

Justus-Liebig-Universität Gießen
Fachbereich Biologie und Chemie

„Eine evolutionäre Analyse individueller strategischer
Entscheidungen in Kooperationssituationen“

Dissertation zur Erlangung des Doktorgrades der
Naturwissenschaftlichen Fachbereiche der
Justus-Liebig-Universität Gießen

vorgelegt von

Dr. Ulrich Frey
aus Nakuru (Kenia)

Stuttgart, September 2015

betreut durch Prof. Dr. Eckart Voland und
Prof. Dr. Volker Wissemann

Für Sonja, Laura, Elinor und Felix

Danksagung

Ich möchte mich zunächst bei meinem Doktorvater Eckart Voland für seine Unterstützung in allen Belangen herzlich bedanken! Ganz besonders habe ich seine Art schätzen gelernt, immer offen gegenüber neuen Ideen zu sein, dabei Disziplingrenzen zu vergessen, weil faszinierende Forschungsfragen dazu tendieren, sich nicht daran zu halten. Seine hilfreichen Diskussionen, die immer wieder aufs Wesentliche zurückführen, haben mir sehr geholfen.

Ein weiteres herzliches Dankeschön geht an Volker Wissemann für seine Bereitschaft, diese Arbeit zu begutachten, ebenso wie an Volkmar Wolters und Dittmar Graf. Vielen Dank auch an Frau Quillfeldt – ihre zahlreichen Anmerkungen haben zur Qualität der Arbeit beigetragen.

Diese Dissertation wäre nicht möglich gewesen, wenn mir Gameforge, und im besonderen Sebastian Mayer nicht ihre Daten zur Verfügung gestellt hätten. Das gilt ebenso für die im Internet verfügbaren Daten zu den Labor-PGG: Danke an James Walker, Martin Sefton and Robert Schupp; Ernst Fehr; Nikos Nikiforakis und Hans-Theo Normann; Simon Gächter and Elke Renner.

Bei der Analyse der Daten, vor allem der Umformung, waren mir besonders Niklas Rappel (zu meinen SQL-Fragen), Carl Salk (zu meinen R-Fragen) und Manfred Hollenhorst (zu meinen statistischen Fragen) behilflich. Es hat Spaß gemacht in einem so angenehmen Umfeld an der Universität Gießen arbeiten zu können – danke an Charlotte, Kai W., Hannah, Kai D., Johannes und Ingrid Weil. Insa danke ich für ihre Geduld.

Schließlich möchte ich mich für die Unterstützung durch mein persönliches Umfeld bedanken: Hier waren meine Eltern sowie Maria und Hans immer für die Kinder da, wenn ich etwas Zeit brauchte!

Ein besonders dickes Dankeschön geht an meinen Bruder Johannes, der immer zur Stelle war, wenn es nötig war! Dank seiner beständigen Hilfe lesen sich meine Texte viel runder.

Schließlich möchte ich mich bei meiner Familie – Sonja, Felix, Elinor und Laura – bedanken, die es immer geschafft haben, meine vielfältigen Frustrationen über Probleme bei der Arbeit mit den Daten in gute Laune umzuwandeln!

