



UvA-DARE (Digital Academic Repository)

Mechanisms for the evolution of prosociality

Graser, C.J.

Publication date
2024

[Link to publication](#)

Citation for published version (APA):

Graser, C. J. (2024). *Mechanisms for the evolution of prosociality*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Appendices

English summary

In broad terms, this thesis looks at several mechanisms for the evolution of prosocial behaviour, using the tools of evolutionary game theory. The term *prosocial behaviour* refers to behaviour which benefits others, but which may come at a cost to the oneself. In evolutionary game theory, interactions between individuals are described mathematically. Depending on the behaviour of all interacting individuals, individuals get assigned *payoffs*, which, in turn, determine their evolutionary success – those individuals with higher payoffs become more abundant in the population, those with lower payoffs become less abundant. In such models, how prosocial behaviour can evolve, therefore, constitutes a bit of a puzzle, as it is the recipients of the prosocial behaviour, and not necessarily those who behave prosocially, who enjoy the selective benefit.

There are different types of evolutionary mechanisms that provide a solution to this puzzle. For instance, if there is *assortment*, meaning that individuals preferentially interact with others who are genetically similar, then behaviour that deals out selective benefits for others, can, nevertheless, contribute to the spread of the genes that code for this behaviour.

The first part of this thesis (Chapter 1) relates to this mechanism. Specifically, Chapter 1 considers a model of preference evolution. Rather than modelling natural selection for behaviours in specific types of interactions (games), models of preference evolution try to find rules – preferences – which explain behaviour across larger classes of games, and look at why or whether individuals who behave according to such preferences are favoured by natural selection.

Chapter 1 focuses on a particular rule, known as "Homo Moralis" preferences (Alger and Weibull, 2013). In any given game, individuals who behave according to this rule choose a strategy which maximises a combination of (1) their own payoff and (2) the joint payoff that they and their partner would get if both chose this strategy. Alger and Weibull (2013) suggest that Homo Moralis preferences, if the relative weight put on (2) is consistent with the degree of assortment in the population, are especially successful in evolution.

Chapter 1 shows that a different rule, "regular altruism", does at least equally well. Rather than requiring this second, hypothetical term of "what behaviour would lead to the ideal outcome if everyone behaved according to it", regular altruists maximise a combination of, again,

(1) their own payoff, but now (2) their interaction partner's payoff. In terms of the evolution of prosociality, the difference between Homo Moralis and regular altruists therefore lies in whether the feature that brings about the prosocial behaviour is a Kantian categorical-imperative-like tendency towards socially efficient behaviour, or whether it is a direct concern for the payoff of the individual one currently interacts with.

Chapter 1 shows that in all games in which Homo Moralis preferences prescribe a well-defined strategy, this strategy is also consistent with regular altruist preferences. However, there are settings in which there is no strategy that maximises Homo Moralis preferences so that Homo Moralis' behaviour is undefined. Here, Chapter 1 shows that natural ways of extending the definition of Homo Moralis preferences to cover these settings either, again, lead to identical behaviour for Homo Moralis and regular altruists, or, if they lead to distinct behaviours, regular altruists win out in direct competition against Homo Moralis.

In Chapter 2 and 3 the thesis moves away from preference evolution and looks at the evolution of prosociality in specific games. The mechanism enabling the evolution of prosocial behaviour in the games in Chapters 2 and 3 is repetition. If individuals interact repeatedly, they can play strategies that threaten to *not* play prosocially in the future rounds of the interaction, if their opponent does not play prosocially in the current round of the interaction.

Chapter 3 considers how selection for prosocial behaviour changes if individuals' actions are subject to errors, so that individuals who intend to behave a certain way sometimes accidentally behave differently. Somewhat counter-intuitively, errors can facilitate the evolution of prosocial behaviours. The reason behind this is that, as described above, with repetition, prosocial outcomes emerge if individuals play strategies which threaten to punish any deviations from prosocial behaviour. If deviating from prosocial behaviour thereby yields a selective disadvantage, without errors, evolution can bring a population to a situation where only individuals who play prosocially are left. However, once such a situation is reached, individuals never have to follow through on their threat of punishing deviations from prosocial behaviour. At that point, individuals who play prosocially but who lack the tendency to punish deviations from prosocial behaviour can *drift* into the population – there is no selective pressure favouring those who do have the tendency to punish. This poses a challenge to the evolutionary stability of prosocial outcomes, because once individuals who do not punish have spread in the population, they can be exploited by less prosocial mutant strategies, who then enjoy a selective advantage and can take over the population.

