51. Engineering cooperation: the behavioral design of reputation systems

GARY BOLTON, BEN GREINER AND AXEL OCKENFELS

Feedback provision

Why do traders give feedback? Feedback is largely a public good, aiding other traders in managing the risk associated with an unfamiliar exchange partner. Models that assume a purely material incentive for action imply low participation rates, yet the rates of feedback giving are often high. The human inclination for reciprocity as described by social preference models (see Cooper and Kagel 2016 for an overview) provides an explanation: Kind feedback is given in response to kind treatment in the transaction, negative feedback in response to negative treatment (Bolton et al. 2004). The reciprocal nature of feedback giving is also evident in feedback provision in two-way feedback systems where sellers and buyers rate each other. A study of eBay's two-way system found that about 70 per cent of eBay traders leave feedback (Bolton et al. 2013; other research finds similar numbers). Mutual feedback is given much more frequently (64 per cent of the time) than would be anticipated if feedback provision were independent (49 per cent of the time). Traders are more likely to provide feedback on a trader with whom they share a common group identity (Bolton, Mans, and Ockenfels 2020), consistent with findings that the propensity for norm enforcement is stronger in within-group interactions (Bernhard et al. 2006; Chen and Li 2009; Chen et al. 2010; Goette et al. 2012; Mussweiler and Ockenfels 2013).

Distortions in reported feedback

This is where the feedback given tends to extremes. Typically, the vast majority is very positive, and most of the remainder is very negative. Analysing eBay data, Nosko and Tadelis (2015) find that traders' average percentage of positive feedback is 99.3 per cent, with a median of 100 per cent. There is evidence that traders who have had a mildly negative experience are less likely to report it (Dellarocas and Wood 2008), possibly due to a preference for leniency when the attribution of errors is ambiguous (Bolton et al. 2019). Similar leniency is observed in employee evaluations and might result from altruistic tendencies (Ockenfels et al. 2015; Kusterer and Sliwka 2022). Deeper dissatisfaction may be a major motivation for submitting a negative review (Lafky 2014).

The potential problem with excessively positive feedback is that it provides an overly confident view of a trader's reliability. Many feedback patterns we see can be linked to two fundamental research findings on patterns of human cooperation: Altruistic punishment promotes cooperation, while counter-punishment hampers it (Ostrom et al. 1992; Fehr and Gächter 2000; 2002; Nikiforakis 2008; Mussweiler and Ockenfels 2013; Balafoutas et al. 2014). A natural way to (altruistically) punish a trader on an Internet platform for not behaving according to the social or trading norm is to leave negative feedback. This way, altruistic punishment of norm-violators creates an incentive for trustworthy behaviour. However, punishments can often be counter-punished, which is known to reduce the effectiveness of punishment to promote cooperation. Indeed, by retaliating against negative feedback with negative feedback, counter-punishment may spoil the reputation of the altruistic punisher, which in turn may deter altruistic punishment in the first place. As a result, the potential of counter-punishment can hamper the effectiveness of reputation mechanisms and thus the performance of markets (Bolton et al. 2013; 2018; 2023).

Future research

The biases observed in feedback giving, along with their negative impact on market trading, highlight the need for improved, behaviourally-informed feedback system designs (Bolton and Ockenfels 2012; Chen et al. 2021). Many Internet reputation systems ask traders to provide feedback with general or open-ended questions, or on scales like the commonly used Likert 5-point scale. Recent work shows that the informativeness of some systems can be improved by choosing questions and scales that are more aligned with the trust issues specific to a marketplace (Bolton et al. forthcoming).

There is also work investigating the use of incentives to motivate raters to provide (more) useful information (see Li 2010; Li and Xiao

2014; and Li et al. 2016 for case studies on Alibaba; Cabral and Li 2015 for field experiments on eBay; and Burtch et al. 2018). Other work advances using big data and artificial intelligence to predict future trader behaviour (Milgrom and Tadelis 2018; Masterov et al. 2015; Luca and Zervas 2016).

Acknowledgement

References

Balafoutas L, Nikiforakis N, Rockenbach B (2014) Direct and indirect punishment among strangers in the field. *Proceedings of the National Academy of Sciences* 111(45):15924–15927.

Bernhard H, Fehr E, Fischbacher U (2006) Group Affiliation and Altruistic Norm Enforcement. *American Economic Review* 96(2):217–221.

Bolton GE, Breuer K, Greiner B, Ockenfels A (2023) Fixing feedback revision rules in online markets. *Journal of Economics and Management Strategy* 32(2):247–256.

Bolton GE, Ferecatu A, Kusterer DJ (forthcoming). Rate This Transaction: Coordinating Mappings in Market Feedback Systems. *Management Science*.

Bolton GE, Greiner B, Ockenfels A (2013) Engineering trust: Reciprocity in the production of reputation information. *Management Science* 59(2):265–285.