Inhaltsverzeichnis

A	Einleitung.....	12
1	Theoretischer Hintergrund: Evolutionäre Mechanismen zur Erklärung von Kooperation.....	12
2	Soziale Dilemmasituationen.....	20
3	Fazit für die Methodik.....	27
4	Ziel und Fragestellung.....	32
B	Stand der Forschung und Arbeitshypothesen.....	35
1	Allgemeine Arbeitshypothesen.....	35
2	Stand der Forschung zur Studie 1 (Spielertypen): Kooperationsverhalten in public goods games	42
3	Arbeitshypothesen zur Studie 1 (Spielertypen).....	45
4	Stand der Forschung zur Studie 2 (Effiziente Bestrafung) – Sanktionen in public goods games, in sozial-ökologischen Systemen und bei Jägern und Sammlern.....	47
4.1	Sanktionen als Mittel, um Kooperation zu erhöhen.....	50
4.2	Effizienz von Sanktionen.....	51
4.3	Sanktionen in sozial-ökologischen Systemen.....	51
4.4	Sanktionen bei Jägern und Sammlern.....	52
5	Arbeitshypothesen zur Studie 2 (Effiziente Bestrafung).....	53
6	Stand der Forschung zur Studie 3 (Erfolg Kooperierender) – Die evolutionäre Funktionslogik sozialer Dilemmata.....	55
7	Arbeitshypothesen zur Studie 3 (Erfolg Kooperierender).....	58
C	Daten und Methoden.....	62
1	Beschreibung der Daten aus PGG-Laborexperimenten.....	62
2	Beschreibung der Daten des Online-Browser-Spiels Ikariam.....	67
2.1	Definition der Spielertypen.....	77
2.2	Definition der abhängigen Variablen (Erfolgsmaß und Investitionsbereitschaft).....	80
3	Beschreibung der Arbeitsschritte zur Datentransformation.....	82
3.1	PGG-Laborexperimente.....	82
3.2	Ikariam.....	84
4	Methoden.....	85
D	Ergebnisse und Diskussion.....	87
1	Ergebnisse und Diskussion Studie 1: Differenzierung von Spielertypen.....	87
1.1	Resultate für Hypothesen 1 und 2: Differenzierung von Spielertypen.....	87
1.2	Diskussion Hypothesen 1 und 2.....	100
1.2.1	Interpretation der Häufigkeiten von Spielertypen.....	100
1.2.2	Evolutionäre Interpretation der Spielertypen.....	103
2	Ergebnisse und Diskussion Studie 2: Effiziente Bestrafung.....	104
2.1	Resultate für Hypothese 1: Kooperative Spielertypen strafen häufiger.....	104
2.2	Diskussion Hypothese 1.....	110
2.3	Resultate für Hypothese 2: Spielertypen und Straftypen treten in bestimmten Kombinationen auf.....	111
2.4	Diskussion Hypothese 2.....	114
2.5	Resultate für Hypothese 3: Gruppenzusammensetzung beeinflusst das Strafverhalten... ..	114
2.6	Diskussion Hypothese 3.....	116
2.7	Resultate für Hypothese 4: Sanktionen sind in ihrem zeitlichen Verlauf unabhängig vom Spielertyp und treten vor allem zu Beginn von Interaktionen auf.....	117
2.8	Diskussion Hypothese 4.....	118
2.9	Resultate für Hypothese 5: Die Kombination aus Sanktionen und hohen Einzahlungen in ein Gemeingut führt zu Wettbewerbsvorteilen für Kooperierende.....	121
2.10	Diskussion Hypothese 5.....	122