Errors solve this complication, as they ensure that even individuals who intend to play prosocially sometimes fail to do so. Therefore punishment behaviour remains payoff-relevant, which provides a selective pressure against the *drifting-in* of strategies who do not punish.

How strong this effect is, naturally, depends on the frequency at which such errors occur.

However, very frequent errors can also be detrimental to the evolution of prosocial outcomes, as the benefits of mutual prosocial interactions are diluted by the frequent triggering of this punishment behaviour. Chapter 3 looks at the tension between those effects, and finds that conditions for the evolution of prosocial behaviours are most favourable at moderate error rates.

In the model in Chapter 2, there are no errors, but compared to Chapter 3 individuals have a richer strategy set to choose from. Not only can they choose their behaviour in a given interaction, but they can also terminate their current interaction to find another individual to interact with.

Having this option to leave gives strategies a further tool to incentivise others to play prosocially. However, it also deals less prosocial strategies the possibility to escape potential punishment after free-riding on someone else's prosocial actions. Chapter 2 argues that on aggregate, the option to leave likely promotes the evolution of prosocial behaviour.

Chapters 2 and 3 both combine theoretical analyses and with stochastic simulations of the evolutionary processes that are modelled. One key task in analyzing such simulations is to identify whether the strategies that become and remain prevalent in a population constitute equilibria, or to see how much of a selective advantage potential invading mutants can have for non-equilibrium resident strategies. Doing this in a time-efficient manner for a large number of resident strategies across many simulation runs requires efficient algorithmic approaches. Chapter 4 presents such an algorithm games considered in Chapters 2 and 3.

Dutch summary

In brede zin bekijkt dit proefschrift verschillende mechanismen voor de evolutie van prosociaal gedrag, met behulp van de instrumenten van de evolutionaire speltheorie. De term prosociaal gedrag verwijst naar gedrag dat anderen ten goede komt, maar dat ten koste kan gaan van jezelf. In evolutionaire speltheorie worden interacties tussen individuen wiskundig beschreven. Afhankelijk van het gedrag van alle interacterende individuen, krijgen individuen beloningen toegewezen, die op hun beurt hun evolutionaire succes bepalen – individuen met hogere beloningen worden talrijker in de populatie, individuen met lagere beloningen worden minder talrijk. In dergelijke modellen is het dus raadselachtig hoe prosociaal gedrag kan evolueren, omdat het de ontvangers van het prosociale gedrag zijn, en niet noodzakelijkerwijs degenen die zich prosociaal gedragen, die het selectieve voordeel genieten.

Er zijn verschillende soorten evolutionaire mechanismen die een oplossing bieden voor deze puzzel. Eén zo'n mechanisme is *assortment*, wat betekent dat individuen bij voorkeur interacties aangaan met anderen die genetisch op hen lijken. Als dat het geval is, kunnen genen die coderen voor prosociaal gedrag zich verspreiden in een populatie, omdat de dragers van deze genen ook bij voorkeur prosociaal gedrag vertonen.

Een dergelijk mechanisme is *assortment*, wat betekent dat individuen bij voorkeur interacteren met anderen die genetisch vergelijkbaar zijn. Als dat het geval is, kunnen genen die coderen voor prosociaal gedrag zich verspreiden in een populatie, omdat de dragers van deze genen ook bij voorkeur de ontvangers zijn van prosociaal gedrag.

Het eerste deel van dit proefschrift (hoofdstuk 1) heeft betrekking op dit mechanisme. Specifiek beschouwt hoofdstuk 1 een model van voorkeursevolutie. In plaats van het modelleren van natuurlijke selectie voor gedrag in specifieke soorten interacties (spellen), proberen modellen van voorkeursevolutie regels – voorkeuren – te vinden die gedrag over grotere klassen van spellen verklaren. Zulke modellen kijken dan waarom of individuen die zich volgens bepaalde voorkeuren gedragen, bevoordeeld worden door natuurlijke selectie.

Hoofdstuk 1 focust op een bepaalde regel, bekend als "Homo Moralis" voorkeuren (Alger and Weibull (2013)). In een gegeven spel kiezen individuen die zich volgens deze regel gedragen een strategie die een combinatie van (1) hun eigen uitbetaling en (2) de gezamenlijke uitbetaling

die zij en hun partner zouden krijgen als ze allebei deze strategie zouden kiezen, maximaliseert. Alger and Weibull (2013) suggereert dat Homo Moralis voorkeuren, als het relatieve gewicht dat aan (2) wordt toegekend consistent is met de mate van assortment in de populatie, bijzonder succesvol zijn in de evolutie.

