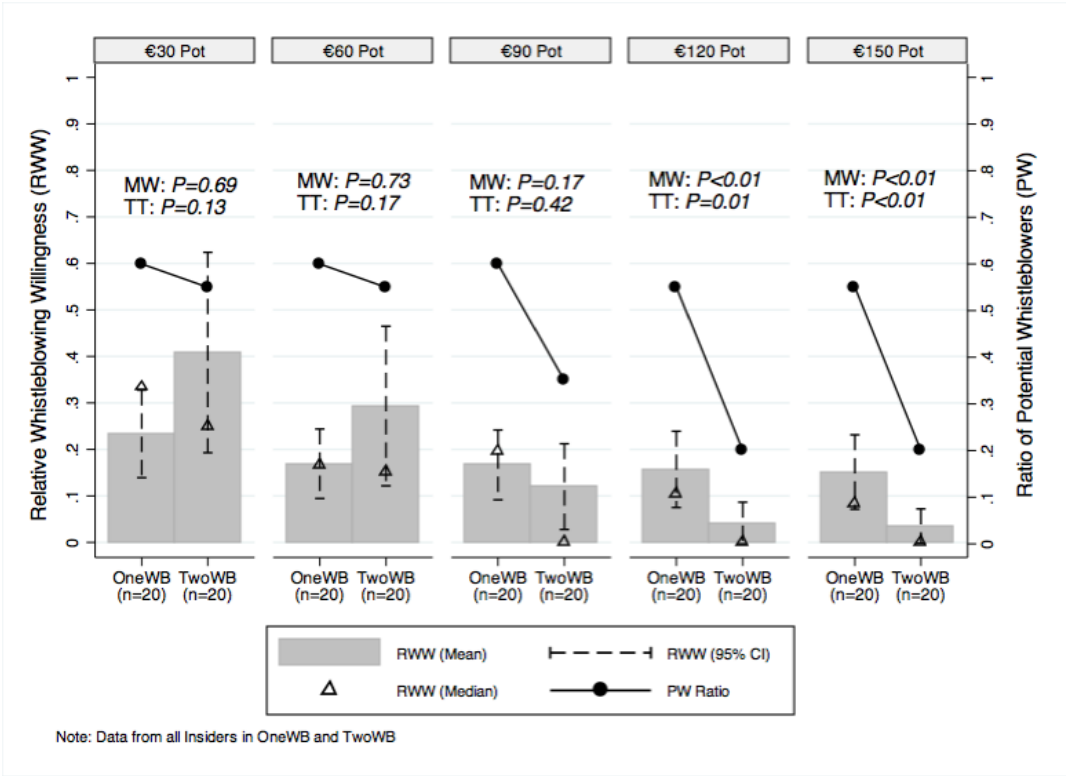


Fig. 1 RWW and PW Ratio by Pots and Treatments (All Insiders). We report in each panel the t-test (TT) and Mann-Whitney Test (MW) p-values for between treatment comparisons of RWW.

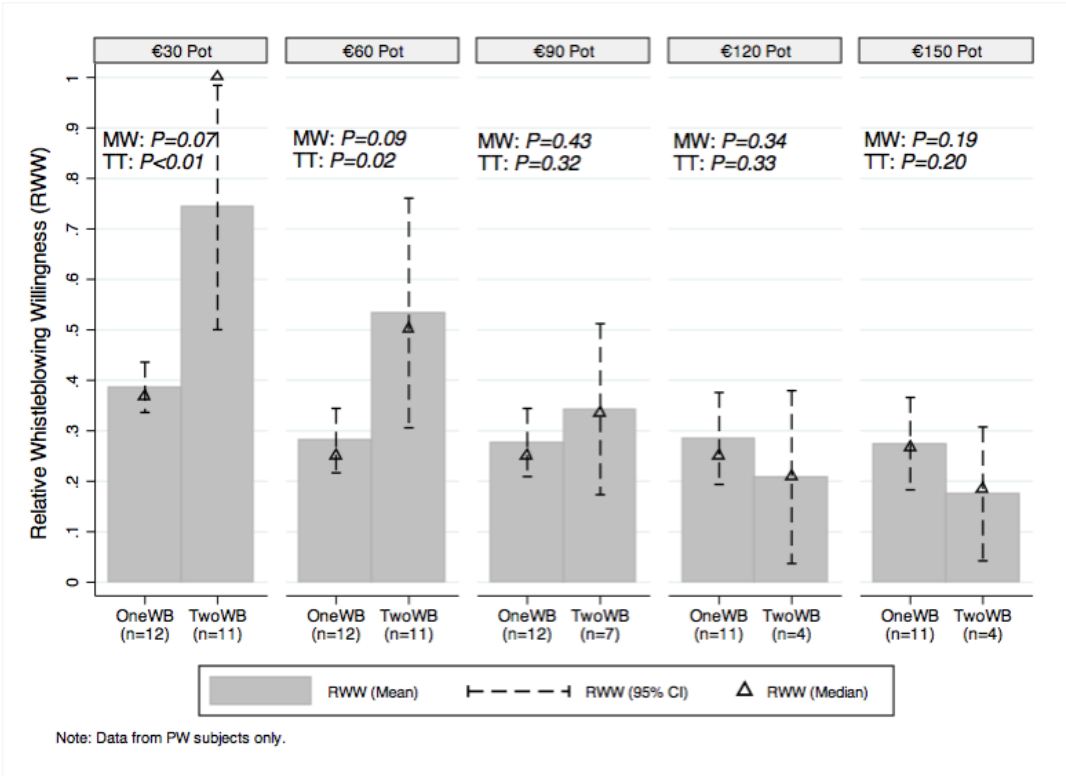


Fig. 2 RWW by Pots and Treatments (PW Subjects only). We report in each panel the t-test (TT) and Mann-Whitney Test (MW) p-values for between treatment comparisons of RWW.