3	Ergebnisse Studie 3: Der Erfolg Kooperierender.....	123
3.1	Resultate PGG für Hypothese 1: Egoistischere Strategien sind – ohne zusätzliche Randbedingungen – kooperierenden Strategien überlegen.....	123
3.2	Diskussion H1 PGG.....	126
3.3	Resultate Ikariam für Hypothese 1: Egoistischere Strategien sind – ohne zusätzliche Randbedingungen – kooperierenden Strategien überlegen.....	126
3.4	Diskussion H1 Ikariam.....	129
3.5	Resultate PGG Hypothese 2: Wenn Kooperierende unter sich sind, so sind sie Trittbrettfahrergruppen überlegen.....	131
3.6	Diskussion H2 PGG.....	137
3.7	Resultate Ikariam Hypothese 2: Wenn Kooperierende unter sich sind, so sind sie Trittbrettfahrergruppen überlegen.....	137
3.8	Diskussion H2 Ikariam.....	142
3.9	Resultate PGG Hypothese 3: Die Staffelung des Gruppenerfolgs sollte K, KR, R, KRT, KT, RT, T sein.....	143
3.10	Resultate Ikariam H3: Die Staffelung des Gruppenerfolgs sollte K, KR, R, KRT, KT, RT, T sein.....	144
3.11	Diskussion H3 Ikariam.....	146
3.12	Resultate PGG Hypothese 4: Im direkten Aufeinandertreffen von Kooperierenden und Trittbrettfahrern ist Erfolg (für T) bzw. Misserfolg (für K) am deutlichsten.....	146
3.13	Diskussion H4 PGG.....	147
3.14	Resultate Ikariam Hypothese 4: Im direkten Aufeinandertreffen von Kooperierenden und Trittbrettfahrern ist Erfolg (für T) bzw. Misserfolg (für K) am deutlichsten.....	148
3.15	Diskussion H4 Ikariam.....	149
3.16	Resultate PGG Hypothese 5: Alle Spielertypen interagieren länger mit Kooperierenden als mit Trittbrettfahrern.....	150
3.17	Resultate Ikariam Hypothese 5: Alle Spielertypen interagieren länger mit Kooperierenden als mit Trittbrettfahrern.....	150
3.18	Diskussion H5 Ikariam.....	152
3.19	Resultate PGG Hypothese 6: Je länger die Interaktionen, desto größer die Unterschiede zwischen Kooperierenden und Trittbrettfahrern.....	152
3.20	Diskussion H6 PGG.....	156
3.21	Resultate Ikariam H6: Je länger die Interaktionen, desto größer die Unterschiede zwischen Kooperierenden und Trittbrettfahrern.....	156
3.22	Diskussion H6 Ikariam.....	158
3.23	Resultate PGG Hypothese 7: Die R1-Strategie mit Trade-off zwischen Investitionen in das öffentliche Gut und persönlichem Vorteil sollte sehr guten Erfolg zeigen.....	159
3.24	Resultate Ikariam Hypothese 7: Die R1-Strategie mit Trade-off zwischen Investitionen in das öffentliche Gut und persönlichem Vorteil sollte sehr guten Erfolg zeigen.....	160
3.25	Diskussion H7 Ikariam.....	163
E	Gesamtfazit und allgemeine Diskussion.....	163
1	Zusammenhang der untersuchten drei Kooperationsstudien.....	163
2	Neue Erkenntnisse.....	164
3	Offene Fragen.....	172
F	Literaturverzeichnis.....	173
G	Anhang.....	186
1	SQL-Skripte.....	186
1.1	Skript 1.....	186
1.2	Skript 2.....	187
2	Internetspeicherorte der Rohdaten der verwendeten Laborexperimente.....	187

3 Ikariam-Umfrage.....	188
4 Berechnung des Step return.....	188
5 Beschreibung der Variablen bei Ikariam.....	191
6 Ausbaustufen der drei Gemeingüter.....	193
7 Weitere Abbildungen.....	197
H Eigenständigkeitserklärung.....	199

