International Institute for
Applied Systems Analysis
IIASA www.iiasa.ac.at

The Emergence of Altruistic Punishment: Via Freedom to Enforcement

Hauert, C., Traulsen, A., Brandt, H., Nowak, M.A. and Sigmund, K.

IIASA Interim Report
December 2007

International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report IR-07-053

The Emergence of Altruistic Punishment: Via Freedom to Enforcement

Christoph Hauert (christoph_hauert@harvard.edu)
Arne Traulsen (traulsen@fas.harvard.edu)
Hannelore Brandt (hannelore.brandt@wu-wien.ac.at)
Martin A. Nowak (martin_nowak@harvard.edu)
Karl Sigmund (karl.sigmund@univie.ac.at)

Approved by

Ulf Dieckmann
Leader, Evolution and Ecology Program

December 2007

IIASA Studies in Adaptive Dynamics No. 137

The Evolution and Ecology Program at IIASA fosters the development of new mathematical and conceptual techniques for understanding the evolution of complex adaptive systems.

Focusing on these long-term implications of adaptive processes in systems of limited growth, the Evolution and Ecology Program brings together scientists and institutions from around the world with IIASA acting as the central node.

Scientific progress within the network is collected in the IIASA Studies in Adaptive Dynamics series.

No. 1 Metz JAJ, Geritz SAH, Meszéna G, Jacobs FJA, van Heerwaarden JS: *Adaptive Dynamics: A Geometrical Study of the Consequences of Nearly Faithful Reproduction.* IIASA Working Paper WP-95-099 (1995). van Strien SJ, Verduyn Lunel SM (eds): Stochastic and Spatial Structures of Dynamical Systems, Proceedings of the Royal Dutch Academy of Science (KNAW Verhandelingen), North Holland, Amsterdam, pp. 183-231 (1996).

No. 2 Dieckmann U, Law R: *The Dynamical Theory of Coevolution: A Derivation from Stochastic Ecological Processes.* IIASA Working Paper WP-96-001 (1996). Journal of Mathematical Biology 34:579-612 (1996).

No. 3 Dieckmann U, Marrow P, Law R: *Evolutionary Cycling of Predator-Prey Interactions: Population Dynamics and the Red Queen.* IIASA Preprint (1995). Journal of Theoretical Biology 176:91-102 (1995).

No. 4 Marrow P, Dieckmann U, Law R: *Evolutionary Dynamics of Predator-Prey Systems: An Ecological Perspective.* IIASA Working Paper WP-96-002 (1996). Journal of Mathematical Biology 34:556-578 (1996).

No. 5 Law R, Marrow P, Dieckmann U: *On Evolution under Asymmetric Competition.* IIASA Working Paper WP-96-003 (1996). Evolutionary Ecology 11:485-501 (1997).

No. 6 Metz JAJ, Mylius SD, Diekmann O: *When Does Evolution Optimize? On the Relation Between Types of Density Dependence and Evolutionarily Stable Life History Parameters.* IIASA Working Paper WP-96-004 (1996).

No. 7 Ferrière R, Gatto M: *Lyapunov Exponents and the Mathematics of Invasion in Oscillatory or Chaotic Populations.* Theoretical Population Biology 48:126-171 (1995).

No. 8 Ferrière R, Fox GA: *Chaos and Evolution.* IIASA Preprint (1996). Trends in Ecology and Evolution 10:480-485 (1995).

No. 9 Ferrière R, Michod RE: *The Evolution of Cooperation in Spatially Heterogeneous Populations.* IIASA Working Paper WP-96-029 (1996). The American Naturalist 147:692-717 (1996).

No. 10 van Dooren TJM, Metz JAJ: *Delayed Maturation in Temporally Structured Populations with Non-Equilibrium Dynamics.* IIASA Working Paper WP-96-070 (1996). Journal of Evolutionary Biology 11:41-62 (1998).

No. 11 Geritz SAH, Metz JAJ, Kisdi É, Meszéna G: *The Dynamics of Adaptation and Evolutionary Branching.* IIASA Working Paper WP-96-077 (1996). Physical Review Letters 78:2024-2027 (1997).

No. 12 Geritz SAH, Kisdi É, Meszéna G, Metz JAJ: *Evolutionary Singular Strategies and the Adaptive Growth and Branching of the Evolutionary Tree.* IIASA Working Paper WP-96-114 (1996). Evolutionary Ecology 12:35-57 (1998).

No. 13 Heino M, Metz JAJ, Kaitala V: *Evolution of Mixed Maturation Strategies in Semelparous Life-Histories: The Crucial Role of Dimensionality of Feedback Environment.* IIASA Working Paper WP-96-126 (1996). Philosophical Transactions of the Royal Society of London Series B 352:1647-1655 (1997).

No. 14 Dieckmann U: *Can Adaptive Dynamics Invade?* IIASA Working Paper WP-96-152 (1996). Trends in Ecology and Evolution 12:128-131 (1997).

No. 15 Meszéna G, Czibula I, Geritz SAH: *Adaptive Dynamics in a 2-Patch Environment: A Simple Model for Allopatric and Parapatric Speciation.* IIASA Interim Report IR-97-001 (1997). Journal of Biological Systems 5:265-284 (1997).

No. 16 Heino M, Metz JAJ, Kaitala V: *The Enigma of Frequency-Dependent Selection.* IIASA Interim Report IR-97-061 (1997). Trends in Ecology and Evolution 13:367-370 (1998).

No. 17 Heino M: *Management of Evolving Fish Stocks.* IIASA Interim Report IR-97-062 (1997). Canadian Journal of Fisheries and Aquatic Sciences 55:1971-1982 (1998).

No. 18 Heino M: *Evolution of Mixed Reproductive Strategies in Simple Life-History Models.* IIASA Interim Report IR-97-063 (1997).

No. 19 Geritz SAH, van der Meijden E, Metz JAJ: *Evolutionary Dynamics of Seed Size and Seedling Competitive Ability.* IIASA Interim Report IR-97-071 (1997). Theoretical Population Biology 55:324-343 (1999).

No. 20 Galis F, Metz JAJ: *Why Are There So Many Cichlid Species? On the Interplay of Speciation and Adaptive Radiation.* IIASA Interim Report IR-97-072 (1997). Trends in Ecology and Evolution 13:1-2 (1998).

No. 21 Boerlijst MC, Nowak MA, Sigmund K: *Equal Pay for all Prisoners/ The Logic of Contrition.* IIASA Interim Report IR-97-073 (1997). American Mathematical Society Monthly 104:303-307 (1997). Journal of Theoretical Biology 185:281-293 (1997).

No. 22 Law R, Dieckmann U: *Symbiosis Without Mutualism and the Merger of Lineages in Evolution.* IIASA Interim Report IR-97-074 (1997). Proceedings of the Royal Society of London Series B 265:1245-1253 (1998).

No. 23 Klinkhamer PGL, de Jong TJ, Metz JAJ: *Sex and Size in Cosexual Plants.* IIASA Interim Report IR-97-078 (1997). Trends in Ecology and Evolution 12:260-265 (1997).

No. 24 Fontana W, Schuster P: *Shaping Space: The Possible and the Attainable in RNA Genotype-Phenotype Mapping.* IIASA Interim Report IR-98-004 (1998). Journal of Theoretical Biology 194:491-515 (1998).

No. 25 Kisdi É, Geritz SAH: *Adaptive Dynamics in Allele Space: Evolution of Genetic Polymorphism by Small Mutations in a Heterogeneous Environment.* IIASA Interim Report IR-98-038 (1998). Evolution 53:993-1008 (1999).

No. 26 Fontana W, Schuster P: *Continuity in Evolution: On the Nature of Transitions.* IIASA Interim Report IR-98-039 (1998). Science 280:1451-1455 (1998).

No. 27 Nowak MA, Sigmund K: *Evolution of Indirect Reciprocity by Image Scoring/ The Dynamics of Indirect Reciprocity.* IIASA Interim Report IR-98-040 (1998). Nature 393:573-577 (1998). Journal of Theoretical Biology 194:561-574 (1998).