Hoofdstuk 1 laat zien dat een andere regel, bekend als "regulier altruïsme" het minstens even goed doet. In plaats van deze tweede, hypothetische term van "welke strategie zou leiden tot de ideale uitkomst als iedereen zich ernaar zou gedragen", maximaliseren reguliere altruïsten een combinatie van, wederom, (1) hun eigen beloning, maar nu (2) de beloning van hun partner. In termen van de evolutie van prosocialiteit ligt het verschil tussen Homo Moralis en reguliere altruïsten dus in de overwegingen die het prosociale gedrag teweegbrengen. Is het een Kantiaanse categorisch-imperatieve neiging tot sociaal efficiënt gedrag, zoals de Homo Moralis voorkeuren suggereren? Of is het, zoals de reguliere altruïstische preferenties suggereren, een directe zorg voor de payoff van het individu waarmee men op dat moment interacteert?

Hoofdstuk 1 laat zien dat in alle spellen waarin Homo Moralis voorkeuren een goed gedefinieerd gedrag voorschrijven, dit gedrag ook consistent is met reguliere altruïstische voorkeuren. Er zijn echter situaties waarin er geen strategie is die de voorkeuren van Homo Moralis maximaliseert, zodat het gedrag van Homo Moralis ongedefinieerd is. Hoofdstuk 1 laat zien dat natuurlijke manieren om de definitie Homo Moralis voorkeuren uit te breiden naar deze settings ofwel opnieuw leiden tot identiek gedrag voor Homo Moralis en reguliere altruïsten, of, als ze leiden tot verschillend gedrag, reguliere altruïsten winnen in directe concurrentie met Homo Moralis.

In hoofdstuk 2 en 3 neemt het proefschrift afstand van voorkeursevolutie en kijkt het naar de evolutie van prosocialiteit in specifieke spellen. Het mechanisme dat de evolutie van sociaal gedrag in de spellen in hoofdstuk 2 en 3 mogelijk maakt, is herhaling. Als individuen herhaaldelijk met elkaar interacteren, kunnen ze strategieën spelen die dreigen niet sociaal te spelen in de toekomstige rondes van de interactie, als hun tegenstander niet sociaal speelt in de huidige ronde van de interactie.

Hoofdstuk 3 beschouwt hoe selectie voor sociaal gedrag verandert als de acties van individuen onderhevig zijn aan fouten, zodat individuen die van plan zijn zich op een bepaalde manier te gedragen zich soms per ongeluk anders gedragen. Enigszins contra-intuïtief kunnen fouten de evolutie van sociaal gedrag bevorderen. De reden hiervoor is dat, zoals hierboven beschreven, bij herhaling sociale uitkomsten ontstaan als individuen strategieën spelen die afwijkingen van sociaal gedrag dreigen af te straffen. Als afwijken van sociaal gedrag daardoor een selectief nadeel oplevert, kan evolutie zonder fouten een populatie naar een situatie brengen waarin alleen individuen overblijven die sociaal spelen. Als zo'n situatie echter eenmaal is bereikt, hoeven individuen hun dreigement om afwijkingen van sociaal gedrag te bestraffen nooit op te volgen. Op dat moment kunnen individuen die sociaal spelen, maar niet de neiging hebben om afwijkingen van sociaal gedrag te bestraffen, de populatie binnen-

drijven – er is geen selectieve druk die diegenen bevoordeelt die wel de neiging hebben om te bestraffen. Dit vormt een uitdaging voor de evolutionaire stabiliteit van prosociale uitkomsten, want als individuen die niet straffen zich eenmaal hebben verspreid in de populatie, kunnen ze worden uitgebuit door minder prosociale mutanten, die dan een selectief voordeel hebben en de populatie kunnen overnemen.

Fouten lossen deze complicatie op, omdat ze ervoor zorgen dat zelfs individuen die van plan zijn om sociaal te spelen dat soms niet doen. Strafgedrag blijft dus payoff-relevant, wat zorgt voor een selectieve druk tegen het binnendrijven van strategieën die niet straffen.