Tabellenverzeichnis

Tabelle 1. Evolutionäre Erklärungsmodelle für die Gemeinguterzeugung.....	41
Tabelle 2. Vergleich der wichtigsten Parameter der fünf verwendeten PGG-Studien.....	63
Tabelle 3. Definition der Spielertypen.....	64
Tabelle 4. Definition der Gruppentypen.....	65
Tabelle 5: Definition der Straf- und Bestraftypen.....	66
Tabelle 6. Deskriptive Statistik der Ikariam-Spielerzahlen im Ländervergleich.....	72
Tabelle 7. Altersverteilung bei deutschen Ikariam-Spielern (Daten aus eigener Online-Umfrage 2014, siehe Anhang G).....	73
Tabelle 8. Deskriptive Statistik für den Ausbau der öffentlichen Güter.....	73
Tabelle 9: Korrelationen zwischen den Ausbaustufen der Gemeingüter nach zehn Monaten Spielzeit (Daten des letzten Snapshots, am 13.2.2014).....	77
Tabelle 10. Vergleich der Anteile an Spielertypen in Prozent aus 8 Studien zu konditionaler Kooperation (Messmethoden sind teilweise unterschiedlich, Typen sind für diesen Überblick teilweise zusammengefasst).....	88
Tabelle 11. Mittelwert Anzahl Städte pro Inseltyp im Ländervergleich ohne den ersten Snapshot (25.4) für Inseln mit mindestens zwei Spielern.....	90
Tabelle 12. Vergleich des Anteils an Gruppentypen für Inseln mit und ohne Einzelspieler ohne den ersten Snapshot vom 25.4.....	92
Tabelle 13. Häufigkeiten von Reziprokatoren in Prozent mit verschiedenen Messmethoden (siehe Methoden).....	93
Tabelle 14. Häufigkeiten von Trittbrettfahrern und Kooperierenden in Prozent mit verschiedenen Messmethoden (siehe Methoden).....	93
Tabelle 15. Mittelwerte über alle 28 Snapshots der Anteile von Spielertypen in Prozent im Ländervergleich.....	99
Tabelle 16: Häufigkeitsanteile der Spielertypen in Prozent im Ländervergleich nach zehn Monaten.....	99
Tabelle 17: Paarweise Vergleiche der Post-hoc-Tests für Unterschiede in der Bestrafung zwischen Spielertypen in PGG.....	109
Tabelle 18: Spearman-Korrelationen (zweiseitig) zwischen Spielertyp, Strafen und erhaltener Bestrafung für drei Studien (n = 112; ** = auf dem 0,01-Niveau signifikant).....	110
Tabelle 19: Pearson-Korrelationen (zweiseitig) zwischen Einzahlungen und Verdienst für drei Studien (** = auf dem 0,01-Niveau signifikant).....	121
Tabelle 20: Paarweise Vergleiche der Post-hoc-Tests für Unterschiede im Erfolgsmaß zwischen Spielertypen in Ikariam.....	129
Tabelle 21. Häufigkeit der Spielertypen bei geänderter Definition (siehe Text).....	130
Tabelle 22. Durchschnittlicher Erfolg (Rathauslevel / Spieldauer) pro Anzahl der Städte nach Inseltyp im Ländervergleich nach zehn Monaten.....	141
Tabelle 23: Paarweise Vergleiche der Post-hoc-Tests für Unterschiede im Erfolgsmaß zwischen Gruppen in Ikariam.....	144
Tabelle 24. Pearson-Korrelationskoeffizienten (einseitig) zwischen der Anzahl der Städte und den Ausbaustufen der Gemeingüter für den letzten Snapshot (13.2.2014) ohne leere Inseln; alle Zusammenhänge sind auf dem 0,01-Niveau signifikant.....	146
Tabelle 25: Deskriptive Statistik für Spieldauer pro Spielertyp.....	150
Tabelle 26: Deskriptive Statistik für den Erfolg (RHL/SD) pro Stadtanzahl pro Spielertyp nach 10 Monaten.....	162
Tabelle 27. Zusammenfassung der Ergebnisse zu den Arbeitshypothesen der drei Studien.....	165
Tabelle 28. Beispielberechnung des step return für acht Spieler bei einem durchschnittlichen Bonus	

von 20%.....	191
Tabelle 29. Information über die Spieler (5 Variablen).....	191
Tabelle 30. Information über die Städte und deren Gebäude (49 Variablen).....	192
Tabelle 31. Information über Einzahlungen in die Gemeingüter (4 Variablen).....	193
Tabelle 32. Information über die Inseln (11 Variablen).....	193
Tabelle 33. Stufen des Gemeinguts Sägewerk.....	195
Tabelle 34. Stufen des Gemeinguts Luxusmine.....	197
Tabelle 35. Stufen des Gemeinguts Wunder.....	197