No. 28 Kisdi É: *Evolutionary Branching Under Asymmetric Competition.* IIASA Interim Report IR-98-045 (1998). Journal of Theoretical Biology 197:149-162 (1999).

No. 29 Berger U: *Best Response Adaptation for Role Games.* IIASA Interim Report IR-98-086 (1998).

No. 30 van Dooren TJM: *The Evolutionary Ecology of Dominance-Recessivity.* IIASA Interim Report IR-98-096 (1998). Journal of Theoretical Biology 198:519-532 (1999).

No. 31 Dieckmann U, O'Hara B, Weisser W: *The Evolutionary Ecology of Dispersal.* IIASA Interim Report IR-98-108 (1998). Trends in Ecology and Evolution 14:88-90 (1999).

No. 32 Sigmund K: *Complex Adaptive Systems and the Evolution of Reciprocation.* IIASA Interim Report IR-98-100 (1998). Ecosystems 1:444-448 (1998).

No. 33 Posch M, Pichler A, Sigmund K: *The Efficiency of Adapting Aspiration Levels.* IIASA Interim Report IR-98-103 (1998). Proceedings of the Royal Society London Series B 266:1427-1435 (1999).

No. 34 Mathias A, Kisdi É: *Evolutionary Branching and Coexistence of Germination Strategies.* IIASA Interim Report IR-99-014 (1999).

No. 35 Dieckmann U, Doebeli M: *On the Origin of Species by Sympatric Speciation.* IIASA Interim Report IR-99-013 (1999). Nature 400:354-357 (1999).

No. 36 Metz JAJ, Gyllenberg M: *How Should We Define Fitness in Structured Metapopulation Models? Including an Application to the Calculation of Evolutionarily Stable Dispersal Strategies.* IIASA Interim Report IR-99-019 (1999). Proceedings of the Royal Society of London Series B 268:499-508 (2001).

No. 37 Gyllenberg M, Metz JAJ: *On Fitness in Structured Metapopulations.* IIASA Interim Report IR-99-037 (1999). Journal of Mathematical Biology 43:545-560 (2001).

No. 38 Meszéna G, Metz JAJ: *Species Diversity and Population Regulation: The Importance of Environmental Feedback Dimensionality.* IIASA Interim Report IR-99-045 (1999).

No. 39 Kisdi É, Geritz SAH: *Evolutionary Branching and Sympatric Speciation in Diploid Populations.* IIASA Interim Report IR-99-048 (1999).

No. 40 Ylikarjula J, Heino M, Dieckmann U: *Ecology and Adaptation of Stunted Growth in Fish.* IIASA Interim Report IR-99-050 (1999). Evolutionary Ecology 13:433-453 (1999).

No. 41 Nowak MA, Sigmund K: *Games on Grids.* IIASA Interim Report IR-99-038 (1999). Dieckmann U, Law R, Metz JAJ (eds): The Geometry of Ecological Interactions: Simplifying Spatial Complexity, Cambridge University Press, Cambridge, UK, pp. 135-150 (2000).

No. 42 Ferrière R, Michod RE: *Wave Patterns in Spatial Games and the Evolution of Cooperation.* IIASA Interim Report IR-99-041 (1999). Dieckmann U, Law R, Metz JAJ (eds): The Geometry of Ecological Interactions: Simplifying Spatial Complexity, Cambridge University Press, Cambridge, UK, pp. 318-332 (2000).

No. 43 Kisdi É, Jacobs FJA, Geritz SAH: *Red Queen Evolution by Cycles of Evolutionary Branching and Extinction.* IIASA Interim Report IR-00-030 (2000). Selection 2:161-176 (2001).

No. 44 Meszéna G, Kisdi É, Dieckmann U, Geritz SAH, Metz JAJ: *Evolutionary Optimisation Models and Matrix Games in the Unified Perspective of Adaptive Dynamics.* IIASA Interim Report IR-00-039 (2000). Selection 2:193-210 (2001).

No. 45 Parvinen K, Dieckmann U, Gyllenberg M, Metz JAJ: *Evolution of Dispersal in Metapopulations with Local Density Dependence and Demographic Stochasticity.* IIASA Interim Report IR-00-035 (2000). Journal of Evolutionary Biology 16:143-153 (2003).

No. 46 Doebeli M, Dieckmann U: *Evolutionary Branching and Sympatric Speciation Caused by Different Types of Ecological Interactions.* IIASA Interim Report IR-00-040 (2000). The American Naturalist 156:S77-S101 (2000).

No. 47 Heino M, Hanski I: *Evolution of Migration Rate in a Spatially Realistic Metapopulation Model.* IIASA Interim Report IR-00-044 (2000). The American Naturalist 157:495-511 (2001).

No. 48 Gyllenberg M, Parvinen K, Dieckmann U: *Evolutionary Suicide and Evolution of Dispersal in Structured Metapopulations.* IIASA Interim Report IR-00-056 (2000). Journal of Mathematical Biology 45:79-105 (2002).

No. 49 van Dooren TJM: *The Evolutionary Dynamics of Direct Phenotypic Overdominance: Emergence Possible, Loss Probable.* IIASA Interim Report IR-00-048 (2000). Evolution 54:1899-1914 (2000).

No. 50 Nowak MA, Page KM, Sigmund K: *Fairness Versus Reason in the Ultimatum Game.* IIASA Interim Report IR-00-57 (2000). Science 289:1773-1775 (2000).

No. 51 de Feo O, Ferrière R: *Bifurcation Analysis of Population Invasion: On-Off Intermittency and Basin Riddling.* IIASA Interim Report IR-00-074 (2000). International Journal of Bifurcation and Chaos 10:443-452 (2000).

No. 52 Heino M, Laaka-Lindberg S: *Clonal Dynamics and Evolution of Dormancy in the Leafy Hepatic Lophozia Silvicola.* IIASA Interim Report IR-01-018 (2001). Oikos 94:525-532 (2001).

No. 53 Sigmund K, Hauert C, Nowak MA: *Reward and Punishment in Minigames.* IIASA Interim Report IR-01-031 (2001). Proceedings of the National Academy of Sciences of the USA 98:10757-10762 (2001).

No. 54 Hauert C, De Monte S, Sigmund K, Hofbauer J: *Oscillations in Optional Public Good Games.* IIASA Interim Report IR-01-036 (2001).

No. 55 Ferrière R, Le Galliard J: *Invasion Fitness and Adaptive Dynamics in Spatial Population Models.* IIASA Interim Report IR-01-043 (2001). Clobert J, Dhondt A, Danchin E, Nichols J (eds): Dispersal, Oxford University Press, pp. 57-79 (2001).

No. 56 de Mazancourt C, Loreau M, Dieckmann U: *Can the Evolution of Plant Defense Lead to Plant-Herbivore Mutualism?* IIASA Interim Report IR-01-053 (2001). The American Naturalist 158:109-123 (2001).

No. 57 Claessen D, Dieckmann U: *Ontogenetic Niche Shifts and Evolutionary Branching in Size-Structured Populations.* IIASA Interim Report IR-01-056 (2001). Evolutionary Ecology Research 4:189-217 (2002).

No. 58 Brandt H: *Correlation Analysis of Fitness Landscapes.* IIASA Interim Report IR-01-058 (2001).

No. 59 Dieckmann U: *Adaptive Dynamics of Pathogen-Host Interacations.* IIASA Interim Report IR-02-007 (2002). Dieckmann U, Metz JAJ, Sabelis MW, Sigmund K (eds): Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management, Cambridge University Press, Cambridge, UK, pp. 39-59 (2002).

No. 60 Nowak MA, Sigmund K: *Super- and Coinfection: The Two Extremes.* IIASA Interim Report IR-02-008 (2002). Dieckmann U, Metz JAJ, Sabelis MW, Sigmund K (eds): Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management, Cambridge University Press, Cambridge, UK, pp. 124-137 (2002).