Bolton GE, Greiner B, Ockenfels A (2018) Dispute resolution or escalation? The strategic gaming of feedback withdrawal options in online markets. *Management Science* 64(9):4009–4031.

Bolton GE, Katok E, Ockenfels A (2004) How effective are electronic reputation mechanisms? An experimental investigation. *Management Science* 50(11):1587–1602.

Bolton GE, Kusterer DJ, Mans J (2019) Inflated reputations: Uncertainty, leniency, and moral wiggle room in trader feedback systems. *Management Science* 65(11):5371–5391.

Bolton GE, Mans J, Ockenfels A (2020) A norm enforcement in markets: Group identity and the volunteering of feedback. *Economic Journal* 130(629):1248–1261.

Bolton GE, Ockenfels A (2012) Behavioral economic engineering. *Journal of Economic Psychology* 33(3):665–676.

Burtch G, Hong Y, Bapna R, Griskevicius V (2018) Stimulating online reviews by combining financial incentives and social norms. *Management Science* 64(5):2065–2082.

Cabral L, Li L (Ivy) (2015) A dollar for your thoughts: Feedback-conditional rebates on eBay. *Management Science* 61(9):2052–2063.

Chen Y, Cramton P, List J, Ockenfels A (2021) Market Design, Human Behavior and Management. *Management Science* 67(9):5317–5348.

Chen Y, Harper FM, Konstan J, Li SX (2010) Social comparisons and contributions to online communities: A field experiment on MovieLens. *American Economic Review* 100(4):1358–1398.

Chen Y, Li SX (2009) Group identity and social preferences. *American Economic Review* 99(1):431–457.

Cooper DJ, Kagel JH (2016) Other-regarding preferences: A selective survey of experimental results. Kagel JH, Roth AE, eds. *Handbook of Experimental Economics, Vol. 2*. (Princeton University Press, Princeton), 217–289.

Dellarocas C, Wood CA (2008) The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management Science* 54(3):460–476.

Fehr E, Gächter S (2000) Cooperation and punishment in public goods experiments. *American Economic Review* 90(4):980–994.

Fehr E, Gächter S (2002) Altruistic punishment in humans. *Nature* 415(6868):137–140.

Goette L, Huffman D, Meier S (2012) The Impact of Social Ties on Group Interactions: Evidence from Minimal Groups and Randomly Assigned Real Groups. *American Economic Journal: Microeconomics* 4(1):101–15.

Kreps DM, Wilson R (1982) Reputation and imperfect information. *Journal of Economic Theory* 27(2):253–279.

Kusterer D, Sliwka D (2022) Social Preferences and Rating Biases in Subjective Performance Evaluations. IZA Discussion Paper No. 15496, http://dx.doi.org/10.2139 /ssrn.4196269

GARY BOLTON, BEN GREINER AND AXEL OCKENFELS

Lafky J. (2014) Why do people rate? Theory and evidence on online ratings. *Games and Economic Behavior* 87(C), 554–570.

Li L. (2010) Reputation, trust, and rebates: How online auction markets can improve their feedback mechanisms. *Journal of Economics & Management Strategy*, *19*(2): 303–331.

Li LI, Tadelis S, Zhou X (2016) Buying reputation as a signal of quality: Evidence from an online marketplace. NBER Working Paper.

Li LI, Xiao E (2014) Money talks: Rebate mechanisms in reputation system design. *Management Science* 60(8):2054–2072.

Luca M, Zervas G (2016) Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science* 62(12):3412–3427.

Masterov DV, Mayer UF, Tadelis S (2015) Canary in the e-commerce coal mine: Detecting and predicting poor experiences using buyer-to-seller messages. *Proceedings of the Sixteenth ACM Conference on Economics and Computation*. EC '15. (ACM, New York, NY), 81–93.

Milgrom PR, Tadelis S (2018) How artificial intelligence and machine learning can impact market design. Agrawal A, Gans J, Goldfarb A, eds. *The Economics of Artificial Intelligence: An Agenda*. National Bureau of Economic Research conference report. (University of Chicago Press, Chicago), 567–585.

Mussweiler, T., & Ockenfels, A. (2013) Similarity increases altruistic punishment in humans. *Proceedings of the National Academy of Sciences*, *110*(48), 19318–19323.

Nikiforakis N (2008) Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics* 92(1–2):91–112.

Nosko C, Tadelis S (2015) The limits of reputation in platform markets: An empirical analysis and field experiment. NBER Working Paper.

Ockenfels A, Sliwka D, Werner P (2015) Bonus payments and reference point violations. *Management Science* 61(7):1496–1513.

Ostrom E, Walker J, Gardner R (1992) Covenants with and without a sword: Self-governance is possible. *American Political Science Review* 86(2):404–417.

GARY BOLTON, BEN GREINER AND AXEL OCKENFELS