Abbildungsverzeichnis

Abbildung 1. Eine Typologie evolutionärer Erklärungen für Kooperation (West et al. 2011, S. 236, übersetzt).....	15
Abbildung 2: Unterschiede in der Organspendebereitschaft je nach opt-in oder opt-out Verfahren.....	25
Abbildung 3: Durchschnittliches Angebot im Ultimatumspiel in 15 Kulturen in Prozent (Henrich et al. 2010).....	29
Abbildung 4. Mittlere Höhe der Bestrafung in PGG in 16 Städten (Herrmann et al. 2008).....	49
Abbildung 5. Screenshot der Hauptbestandteile der Gemeingutsituation in Ikariam.....	69
Abbildung 6: Entwicklung der Spielerzahlen im Zeitverlauf im Ländervergleich.....	72
Abbildung 7. Verteilung der Ausbaustufen der Sägewerke im Ländervergleich am 13.2.2014 (28. Snapshot).....	74
Abbildung 8. Verteilung der Ausbaustufen der Luxusminen im Ländervergleich am 13.2.2014 (28. Snapshot).....	75
Abbildung 9. Verteilung der Ausbaustufen der Wunder im Ländervergleich am 13.2.2014 (28. Snapshot).....	75
Abbildung 10. Entwicklung der Ausbaustufen der Sägewerke im Ländervergleich über die Zeit....	76
Abbildung 11. Häufigkeitsverteilung Spieler am 13.02.2014 bezüglich der abhängigen Variable Rathauslevel / Spieldauer in Ikariam über alle Länder.....	81
Abbildung 12: Häufigkeitsverteilung der Investitionen in die Gemeingüter: Prozentualer Anteil der investierten Ressourcen am 13.2.2014 (28. Snapshot) über alle Länder.....	82
Abbildung 13: QQ-Diagramm für Investitionen in das Gemeingut in Prozent eigener Ressourcen über alle Länder für den letzten Snapshot (13.2).....	86
Abbildung 14. Häufigkeit von Spielertypen aggregiert über alle fünf Labor-PGG-Studien.....	89
Abbildung 15. Häufigkeit von Spielertypen nach Treatment differenziert (PNP = Partner-No Punishment, PP = Partner-Punishment, SNP = Stranger-No Punishment, SP = Stranger-Punishment); Daten aus allen fünf Labor-PGG-Studien nach Treatment aggregiert.....	89
Abbildung 16: Verteilungsmaße für die Anzahl der Städte pro Land für KRT-Inseln.....	91
Abbildung 17. Häufigkeit der Gruppentypen für Inseln mit mindestens zwei Spielern für alle Snapshots.....	92
Abbildung 18: Lagemaße für Trittbrettfahrer (0%) im Ländervergleich.....	94
Abbildung 19: Lagemaße für Kooperierende ($\geq 50\%$) im Ländervergleich.....	95
Abbildung 20. Anteil Kooperierender über die Zeit im Ländervergleich.....	96
Abbildung 21: Anteil Trittbrettfahrer über die Zeit im Ländervergleich.....	97
Abbildung 22: Anteil Reziprokatoren R2 über die Zeit im Ländervergleich.....	98
Abbildung 23: Anteil Reziprokatoren R1 über die Zeit im Ländervergleich.....	98
Abbildung 24: Summe aller Strafpunkte pro Spielertyp (durch Anzahl Spieler pro Spielertyp).....	105
Abbildung 25: Gesamt gegebene Strafpunkte pro Spielertyp und Studie.....	105
Abbildung 26: Anzahl aller Strafen pro Spielertyp (durch Anzahl Spieler pro Spielertyp).....	106
Abbildung 27: Summe aller erhaltenen Strafpunkte pro Spielertyp (durch Anzahl Spieler pro Spielertyp).....	107
Abbildung 28: Gesamt erhaltene Strafpunkte pro Spielertyp und Studie.....	108
Abbildung 29: Häufigkeit der Bestrafung pro Spielertyp (durch Anzahl Spieler pro Spielertyp).....	109
Abbildung 30: Anteil der Straftypen an den sieben Spielertypen in Prozent (Fehr et al. 2000).....	112
Abbildung 31: Anteil der Straftypen an den sieben Spielertypen in Prozent (Fehr et al. 2002).....	112
Abbildung 32: Anteil der Straftypen an den sieben Spielertypen in Prozent (Nikiforakis et al. 2008).....	113
Abbildung 33: Bestrafung pro Spielertyp und Gruppentyp (pro Häufigkeit Spielertyp).....	115
Abbildung 34: Erhaltene Bestrafung pro Spielertyp und Gruppentyp (pro Häufigkeit Spielertyp).....	116
Abbildung 35: Strafpunkte pro Spielertyp und Runde (Fehr et al. 2000).....	117
Abbildung 36: Strafpunkte pro Spielertyp und Runde (Fehr et al. 2002).....	118