No. 61 Sabelis MW, Metz JAJ: *Evolution Management: Taking Stock - Relating Theory to Experiment.* IIASA Interim Report IR-02-009 (2002). Dieckmann U, Metz JAJ, Sabelis MW, Sigmund K (eds): Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management, Cambridge University Press, Cambridge, UK, pp. 379-398 (2002).

No. 62 Cheptou P, Dieckmann U: *The Evolution of Self-Fertilization in Density-Regulated Populations .* IIASA Interim Report IR-02-024 (2002). Proceedings of the Royal Society of London Series B 269:1177-1186 (2002).

No. 63 Bürger R: *Additive Genetic Variation Under Intraspecific Competition and Stabilizing Selection: A Two-Locus Study.* IIASA Interim Report IR-02-013 (2002). Theoretical Population Biology 61:197-213 (2002).

No. 64 Hauert C, De Monte S, Hofbauer J, Sigmund K: *Volunteering as Red Queen Mechanism for Co-operation in Public Goods Games.* IIASA Interim Report IR-02-041 (2002). Science 296:1129-1132 (2002).

No. 65 Dercole F, Ferrière R, Rinaldi S: *Ecological Bistability and Evolutionary Reversals under Asymmetrical Competition.* IIASA Interim Report IR-02-053 (2002). Evolution 56:1081-1090 (2002).

No. 66 Dercole F, Rinaldi S: *Evolution of Cannibalistic Traits: Scenarios Derived from Adaptive Dynamics.* IIASA Interim Report IR-02-054 (2002). Theoretical Population Biology 62:365-374 (2002).

No. 67 Bürger R, Gimelfarb A: *Fluctuating Environments and the Role of Mutation in Maintaining Quantitative Genetic Variation.* IIASA Interim Report IR-02-058 (2002). Genetical Research 80:31-46 (2002).

No. 68 Bürger R: *On a Genetic Model of Intraspecific Competition and Stabilizing Selection.* IIASA Interim Report IR-02-062 (2002). Amer. Natur. 160:661-682 (2002).

No. 69 Doebeli M, Dieckmann U: *Speciation Along Environmental Gradients.* IIASA Interim Report IR-02-079 (2002). Nature 421:259-264 (2003).

No. 70 Dercole F, Irisson J, Rinaldi S: *Bifurcation Analysis of a Prey-Predator Coevolution Model.* IIASA Interim Report IR-02-078 (2002). SIAM Journal on Applied Mathematics 63:1378-1391 (2003).

No. 71 Le Galliard J, Ferrière R, Dieckmann U: *The Adaptive Dynamics of Altruism in Spatially Heterogeneous Populations.* IIASA Interim Report IR-03-006 (2003). Evolution 57:1-17 (2003).

No. 72 Taborsky B, Dieckmann U, Heino M: *Unexpected Discontinuities in Life-History Evolution under Size-Dependent Mortality.* IIASA Interim Report IR-03-004 (2003). Proceedings of the Royal Society of London Series B 270:713-721 (2003).

No. 73 Gardmark A, Dieckmann U, Lundberg P: *Life-History Evolution in Harvested Populations: The Role of Natural Predation.* IIASA Interim Report IR-03-008 (2003). Evolutionary Ecology Research 5:239-257 (2003).

No. 74 Mizera F, Meszéna G: *Spatial Niche Packing, Character Displacement and Adaptive Speciation Along an Environmental Gradient.* IIASA Interim Report IR-03-062 (2003). Evolutionary Ecology Research 5:363-382 (2003).

No. 75 Dercole F: *Remarks on Branching-Extinction Evolutionary Cycles.* IIASA Interim Report IR-03-077 (2003). Journal of Mathematical Biology 47:569-580 (2003).

No. 76 Hofbauer J, Sigmund K: *Evolutionary Game Dynamics.* IIASA Interim Report IR-03-078 (2003). Bulletin of the American Mathematical Society 40:479-519 (2003).

No. 77 Ernande B, Dieckmann U, Heino M: *Adaptive Changes in Harvested Populations: Plasticity and Evolution of Age and Size at Maturation.* IIASA Interim Report IR-03-058 (2003). Proceedings of the Royal Society of London Series B-Biological Sciences 271:415-423 (2004).

No. 78 Hanski I, Heino M: *Metapopulation-Level Adaptation of Insect Host Plant Preference and Extinction-Colonization Dynamics in Heterogeneous Landscapes.* IIASA Interim Report IR-03-028 (2003). Theoretical Population Biology 63:309-338 (2003).

No. 79 van Doorn G, Dieckmann U, Weissing FJ: *Sympatric Speciation by Sexual Selection: A Critical Re-Evaluation.* IIASA Interim Report IR-04-003 (2004). American Naturalist 163:709-725 (2004).

No. 80 Egas M, Dieckmann U, Sabelis MW: *Evolution Restricts the Coexistence of Specialists and Generalists - the Role of Trade-off Structure.* IIASA Interim Report IR-04-004 (2004). American Naturalist 163:518-531 (2004).

No. 81 Ernande B, Dieckmann U: *The Evolution of Phenotypic Plasticity in Spatially Structured Environments: Implications of Intraspecific Competition, Plasticity Costs, and Environmental Characteristics.* IIASA Interim Report IR-04-006 (2004). Journal of Evolutionary Biology 17:613-628 (2004).

No. 82 Cressman R, Hofbauer J: *Measure Dynamics on a One-Dimensional Continuous Trait Space: Theoretical Foundations for Adaptive Dynamics.* IIASA Interim Report IR-04-016 (2004).

No. 83 Cressman R: *Dynamic Stability of the Replicator Equation with Continuous Strategy Space.* IIASA Interim Report IR-04-017 (2004).

No. 84 Ravigné V, Olivieri I, Dieckmann U: *Implications of Habitat Choice for Protected Polymorphisms.* IIASA Interim Report IR-04-005 (2004). Evolutionary Ecology Research 6:125-145 (2004).

No. 85 Nowak MA, Sigmund K: *Evolutionary Dynamics of Biological Games.* IIASA Interim Report IR-04-013 (2004). Science 303:793-799 (2004).

No. 86 Vukics A, Asbóth J, Meszéna G: *Speciation in Multidimensional Evolutionary Space.* IIASA Interim Report IR-04-028 (2004). Physical Review 68:041-903 (2003).

No. 87 de Mazancourt C, Dieckmann U: *Trade-off Geometries and Frequency-dependent Selection.* IIASA Interim Report IR-04-039 (2004). American Naturalist 164:765-778 (2004).

No. 88 Cadet CR, Metz JAJ, Klinkhamer PGL: *Size and the Not-So-Single Sex: Disentangling the Effects of Size on Sex Allocation.* IIASA Interim Report IR-04-084 (2004). American Naturalist 164:779-792 (2004).

No. 89 Rueffler C, van Dooren TJM, Metz JAJ: *Adaptive Walks on Changing Landscapes: Levins' Approach Extended.* IIASA Interim Report IR-04-083 (2004). Theoretical Population Biology 65:165-178 (2004).

No. 90 de Mazancourt C, Loreau M, Dieckmann U: *Understanding Mutualism When There is Adaptation to the Partner.* IIASA Interim Report IR-05-016 (2005). Journal of Ecology 93:305-314 (2005).

No. 91 Dieckmann U, Doebeli M: *Pluralism in Evolutionary Theory.* IIASA Interim Report IR-05-017 (2005). Journal of Evolutionary Biology 18:1209-1213 (2005).

No. 92 Doebeli M, Dieckmann U, Metz JAJ, Tautz D: *What We Have Also Learned: Adaptive Speciation is Theoretically Plausible.* IIASA Interim Report IR-05-018 (2005). Evolution 59:691-695 (2005).

No. 93 Egas M, Sabelis MW, Dieckmann U: *Evolution of Specialization and Ecological Character Displacement of Herbivores Along a Gradient of Plant Quality.* IIASA Interim Report IR-05-019 (2005). Evolution 59:507-520 (2005).