Hoe sterk dit effect is, hangt natuurlijk af van de frequentie waarmee zulke fouten voorkomen. Zeer frequente fouten kunnen echter ook nadelig zijn voor de evolutie van prosociale uitkomsten, omdat de voordelen van wederzijdse prosociale interacties verwateren door het veelvuldig uitlokken van dit strafgedrag. Hoofdstuk 3 bekijkt het spanningsveld tussen deze effecten en stelt vast dat de omstandigheden voor de evolutie van sociaal gedrag het gunstigst zijn bij gematigde foutenpercentages.

In het model in hoofdstuk 2 zijn er geen fouten, maar vergeleken met hoofdstuk 3 hebben individuen een rijkere strategie set om uit te kiezen. Ze kunnen niet alleen hun gedrag in een bepaalde interactie kiezen, maar ze kunnen ook hun huidige interactie beëindigen om een ander individu te vinden om mee te interageren.

Deze optie om te vertrekken geeft strategieën een extra middel om anderen te stimuleren om sociaal te spelen. Het geeft minder sociale strategieën echter ook de mogelijkheid om een mogelijke straf te ontlopen als ze meeliften op de sociale acties van een ander. Hoofdstuk 3 stelt dat op aggregate de optie om te vertrekken waarschijnlijk de evolutie van sociaal gedrag bevordert.

Hoofdstuk 2 en 3 combineren beide theoretische analyses met stochastische simulaties van de gemodelleerde evolutionaire processen. Een belangrijke taak bij het analyseren van zulke simulaties is om vast te stellen of de strategieën die overheersen in een populatie evenwichten vormen, of om te zien hoeveel selectief voordeel potentiële binnenvallende mutanten kunnen hebben. Om dit op een tijdsefficiënte manier te doen voor een groot aantal residente strategieën over vele simulatieruns zijn efficiënte algoritmische benaderingen nodig. Hoofdstuk 4 presenteert zo'n algoritme voor de spellen die in hoofdstuk 2 en 3 zijn behandeld.

Bibliography

- Adams, D., 1979. *The Hitchhiker's Guide to the Galaxy* (Hitchhiker's Guide to the Galaxy, 1). Del Rey.
- Akdeniz, A., Graser, C., van Veelen, M., 2020. Homo moralis and regular altruists—preference evolution for when they disagree .
- Akdeniz, A., Graser, C., van Veelen, M., 2023. Homo moralis and regular altruists ii. Tinbergen Institute Discussion Paper 2023-025/I .
- Akdeniz, A., van Veelen, M., 2020. The cancellation effect at the group level. *Evolution* 74, 1246–1254.
- Akdeniz, A., van Veelen, M., 2023. Evolution and the ultimatum game. *Games and Economic Behavior* 142, 570–612.
- Alger, I., Weibull, J.W., 2012. A generalization of hamilton's rule—love others how much? *Journal of Theoretical Biology* 299, 42–54.
- Alger, I., Weibull, J.W., 2013. Homo moralis—preference evolution under incomplete information and assortative matching. *Econometrica* 81, 2269–2302.
- Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. *science* 211, 1390–1396.
- Bell, A.V., Richerson, P.J., McElreath, R., 2009. Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *Proceedings of the National Academy of Sciences* 106, 17671–17674.
- Bendor, J., Swistak, P., 1995. Types of evolutionary stability and the problem of cooperation. *Proceedings of the National Academy of Sciences* 92, 3596–3600.
- Bergman, T.J., Beehner, J.C., Cheney, D.L., Seyfarth, R.M., 2003. Hierarchical classification by rank and kinship in baboons. *Science* 302, 1234–1236.
- Binmore, K.G., Samuelson, L., 1992. Evolutionary stability in repeated games played by finite automata. *Journal of economic theory* 57, 278–305.
- Boyd, R., 1989. Mistakes allow evolutionary stability in the repeated prisoner's dilemma game. *Journal of theoretical Biology* 136, 47–56.
- Boyd, R., Lorberbaum, J.P., 1987. No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature* 327, 58–59.
- Call, J., Tomasello, M., 2008. Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences* 12, 187.