Abbildung 37: Strafpunkte pro Spielertyp und Runde (Nikiforakis et al. 2008).....	118
Abbildung 38: Abnehmende Kosten für Bestrafung im Zeitverlauf nach Effektivität der Sanktionen	119
Abbildung 39: Abnehmende Kosten für Bestrafung im Zeitverlauf im Studienvergleich.....	119
Abbildung 40: Verlust des Verdienstes in Prozent vor und nach den Bestrafungskosten für Spielertypen (getrennt nach Treatment und als aggregierte Werte).....	122
Abbildung 41: Durchschnittlicher prozentualer Anteil am maximal möglichen Verdienst für alle Spielertypen in PGG.....	124
Abbildung 42: Durchschnittlicher prozentualer Anteil am maximal möglichen Verdienst für alle Spielertypen nach Treatment in PGG (im Uhrzeigersinn: Partner-P; Stranger-P; Stranger-NP; Partner-NP).....	125
Abbildung 43: Häufigkeitsverteilung der Spielertypen in Ikariam im Ländervergleich für alle 28 Snapshots (10 Monate).....	127
Abbildung 44: Erfolgsmaß (mittlerer Rathauslevel pro Spieldauer) im Ländervergleich für sieben Spielertypen ohne den ersten Snapshot in Ikariam.....	128
Abbildung 45: Häufigkeit Gruppentypen aggregiert über alle Labor-PGG.....	133
Abbildung 46: Häufigkeit der Gruppentypen nach Treatment in % über alle Labor-PGG aggregiert nach Treatment.....	134
Abbildung 47: Erreichte prozentualen Anteile des maximalen Verdienst über alle Runden und vier PGG-Studien pro Gruppentyp.....	135
Abbildung 48: Durchschnittlicher Verdienst pro Gruppentyp über vier PGG-Studien nach Treatment (im Uhrzeigersinn: Partner-P; Stranger-P; Stranger-NP; Partner-NP).....	136
Abbildung 49: Häufigkeit der Gruppentypen in Ikariam im Ländervergleich ohne den ersten Snapshot.....	138
Abbildung 50: Durchschnittliche Häufigkeit der Gruppentypen im Zeitverlauf für alle Länder (Mittelwerte).....	138
Abbildung 51: Durchschnittlicher Erfolg (Rathauslevel / Spieldauer) nach Gruppentyp im Ländervergleich für den letzten Snapshot (im Uhrzeigersinn: DE, EN, TR, GR).....	139
Abbildung 52: Durchschnittlicher Erfolg (Rathauslevel / Spieldauer) pro Anzahl der Städte nach Gruppentyp im Ländervergleich (im Uhrzeigersinn: DE, EN, TR, GR).....	140
Abbildung 53: Test auf Unterschiede im Erfolg (Rathauslevel / Spieldauer) pro Anzahl der Städte nach Gruppentyp für alle Länder am 13.2.....	141
Abbildung 54: Vergleich der zwei Erfolgsmaße (RHL/SD/Stadtanzahl und RHL/SD) für den Erfolg der Gruppentypen für den letzten Snapshot.....	145
Abbildung 55: Test auf Unterschiede im Erfolg der Spielertypen in fünf Gruppenzusammensetzungen in vier Labor-PGG in Prozent des maximal möglichen Verdienstes (links für P, rechts für NP-Treatments).....	147
Abbildung 56: Häufigkeit von Spielertypen in Gruppen in Ikariam.....	148
Abbildung 57: Verteilung des Erfolges zwischen Spielertypen für K- und T-Gruppen in Ikariam (für Griechenland und den 13.02).....	149
Abbildung 58: Durchschnittliche Spieldauer pro Spielertyp in Tagen im Ländervergleich für Inseln mit mehr als einem Spieler.....	151
Abbildung 59: Verdienst pro Runde in Prozent des maximal möglichen Verdienstes pro Spielertyp über die Zeit in vier PGG-Studien für 10 Runden (nur P Treatments).....	153
Abbildung 60: Verdienst pro Runde in Prozent des maximal möglichen Verdienstes pro Spielertyp über die Zeit in vier PGG-Studien für 10 Runden (nur NP Treatments).....	154
Abbildung 61: Verdienst pro Runde in Prozent des maximal möglichen Verdienstes pro Spielertyp von Runde 11 bis 50 für das Treatment PNP aus Gächter et al. 2008	155
Abbildung 62: Verdienst pro Runde in Prozent des maximal möglichen Verdienstes pro Spielertyp von Runde 11 bis 50 für das Treatment PP aus Gächter et al. 2008	155