No. 94 Le Galliard J, Ferrière R, Dieckmann U: *Adaptive Evolution of Social Traits: Origin, Trajectories, and Correlations of Altruism and Mobility.* IIASA Interim Report IR-05-020 (2005). American Naturalist 165:206-224 (2005).

No. 95 Doebeli M, Dieckmann U: *Adaptive Dynamics as a Mathematical Tool for Studying the Ecology of Speciation Processes.* IIASA Interim Report IR-05-022 (2005). Journal of Evolutionary Biology 18:1194-1200 (2005).

No. 96 Brandt H, Sigmund K: *The Logic of Reprobation: Assessment and Action Rules for Indirect Reciprocity.* IIASA Interim Report IR-04-085 (2004). Journal of Theoretical Biology 231:475-486 (2004).

No. 97 Hauert C, Haiden N, Sigmund K: *The Dynamics of Public Goods.* IIASA Interim Report IR-04-086 (2004). Discrete and Continuous Dynamical Systems - Series B 4:575-587 (2004).

No. 98 Meszéna G, Gyllenberg M, Jacobs FJA, Metz JAJ: *Link Between Population Dynamics and Dynamics of Darwinian Evolution.* IIASA Interim Report IR-05-026 (2005). Physical Review Letters 95:Article 078105 (2005).

No. 99 Meszéna G: *Adaptive Dynamics: The Continuity Argument.* IIASA Interim Report IR-05-032 (2005).

No. 100 Brännström NA, Dieckmann U: *Evolutionary Dynamics of Altruism and Cheating Among Social Amoebas.* IIASA Interim Report IR-05-039 (2005). Proceedings of the Royal Society London Series B 272:1609-1616 (2005).

No. 101 Meszéna G, Gyllenberg M, Pasztor L, Metz JAJ: *Competitive Exclusion and Limiting Similarity: A Unified Theory.* IIASA Interim Report IR-05-040 (2005).

No. 102 Szabo P, Meszéna G: *Limiting Similarity Revisited.* IIASA Interim Report IR-05-050 (2005).

No. 103 Krakauer DC, Sasaki A: *The Greater than Two-Fold Cost of Integration for Retroviruses.* IIASA Interim Report IR-05-069 (2005).

No. 104 Metz JAJ: *Eight Personal Rules for Doing Science.* IIASA Interim Report IR-05-073 (2005). Journal of Evolutionary Biology 18:1178-1181 (2005).

No. 105 Beltman JB, Metz JAJ: *Speciation: More Likely Through a Genetic or Through a Learned Habitat Preference?* IIASA Interim Report IR-05-072 (2005). Proceedings of the Royal Society of London Series B 272:1455-1463 (2005).

No. 106 Durinx M, Metz JAJ: *Multi-type Branching Processes and Adaptive Dynamics of Structured Populations.* IIASA Interim Report IR-05-074 (2005). Haccou P, Jager P, Vatutin V (eds): Branching Processes: Variation, Growth and Extinction of Populations, Cambridge University Press, Cambridge, UK, pp. 266-278 (2005).

No. 107 Brandt H, Sigmund K: *The Good, the Bad and the Discriminator - Errors in Direct and Indirect Reciprocity.* IIASA Interim Report IR-05-070 (2005). Journal of Theoretical Biology 239:183-194 (2006).

No. 108 Brandt H, Hauert C, Sigmund K: *Punishing and Abstaining for Public Goods.* IIASA Interim Report IR-05-071 (2005). Proceedings of the National Academy of Sciences of the United States of America 103:495-497 (2006).

No. 109 Ohtsuki A, Sasaki A: *Epidemiology and Disease-Control Under Gene-for-Gene Plant-Pathogen Interaction.* IIASA Interim Report IR-05-068 (2005).

No. 110 Brandt H, Sigmund K: *Indirect Reciprocity, Image-Scoring, and Moral Hazard.* IIASA Interim Report IR-05-078 (2005). Proceedings of the National Academy of Sciences of the United States of America 102:2666-2670 (2005).

No. 111 Nowak MA, Sigmund K: *Evolution of Indirect Reciprocity.* IIASA Interim Report IR-05-079 (2005). Nature 437:1292-1298 (2005).

No. 112 Kamo M, Sasaki A: *Evolution Towards Multi-Year Periodicity in Epidemics.* IIASA Interim Report IR-05-080 (2005). Ecology Letters 8:378-385 (2005).

No. 113 Dercole F, Ferrière R, Gragnani A, Rinaldi S: *Coevolution of Slow-fast Populations: Evolutionary Sliding, Evolutionoary Pseudo-equilibria, and Complex Red Queen Dynamics.* IIASA Interim Report IR-06-006 (2006). Proceedings of the Royal Society B 273:983-990 (2006).

No. 114 Dercole F: *Border Collision Bifurcations in the Evolution of Mutualistic Interactions.* IIASA Interim Report IR-05-083 (2005). International Journal of Bifurcation and Chaos 15:2179-2190 (2005).

No. 115 Dieckmann U, Heino M, Parvinen K: *The Adaptive Dynamics of Function-Valued Traits.* IIASA Interim Report IR-06-036 (2006). Journal of Theoretical Biology 241:370-389 (2006).

No. 116 Dieckmann U, Metz JAJ: *Surprising Evolutionary Predictions from Enhanced Ecological Realism.* IIASA Interim Report IR-06-037 (2006). Theoretical Population Biology 69:263-281 (2006).

No. 117 Dieckmann U, Brännström NA, HilleRisLambers R, Ito H: *The Adaptive Dynamics of Community Structure.* IIASA Interim Report IR-06-038 (2006). Takeuchi Y, Iwasa Y, Sato K (eds): Mathematics for Ecology and Environmental Sciences, Springer, Berlin Heidelberg, pp. 145-177 (2007).

No. 118 Gardmark A, Dieckmann U: *Disparate Maturation Adaptations to Size-dependent Mortality.* IIASA Interim Report IR-06-039 (2006). Proceedings of the Royal Society London Series B 273:2185-2192 (2006).

No. 119 van Doorn G, Dieckmann U: *The Long-term Evolution of Multi-locus Traits Under Frequency-dependent Disruptive Selection.* IIASA Interim Report IR-06-041 (2006). Evolution 60:2226-2238 (2006).

No. 120 Doebeli M, Blok HJ, Leimar O, Dieckmann U: *Multimodal Pattern Formation in Phenotype Distributions of Sexual Populations.* IIASA Interim Report IR-06-046 (2006). Proceedings of the Royal Society London Series B 274:347-357 (2007).

No. 121 Dunlop ES, Shuter BJ, Dieckmann U: *The Demographic and Evolutionary Consequences of Selective Mortality: Predictions from an Eco-genetic Model of the Smallmouth Bass.* IIASA Interim Report IR-06-060 (2006). Transactions of the American Fisheries Society 136:749-765 (2007).

No. 122 Metz JAJ: *Fitness.* IIASA Interim Report IR-06-061 (2006).

No. 123 Brandt H, Ohtsuki H, Iwasa Y, Sigmund K: *A Survey on Indirect Reciprocity.* IIASA Interim Report IR-06-065 (2006). Takeuchi Y, Iwasa Y, Sato K (eds): Mathematics for Ecology and Environmental Sciences, Springer, Berlin Heidelberg, pp. 21-51 (2007).

No. 124 Dercole F, Loiacono D, Rinaldi S: *Synchronization in Ecological Networks: A Byproduct of Darwinian Evolution?* IIASA Interim Report IR-06-068 (2006). International Journal of Bifurcation and Chaos 7:2435-2446 (2007).

No. 125 Dercole F, Dieckmann U, Obersteiner M, Rinaldi S: *Adaptive Dynamics and Technological Change.* IIASA Interim Report IR-06-070 (2006).

No. 126 Rueffler C, van Dooren TJM, Metz JAJ: *The Evolution of Resource Specialization Through Frequency-Dependent and Frequency-Independent Mechanisms.* IIASA Interim Report IR-06-073 (2006). American Naturalist 167:81-93 (2006).