- Carmichael, H.L., MacLeod, W.B., 1997. Gift Giving and the Evolution of Cooperation. *International Economic Review* 38, 485.
- Cooper, D.J., 1996. Supergames played by finite automata with finite costs of complexity in an evolutionary setting. *Journal of Economic Theory* 68, 266–275.
- Dal Bó, P., Fréchette, G.R., 2011. The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review* 101, 411–429.
- Dal Bó, P., Fréchette, G.R., 2018. On the determinants of cooperation in infinitely repeated games: A survey. *Journal of Economic Literature* 56, 60–114.
- Dal Bó, P., Fréchette, G.R., 2019. Strategy choice in the infinitely repeated prisoner's dilemma. *American Economic Review* 109, 3929–3952.
- Dal Bó, P., Pujals, E.R., 2020. The evolutionary robustness of forgiveness and cooperation. arXiv preprint arXiv:1205.0958 – Updated version (2020) at http://www.econ.brown.edu/Faculty/Pedro_Dal_Bo/forgivenessandcooperation.pdf.
- Dawkins, R., 1976. *The selfish gene*. Oxford University Press.
- Fischer, A., Pollack, J., Thalmann, O., Nickel, B., Pääbo, S., 2006. Demographic history and genetic differentiation in apes. *Current Biology* 16, 1133–1138.
- Frank, R.H., 1987. If homo economicus could choose his own utility function, would he want one with a conscience? *The American Economic Review* , 593–604.
- Frank, R.H., 1988. *Passions within reason: The strategic role of the emotions*. WW Norton & Co.
- Fudenberg, D., Maskin, E., 1990. Evolution and cooperation in noisy repeated games. *The American Economic Review* 80, 274–279.
- Fujiwara-Greve, T., Okuno-Fujiwara, M., 2009. Voluntarily separable repeated prisoner's dilemma. *Review of Economic Studies* 76, 993–1021.
- Fujiwara-Greve, T., Okuno-Fujiwara, M., 2016. Diverse behavior patterns in a symmetric society with voluntary partnerships. Working paper .
- Fujiwara-Greve, T., Okuno-Fujiwara, M., 2017. Long-term cooperation and diverse behavior patterns under voluntary partnerships. Working paper .
- Fujiwara-Greve, T., Okuno-Fujiwara, M., Suzuki, N., 2015. Efficiency may improve when defectors exist. *Economic Theory* 60, 423–460.
- García, J., van Veelen, M., 2016. In and out of equilibrium I: Evolution of strategies in repeated games with discounting. *Journal of Economic Theory* 161, 161–189.
- García, J., van Veelen, M., 2018. No strategy can win in the repeated prisoner's dilemma: linking game theory and computer simulations. *Frontiers in Robotics and AI* 5, 102.
- Güth, W., Yaari, M., 1992. An evolutionary approach to explain reciprocal behavior in a simple strategic game. *U. Witt. Explaining Process and Change—Approaches to Evolutionary Economics*. Ann Arbor , 23–34.

- Hamilton, W.D., 1964. The genetical evolution of social behaviour. I and II. *Journal of theoretical biology* 7, 1–16, 17–52.
- Honhon, D., Hyndman, K., 2020. Flexibility and reputation in repeated prisoner's dilemma games. *Management Science* 66, 4998–5014.
- Hopcroft, J., 1971. An $n \log n$ algorithm for minimizing states in a finite automaton, in: *Theory of machines and computations*. Elsevier, pp. 189–196.
- Imhof, L.A., Fudenberg, D., Nowak, M.A., 2005. Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences* 102, 10797–10800.
- Izquierdo, S.S., Izquierdo, L.R., van Veelen, M., 2021. Repeated games with endogenous separation. Working paper .
- Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. *Econometrica: Journal of the Econometric Society* , 29–56.
- Kay, T., Keller, L., Lehmann, L., 2020. The evolution of altruism and the serial rediscovery of the role of relatedness. *Proceedings of the National Academy of Sciences* 117, 28894–28898.
- Krupenye, C., Call, J., 2019. Theory of mind in animals: Current and future directions. *Wiley Interdisciplinary Reviews: Cognitive Science* 10, e1503.
- Langergraber, K., Schubert, G., Rowney, C., Wrangham, R., Zommers, Z., Vigilant, L., 2011. Genetic differentiation and the evolution of cooperation in chimpanzees and humans. *Proceedings of the Royal Society B: Biological Sciences* 278, 2546–2552.
- Langergraber, K.E., Mitani, J.C., Vigilant, L., 2007. The limited impact of kinship on cooperation in wild chimpanzees. *Proceedings of the National Academy of Sciences* 104, 7786–7790.
- Lee, N., 2020. An experiment: Voluntary separation in indefinitely repeated prisoner's dilemma game. Available at SSRN 3207669 .
- Lieberman, D., Tooby, J., Cosmides, L., 2007. The architecture of human kin detection. *Nature* 445, 727–731.
- Miettinen, T., Kosfeld, M., Fehr, E., Weibull, J., 2020. Revealed preferences in a sequential prisoners' dilemma: A horse-race between six utility functions. *Journal of Economic Behavior & Organization* 173, 1–25.
- Nowak, M.A., 2006. *Evolutionary dynamics*. Harvard University Press.
- Nowak, M.A., Sasaki, A., Taylor, C., Fudenberg, D., 2004. Emergence of cooperation and evolutionary stability in finite populations. *Nature* 428, 646–650.
- Parr, L.A., de Waal, F.B., 1999. Visual kin recognition in chimpanzees. *Nature* 399, 647–648.
- Penn, D.C., Holyoak, K.J., Povinelli, D.J., 2008. Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences* 31, 109–130.
- Robson, A.J., 1990. Efficiency in evolutionary games: Darwin, nash and the secret handshake. *Journal of theoretical Biology* 144, 379–396.