No. 127 Rueffler C, Egas M, Metz JAJ: *Evolutionary Predictions Should be Based on Individual Traits.* IIASA Interim Report IR-06-074 (2006). American Naturalist 168:148-162 (2006).

No. 128 Kamo M, Sasaki A, Boots M: *The Role of Trade-Off Shapes in the Evolution of Virulence in Spatial Host-Parasite Interactions: An Approximate Analytical Approach .* IIASA Interim Report IR-06-075 (2006).

No. 129 Boots M, Kamo M, Sasaki A: *The Implications of Spatial Structure Within Populations to the Evolution of Parasites .* IIASA Interim Report IR-06-078 (2006).

No. 130 Andreasen V, Sasaki A: *Shaping the Phylogenetic Tree of Influenza by Cross-Immunity.* IIASA Interim Report IR-06-079 (2006).

No. 131 Rueffler C, van Dooren TJM, Metz JAJ: *The Interplay Between Behavior and Morphology in the Evolutionary Dynamics of Resource Specialization.* IIASA Interim Report IR-06-082 (2006). American Naturalist 169:E34-E52 (2007).

No. 132 Rueffler C, van Dooren TJM, Metz JAJ: *The Evolution of Simple Life-Histories: Steps Towards a Classification.* IIASA Interim Report IR-06-083 (2006).

No. 133 Durinx M, Metz JAJ, Meszéna G: *Adaptive Dynamics for Physiologically Structured Population Models.* IIASA Interim Report IR-07-027 (2007).

No. 134 Ito H, Dieckmann U: *A New Mechanism for Recurrent Adaptive Radiations.* IIASA Interim Report IR-07-048 (2007). American Naturalist 170:E96-E111 (2007).

No. 135 Troost T, Kooi B, Dieckmann U: *Joint evolution of predator body size and prey-size preference.* IIASA Interim Report IR-07-050 (2007).

No. 136 Nowak MA, Sigmund K: *How Populations Cohere: Five Rules for Cooperation.* IIASA Interim Report IR-07-052 (2007). May RM, McLean A (eds): Theoretical Ecology: Principles and Applications, Oxford UP, Oxford, pp. 7-16 (2007).

No. 137 Hauert C, Traulsen A, Brandt H, Nowak MA, Sigmund K: *The Emergence of Altruistic Punishment: Via Freedom to Enforcement.* IIASA Interim Report IR-07-053 (2007). Science 613:1905-1907 (2007).

Issues of the IIASA Studies in Adaptive Dynamics series can be obtained at www.iiasa.ac.at/Research/EEP/Series.html or by writing to eep@iiasa.ac.at.

The emergence of altruistic punishment: via freedom to enforcement

Christoph Hauert [1], Arne Traulsen [1], Hannelore Brandt [2], Martin A. Nowak [1] & Karl Sigmund [3,4,*]

[1] Program for Evolutionary Dynamics, Department of Organismic and Evolutionary Biology, Department of Mathematics, Harvard University, One Brattle Square, Cambridge, MA 02138, USA

[2] Vienna University of Economics and Business Administration A-1090 Vienna, Austria

[3] Faculty of Mathematics, University of Vienna A-1090 Vienna, Austria,

[4] International Institute for Applied Systems Analysis A-2361 Laxenburg, Austria

Keywords: evolutionary game theory; public goods games; cooperation; altruistic punishment; voluntary interactions;

Word count: abstract: 185; body: 3471 (including figure legends); 3 figures.

Corresponding author:

Karl Sigmund

Faculty of Mathematics, University of Vienna, A-1090 Vienna, Austria

e-mail: karl.sigmund@univie.ac.at, phone: +43 (0)1 4277 506 12, fax: +43 (0)1 4277 9 506

In human societies, cooperative behaviour in public goods interactions is usually enforced through institutions that impose sanctions on free-riders. Many experiments on public goods games have shown that in the absence of such institutions, individuals are often willing to punish defectors, even at a cost to themselves, effectively 'taking the law into their own hands'[1–11]. Theoretical models confirm that social norms prescribing the punishment of deviant behaviour are stable: once established, they prevent invasion by dissident minorities[12–15]. But how can such costly punishing behaviour gain a foothold in the population? A surprisingly simple model shows that if individuals have the option to stand aside and abstain from the public goods interaction, this paves the way for the emergence and establishment of cooperative behaviour based on the punishment of defectors. Thus the freedom to withdraw from the public enterprise leads to a self-enforcing prosocial norm. Paradoxically, the option of individual autarky may be an important step for the emergence of institutions punishing the non-cooperation of their members. Conversely, public goods interactions which are obligatory rather than voluntary are unlikely to gain a foothold in the population.

An impressive body of evidence shows that many humans are willing to pay a personal cost in order to punish wrong-doers[1–10]. In particular, punishment is a very effective mechanism to ensure cooperation in public goods interactions. All human populations seem willing to use costly punishment to varying degree, and their willingness to punish correlates with the propensity for altruistic contributions[11]. This raises an evolutionary problem: in joint enterprises, free-riding individuals who do not contribute, but exploit the public goods, fare better than those who pay the cost of contributing. If successful behaviour spreads, these defectors will eventually take over, until there is nothing left to exploit. Punishment reduces the defector's payoff, and thus may solve the social dilemma. But since punishment is costly, it also reduces the punishers' payoff. This raises what has been called a 'second order social dilemma'. Costly punishing is an altruistic act. Individuals who contribute, but do not punish, are better off than the punishers. The frequency of punishers will dwindle and the defectors return.

This second order defection can be punished in turn, and prevented to spread. Thus a free-riding minority will be seriously harmed in the case of altruistic punishment (while it will be pampered in the case of altruistic cooperation). Any social norm that includes the rule to punish those who deviate is evolutionarily stable: once established, it cannot be displaced by an invading minority of dissidents[12]. But how can such punishing behaviour gain a foothold in the population? The trait has to be rare, initially, and thus will incur huge costs by ceaselessly punishing. The emergence of altruistic punishing behaviour is acknowledged to be a major puzzle in the evolution of cooperation. "We seem to have replaced the problem of explaining cooperation with that of explaining altruistic punishment"[16].

We will show that the puzzle disappears if one assumes that individuals can voluntarily decide whether to take part in the joint enterprise or not. If they do not participate, they can obtain an autarkic payoff independent of the other players' behaviour. Thus we consider four strategies. The *loners* are those who do not participate in the public enterprise (they need not be solitary individuals: they just abstain from the public goods game). Those who participate include the *defectors*, who do not contribute but exploit the contributions of the others; the *cooperators*, who contribute, but do not punish; and the *punishers*, who not only contribute to the public good, but punish the defectors, and possibly also those who fail to punish the defectors. In such a model, punishers will invade and take over. In the absence of the loner's option, however, they will often not be able to invade, and the population will be dominated by defectors. This means that if participation in the joint enterprise is voluntary, cooperation-enforcing behaviour emerges. If participation is obligatory (i.e. loners are excluded), then the defectors will win.

This intriguing result was originally presented by Fowler[17]. But his argument was based on a model with serious shortcomings, which does not justify the conclusions[18]. Here we propose a model vindicating Fowler's intuition. If individual autarky is an option, social norms enforcing cooperation through punishment will emerge and come close to fixation, whereas obligatory participation in the public goods game leads to take-over by defectors. These theoretical findings agree well with the results of recent experiments, and may offer a solution to one of the most persistent problems in the evolution of cooperation.

We consider a well-mixed population of constant size M. From time to time, a random sample of size N is selected and offered the option to participate in a public goods game. Those who agree to do so can decide whether or not to contribute an investment of value c to themselves. The individual contributions are added up and multiplied with a factor $r > 1$. The resulting sum is then divided equally among all participants of the public goods game. After this interaction, each contributor can impose a fine β upon each defector, at a personal cost γ for each fine. Moreover, those who punish can also impose a fine on those who contributed, but failed to punish non-contributors. We shall assume for simplicity that for this second type of punishment, fines and costs are reduced by a factor α, with $0 \leq \alpha \leq 1$. By x we denote the total number of cooperators (who contribute in the public goods game but do not punish), by y that of defectors (who participate in the public goods game but do not contribute), by z that of loners (unwilling to participate) and by w the number of punishers (who contribute, and punish the defectors as well as the cooperators who did not punish the defectors in their group). Thus $M = x + y + z + w$. We do not consider more complex strategies (basing their decision, for instance, on the size and composition of the group, or on past experience).

Each loner receives a constant payoff σ. Among the random sample of size N, there will be N_x cooperators, N_y defectors, N_z loners and N_w punishers. These are random variables distributed according to a multivariate distribution which describes sampling without replacement. The group of those willing to participate in the public goods game has size $S := N_x + N_y + N_w$. If $S \leq 1$ then the public goods game does not take place. A player who volunteered for it receives the loner's payoff σ. If $S > 1$, each participant of the public goods game obtains an income $r(N_x + N_w)c/S$. The payoff for the contributors (i.e. the cooperators and the punishers) is reduced by c. The payoff for the defectors is reduced by βN_w. The payoff for the cooperators is reduced by $\alpha\beta N_w$, provided $N_y > 0$ (if there are no defectors in the group, non-punishing behavior will go unnoticed). The payoff for punishers is reduced by γN_y and, if $N_y > 0$, by

$\alpha\gamma N_x$.

This concludes the description of the strategic interaction[18]. We next specify how strategies are transmitted within the population[19]. We define each players' fitness as $1-s+sP$, the convex combination of the 'baseline fitness', which is normalised to 1 for all players, and the payoff P from the optional public goods game with punishment. The relative importance of each component is determined by the selection strength s: small s means weak selection. We shall assume that occasionally, a randomly chosen player can change strategy by adopting the strategy of a player picked with a probability proportional to that player's fitness. This mimics a learning process similar to the Moran process describing natural selection: more successful players are copied more frequently. In addition, we shall assume that with a small probability μ, a player can switch to another strategy irrespective of its payoff (this "mutation term" corresponds to blindly experimenting with anything different).

The analysis of the corresponding stochastic dynamics is greatly simplified in the limiting case $\mu \to 0$. The population consists almost always of one or two types at the most. This holds because for $\mu = 0$ the four monomorphic states are absorbing, and for very small μ the fate of a mutant (i.e. its elimination or fixation) is settled before the next mutant appears. The probabilities ρ_{ij} that a single i-player in a population of j-players reaches fixation can be calculated for $i, j \in \{x, y, z, w\}$ (see online supporting material). This defines a transition matrix among the four monomorphic states of the system, and hence a unique stationary distribution. For small mutation rates μ, this distribution specifies how likely the system is to be in the corresponding pure state, or in its vicinity. Computer simulations show that the approximation also holds for larger mutation rates (on the order of $1/M$).

The outcome is striking: in the limit of rare mutations, the system is most of the time in the homogeneous state with punishers only, irrespective of the initial composition of the population. For large populations ($M = 1000$ can be considered large for most of our prehistory) and small mutation rates, the system spends more than 80 percent in or near the punisher state. This prevalence diminishes only for very small selection strengths (Fig. 1).

In the case of an obligatory public goods game, i.e. in the absence of the loner's option, the situation is very different: in the limit of rare mutations, the system spends most of the time in or near the state with defectors only. For the same parameter values as before, the state is for 90 percent of the time dominated by defectors, and there is hardly any economic benefit from the public good (Fig. 2a). If all contributors punish (i.e. cooperators are excluded), the result remains essentially the same: even if defectors prevail for obligatory public goods games, they are eliminated if the public goods game is voluntary.

Volunteering in the absence of the punishment leads to a more cooperative outcome than for the obligatory game but not to the fixation of the cooperative state. The system exhibits a strong tendency to cycle (from cooperators to defectors to loners and back to cooperators). Roughly speaking, almost half of the time the state is dominated by loners. An outcome dominated by cooperators is almost as likely, whereas domination by defectors is relatively rare. In the limiting case of weak selection, the population even cooperates most of the time (Fig. 2b).

Whereas the limiting case of small mutation rates can be studied analytically (c.f. Figs. 1, 2, and the supplementary information), the case of substantial mutation rates can only be handled by numerical simulations. Complementing interactive online tutorials are provided at *http://homepage.univie.ac.at/hannelore.brandt/publicgoods/* and the VirtualLabs at *http://www.univie.ac.at/virtuallabs*. These show that the outcome is robust within a wide range of parameter values. With cooperators, loners and defectors only, the latter do worst, whereas the former two perform comparably well. With cooperators, punishers and defectors, but no loners, punishers do not prevail, except for large mutation rates. In that case, the mutational drift supplying defectors keeps the punishers active and prevents them from being undermined by cooperators. If all four types are admitted, punishers prevail.

In an obligatory public goods game with cooperators and defectors only, the latter obviously win. The loner's option allows cooperators to persist (although they cannot dominate for a substantial period). The reason is a simple rock-paper-scissors mechanism[20–22]. If there are many defectors, loners will spread. When loners abound, many of the random samples will result in only small groups of players willing to participate. If the groups are sufficiently small, the average payoff for cooperators will be larger than that for defectors, despite the fact that within each group, the latter have a higher payoff than the former; it even pays for the individual defector to switch to cooperation. Thus if the group is sufficiently small ($S < r$), there is no social dilemma. This is a fleeting state only: quickly, cooperators spread, group size increases and the social dilemma returns. But the recurrent eclipse of the social dilemma allows punishers eventually to step in and take over (see Fig. 3).

We have assumed in our model that a punisher faced with twice as many defectors metes out twice as many fines. This assumption can be modified without affecting the conclusions. As it stands, it makes the life of a rare punisher particularly difficult. It is all the more remarkable that punishers can invade nevertheless.

Whether cooperators who fail to punish are punished or not plays a surprisingly small role. The parameter α has little influence on the numerical simulations, and does not show up in the formulas (see supplementary information). The reason is that in the limiting case (μ very small), the three types of punishers, cooperators and defectors rarely co-exist: hence punishers cannot hold cooperators to account for not punishing defectors. In the case $\alpha = 0$, the second order social dilemma always holds: punishing is costly, and contributors failing to punish can get away with it. Nevertheless, punishing behaviour can emerge and prevail, because the first order social dilemma occasionally breaks down. It is of interest in this context that experimental evidence for the punishment of non-punishers (i.e. for non-vanishing α) seems to be lacking[13].

For weak selection, an analytical condition for the dominance of punishers in the absence of loners can be derived: $3(N - r) < N(N - 1)(\beta - \gamma)$ (see supplementary information). This condition is satisfied in Fig. 2a and is reflected in the dominance of punishers for small s. Moreover, if defectors are allowed to retaliate (in which case β is as large as γ) punishers never dominate the population and loners are needed to establish cooperation. However, also note that for strong selection it is clear that defectors always dominate because selection acts against invasion attempts of cooperators as well as punishers (see Fig. 3 and supplementary information).

We could also assume that punishers penalise non-participants. The fine could be $\delta\beta$ and the cost to the punisher $\delta\gamma$, with $0 \leq \delta \leq 1$. Again, this has no great effect on the outcome. If loners are frequent, many samples will contain no punisher. If punishers are frequent, defectors are kept in check and non-participants do poorly, with or without being punished. The most significant difference seems to be that if punishers pay a heavy cost for penalising loners (high δ and γ), then cooperators are needed to overcome the dominance of loners and catalyse the take-over by punishers.

Differences between first and second order social dilemma have been pointed out before in a model[14] based on a group selection scenario and exploiting the fact that when punishers are common, individual level selection against them is weak and may be overcome by selection among groups. Several other models confirm that the punishment of defectors is stable, if it is the prevalent norm. For example by assuming some degree of conformism in the population[15]: individuals preferentially copy what is frequent. Similarly, cooperation can also be stabilised through indirect reciprocity[23], but in each of these cases, the emergence of the pro-social norm remains unclear[24,25].