- Scally, A., Yngvadottir, B., Xue, Y., Ayub, Q., Durbin, R., Tyler-Smith, C., 2013. A genome-wide survey of genetic variation in gorillas using reduced representation sequencing. *PLoS One* 8.
- Selten, R., Hammerstein, P., 1984. Gaps in harley's argument on evolutionarily stable learning rules and in the logic of "tit for tat". *Behavioral and Brain Sciences* 7, 115–116.
- Silk, J.B., 2009. Nepotistic cooperation in non-human primate groups. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 3243–3254.
- Sugden, R., 1986. *The economics of welfare, rights and co-operation*. Oxford: Basil Blackwell.
- Taylor, P.D., 1992a. Altruism in viscous populations—an inclusive fitness model. *Evolutionary Ecology* 6, 352–356.
- Taylor, P.D., 1992b. Inclusive fitness in a homogeneous environment, in: *Proceedings of the Royal Society of London. Series B: Biological Sciences*, The Royal Society. pp. 299–302.
- van Veelen, M., 2006. Why kin and group selection models may not be enough to explain human other-regarding behaviour. *Journal of Theoretical Biology* 242, 790–797.
- van Veelen, M., 2009. Group selection, kin selection, altruism and cooperation: when inclusive fitness is right and when it can be wrong. *Journal of Theoretical Biology* 259, 589–600.
- van Veelen, M., 2011. The replicator dynamics with n players and population structure. *Journal of Theoretical Biology* 276, 78–85.
- van Veelen, M., 2012. Robustness against indirect invasions. *Games and Economic Behavior* 74, 382–393.
- van Veelen, M., 2018. Can Hamilton's rule be violated? *eLife* 7, e41901.
- van Veelen, M., Allen, B., Hoffman, M., Simon, B., Veller, C., 2017. Hamilton's rule. *Journal of Theoretical Biology* 414, 176–230.
- van Veelen, M., García, J., 2019. In and out of equilibrium II: Evolution in repeated games with discounting and complexity costs. *Games and Economic Behavior* 115, 113–130.
- van Veelen, M., García, J., Rand, D.G., Nowak, M.A., 2012. Direct reciprocity in structured populations. *Proceedings of the National Academy of Sciences USA* 109, 9929–9934.
- Vesely, F., Yang, C.L., 2010. On optimal and neutrally stable population equilibrium in voluntary partnership prisoner's dilemma games. SSRN .
- Vesely, F., Yang, C.L., 2012. Breakup, secret handshake and neutral stability in repeated prisoner's dilemma with option to leave: A note. SSRN .
- Volij, O., 2002. In defense of defect. *Games and Economic Behavior* 39, 309–321.
- Weibull, J.W., 1997. *Evolutionary game theory*. MIT press.
- Wilson, D.S., Pollock, G.B., Dugatkin, L.A., 1992. Can altruism evolve in purely viscous populations? *Evolutionary Ecology* 6, 331–341.

- Wu, J., Axelrod, R., 1995. How to cope with noise in the iterated prisoner's dilemma. *Journal of Conflict resolution* 39, 183–189.
- Young, H.P., 1993. The evolution of conventions. *Econometrica: Journal of the Econometric Society* , 57–84.
- Zhang, B.Y., Fan, S.J., Li, C., Zheng, X.D., Bao, J.Z., Cressman, R., Tao, Y., 2016. Opting out against defection leads to stable coexistence with cooperation. *Scientific reports* 6, 35902.