Our model, in contrast, shows that even when initially rare, punishing behaviour can be selectively advantageous, and is likely to become fixed. We consider the most challenging scenario, namely a single well-mixed population whose members imitate preferentially what fares better, not what is more common. The effects of group selection and conformist transmission will further the maintenance of this pro-social norm, once it is established.

The spread of initially rare punishers is also the outcome in Ref. 14. But that model, based on an infinitely large population, assumes that single cooperators can play the public goods game, and obtain a payoff which is higher than that of loners (as high, in fact, as if the whole population contributes to the public good). This neglects the fact that contributing to a public goods game is a risky investment whose return depends on what other players are doing. By contrast, our model leads, in the limiting case of an infinitely large population, to a bistable outcome[18]. Depending on the initial condition, the state either ends up in a Nash equilibrium consisting of cooperators and punishers only, or leads to endless oscillations of loners, defectors and cooperators, without any punishers. Bistability also holds for the reputation-based model [26]. Both approaches do not favour the spread of a minority of punishers. Their emergence is boosted, in finite populations, by stochastic effects. Voluntary participation, by reducing group size if defectors abound, promotes these stochastic effects.

Recent experiments show that if players can choose between joining a public goods game either with or without punishment, they prefer the former[27]. The interpretation seems clear: whoever freely accepts that defection is punished is unlikely to be a defector. It is thus less risky to join such a group. Players voluntarily commit themselves to sanctioning rules. This voluntary submission to a sanctioning regime is not always immediate, however: in the majority of cases it requires a few preliminary rounds. Many players appear to have initial reservations against a sanctioning regime and need a learning phase. In another series of experiments, it has been shown that threatened punishment can decrease the level of cooperation in trust games[28]. Moreover, players reduce their punishing behaviour if they have a less costly option (such as excluding defectors from indirect reciprocity networks), but they do not give it up: rather, they punish in a more focussed way[29]. Experimental evidence for altruistic punishment can also be found in the ultimatum game (rejecting an unfair offer is costly to both players)[2] and in indirect reciprocity (by not helping defectors, players reduce their own chances of being helped)[30].

Reports from present-day hunter-gatherer societies often stress their egalitarian and 'democratic' features: individuals have a great deal of freedom[31]. This creates favourable conditions for voluntary participation. Opting for the 'loner' strategy does not mean living an eremit's life; it means not participating in a collective hunt, for instance, but collecting mushrooms instead. On the other hand, ostracism was probably an early form of severe punishment. There seems to be a smooth transition between choosing not to take part in a joint enterprise and being excluded from it. Together, these two alternatives may explain the emergence of rule-enforcing institutions promoting pro-social behaviour - following Hardins recipe for overcoming the *tragedy of* *the commons*: mutual coercion, mutually agreed upon[32]. However, we must emphasise that there are public good games where no one can stand aside: the preservation of our climate is one example[33]. In such games, participation is obligatory – and defection widespread.

References

1. Fehr, E. and Gächter, S. Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).
2. Fehr, E. and Fischbacher, U. The nature of human alruism. *Nature* **425**, 785–791 (2003).
3. Hammerstein, P., editor. *Genetic and cultural evolution of cooperation*. MIT Press, Cambridge MA, (2003).
4. Price, M. E., Cosmides, L., and Tooby, J. Punitive sentiment as an anti-free rider psychological device. *Evol. Hum. Behav.* **23**, 203–231 (2002).
5. Masclet, D., Noussair, C., Tucker, S., and Villeval, M. C. Monetary and non-monetary punishment in the voluntary contributions mechanism. *Am. Econ. Rev.* **93**, 366–380 (2003).
6. Shinada, M., Yamagishi, T., and Ohmura, Y. False friends are worse than bitter enemies: ”altruistic” punishment of in-group members. *Evol. Hum. Behav.* **25**, 379–393 (2004).
7. Gintis, H., Bowles, S., Boyd, R., and Fehr, E., editors. *Moral sentiments and material interests: the foundations of cooperation in economic life*. MIT Press, Cambridge MA, (2005).
8. de Quervain, D. J.-F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., and Fehr, E. The neural basis of altruistic punishment. *Science* **305**, 1254–1258 (2004).
9. Camerer, C. and Fehr, E. When does ”economic man” dominate social behaviour? *Science* **311**, 47–52 (2006).

10. Nakamaru, M. and Iwasa, Y. The evolution of altruism and punishment: role of the selfish punisher. *J. theor. Biol.* , (to appear) (2006).

11. Henrich, J. Costly punishment across human societies. *Science* **312**, 176–177 (2006).

12. Boyd, R. and Richerson, P. J. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology* **13**, 171–195 (1992).

13. Kiyonari, T. and Barclay, P. Second-order punishment and reward in public goods games. , (to appear) (2006).

14. Boyd, R., Gintis, H., Bowles, S., and Richerson, P. The evolution of altruistic punishment. *Proc. Natl. Acad. Sci. USA* **100**, 3531–3535 (2003).

15. Henrich, J. and Boyd, R. Why people punish defectors. *J. theor. Biol.* **208**, 79–89 (2001).

16. Colman, A. The puzzle of cooperation. *Nature* **440**, 744–745 (2006).

17. Fowler, J. H. Altruistic punishment and the origin of cooperation. *Proc. Nat. Acad. Sci. USA* **102**(19), 7047–7049 (2005).

18. Brandt, H., Hauert, C., and Sigmund, K. Punishing and abstaining for public goods. *Proc. Natl. Acad. Sci. USA* **103**(2), 495–497 (2006).

19. Nowak, M. A., Sasaki, A., Taylor, C., and Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650 (2004).

20. Hauert, C., De Monte, S., Hofbauer, J., and Sigmund, K. Volunteering as red queen mechanism for cooperation in public goods games. *Science* **296**, 1129–1132 (2002).

21. Semmann, D., Krambeck, H.-J., and Milinski, M. Volunteering leads to rock-paper-scissors dynamics in a public goods game. *Nature* **425**, 390–393 (2003).

22. Hauert, C., De Monte, S., Hofbauer, J., and Sigmund, K. Replicator dynamics in optional public goods games. *J. theor. Biol* **218**, 187–194 (2002).

23. Panchanathan, K. and Boyd, R. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* **432**, 499–502 (2004).

24. Fowler, J. H. Human cooperation: Second-order free-riding problem solved? *Nature* **437**, E8 (2005).

25. Milinski, M., Semmann, D., and Krambeck, H.-J. Reputation helps solve the 'tragedy of the commons'. *Nature* **415**, 424–426 (2002).

26. Sigmund, K., Hauert, C., and Nowak, M. A. Reward and punishment. *Proc. Natl. Acad. Sci.* **98**, 10757–10762 (2001).

27. Gürerk, O., Irlenbush, B., and Rockenbach, B. The competitve advantage of sanctioning institutions. *Science* **312**, 108–111 (2006).

28. Fehr, E. and Rockenbach, B. Detrimental effects of sanctions on human altruism. *Nature* **422**, 137–140 (2003).

29. Rockenbach, B. and Milinski, M. Indirect reciprocity resolves the efficiency dilemma of costly punishment. *Nature* , (to appear) (2006).

30. Wedekind, C. and Milinski, M. Cooperation through image scoring in humans. *Science* **288**, 850–852 (2000).

31. Johnson, A. W. and Earle, T. *The Evolution of Human Societies: from Foraging Group to Agrarian State*. Stanford UP, Stanford CA, (1987).

32. Hardin, G. The tragedy of the commons. *Science* **162**, 1243–1248 (1968).

33. Milinski, M., Semmann, D., Krambeck, H.-J., and Marotzke, M. Stabilizing the earths climate is not a losing game: Supporting evidence from public goods experiments. *Proc. Natl. Acad. Sci. USA* **103**, 3994–3998 (2006).

Figures

Hauert et al. - Figure 1

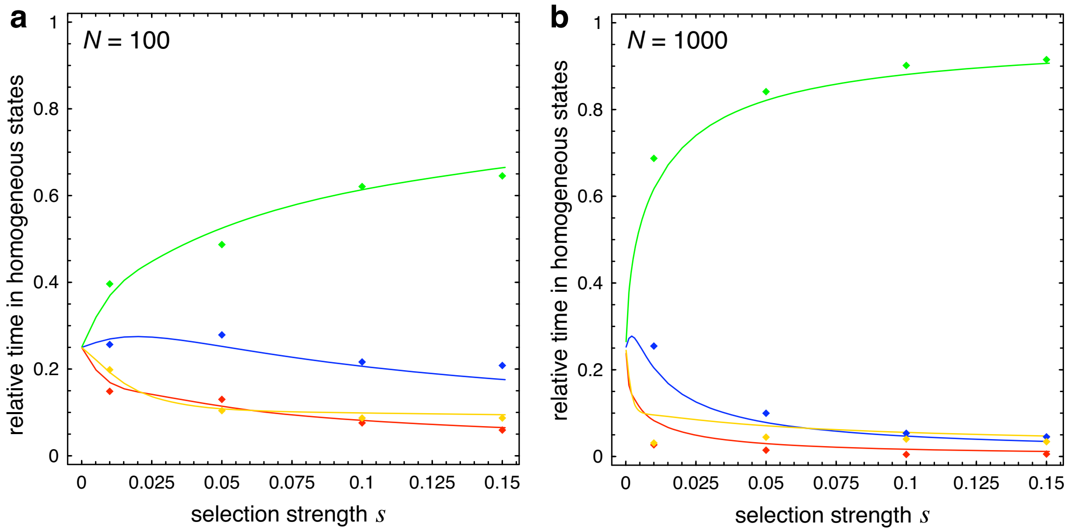

Figure 1: Punishment and abstaining in public goods games in finite populations. In the limit of rare mutations ($\mu \to 0$), the dynamics is restricted to transitions between the four homogeneous states with all cooperators (blue), defectors (red), loners (yellow) or punishers (green). The two panels depict the probabilities of each state as a function of the selection strength s for population sizes $M = 100$ (**a**) and $M = 1000$ (**b**). Simulation data for small mutation rates confirms the analytical results (colored dots). In contrast to the analysis, the fitness of individuals in the simulations is determined by a single random interaction rather than the average. This is source of stochasticity and, together with the mutation rate, is responsible for the small differences between analytical results and the simulation data. Parameters: $N = 5, r = 3, \sigma = 1, \gamma = 1, \beta = 2, \alpha = 0.1, s_{\text{max}} = 0.151$; Simulations: **a** mutation rate $\mu = 10^{-4}$, sampling time $T = 10^7$, **b** $\mu = 10^{-3}$, $T = 10^6$.

Hauert et al. - Figure 2

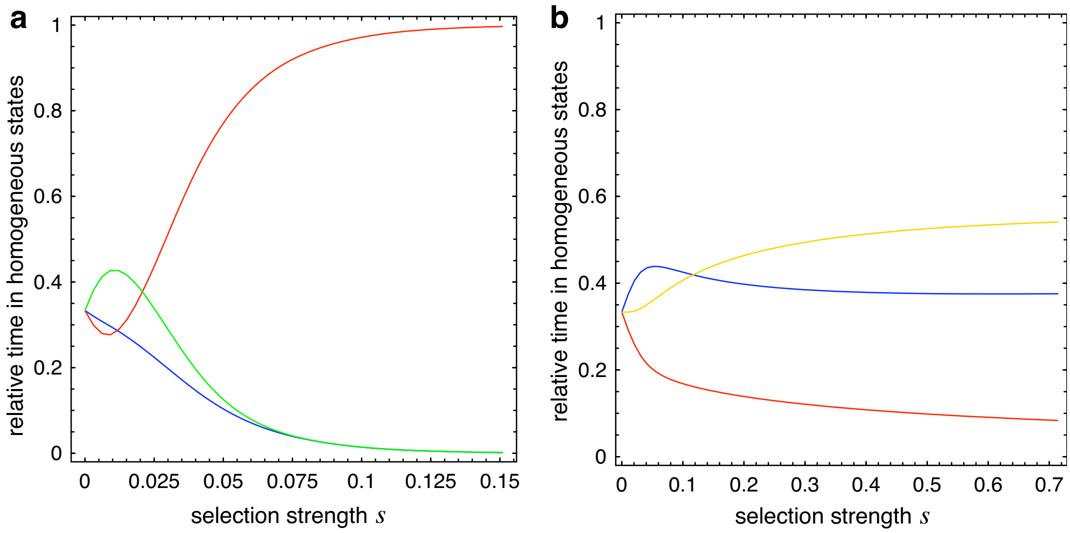

Figure 2: Punishment in obligatory public goods games (**a**) and voluntary participation in public goods games without punishment (**b**). In the limit of rare mutations, the dynamics reduces to transitions between homogeneous states with all cooperators (blue), defectors (red), punishers (green) or loners (yellow). The probabilities for each state are shown as a function of the selection strength s. In the limit of neutral evolution ($s = 0$) the strategic differences disappear and all three respective states become equally likely. In **a** the system is usually found in a state with all defectors, except for weak selection where punishers manage to get the upper hand. In contrast, in **b**, the system spends significantly more time in the cooperator or loner states than in the defector state. Parameters: $N = 5, r = 3, \sigma = 1, \gamma = 1, \beta = 2, \alpha = 0.1, M = 100$; **a** $s_{\text{max}} = 0.151$; **b** $s_{\text{max}} = 0.714$.

Hauert et al. - Figure 3

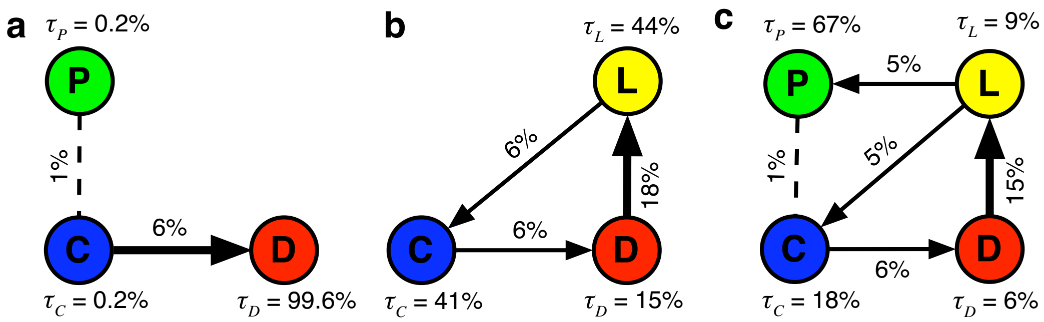

Figure 3: Schematic dynamics for **a** obligatory public goods games with punishment, **b** voluntary public goods games, and **c** the combined effects of volunteering and punishment in finite populations. In the limit of rare mutations, the dynamics is reduced to transition probabilities between the homogeneous states with all cooperators (C), defectors (D), loners (L) or punishers (P). The three panels depict all transition probabilities $> 0.01\%$ together with the relative time τ spent in each state for maximal selection strength. **a** In the absence of loners, defectors dominate despite punishment. The cooperator and punisher state are connected by a neutral edge with a transition probability of $1/M$ in either direction (dashed line). The defector state is essentially stable with transition probabilities to the cooperators state of $< 10^{-4}$ and still many orders of magnitude smaller to reach the punisher state. **b** In voluntary public goods games the transition probabilities illustrate the cyclic dominance of the three strategies and illustrate that the system spends little time in the defector state because of the large transition probability D $\rightarrow$ L. **c** Combining the two mechanisms illustrates the pivotal role of loners where the system can embark on another cooperator-defector-loner cycle or switch to the punisher state. Parameters: $N = 5, r = 3, \sigma = 1, \gamma = 1, \beta = 2, M = 100, s = 0.151.$