Abbildung 63. Median des Erfolges (RHL/SD) über alle Länder aggregiert im Zeitverlauf mit dem ersten Snapshot.....	157
Abbildung 64. Median des Erfolges (RHL/SD) pro Anzahl der Städte über alle Länder aggregiert im Zeitverlauf mit dem ersten Snapshot.....	157
Abbildung 65: Durchschnittlicher Erfolg (RHL/SD) pro Stadtanzahl im Zeitverlauf für die Türkei ohne den ersten Snapshot.....	158
Abbildung 66. Test auf Unterschiede im Erfolg von Reziprokatoren (R1-R5) pro Gruppentyp (Verdienst in % des maximal möglichen Verdienstes für vier PGG-Studien).....	160
Abbildung 67. Erfolg von Spielertypen (RHL durch SD pro Stadtanzahl) und Median im Ländervergleich (alle Snapshots ohne 25.04).....	161
Abbildung 68: Erfolg (Rathauslevel pro Spieldauer und Stadtanzahl) pro Spielertyp und Land (im Uhrzeigersinn: DE, FR, TR, GR) nach 10 Monaten.....	162
Abbildung 69: Durchschnittlicher Erfolg (RHL/SD) pro Stadtanzahl im Zeitverlauf für Deutschland ohne den ersten Snapshot.....	197
Abbildung 70: Durchschnittlicher Erfolg (RHL/SD) pro Stadtanzahl im Zeitverlauf für England ohne den ersten Snapshot.....	198
Abbildung 71: Durchschnittlicher Erfolg (RHL/SD) pro Stadtanzahl im Zeitverlauf für Frankreich ohne den ersten Snapshot.....	198
Abbildung 72: Durchschnittlicher Erfolg (RHL/SD) pro Stadtanzahl im Zeitverlauf für Griechenland ohne den ersten Snapshot.....	199

A Einleitung

1 **Theoretischer Hintergrund: Evolutionäre Mechanismen zur Erklärung von Kooperation**

Seit Darwin hat die Frage, warum es kooperatives Verhalten gibt (sogar bei einfachsten Lebewesen wie Bakterien, siehe z. B. Brockhurst et al. 2008), zu zahlreichen Theorien geführt, wie dieses evolutionäre Rätsel zu lösen ist. Da unter dem Begriff Kooperation viele unterschiedliche Konzepte verstanden werden, folgt hier zunächst die in dieser Arbeit verwendete Definition:

„ein Verhalten, das einem anderen Individuum Nutzen bringt und wegen dieses Nutzens auf den Empfänger selektiert wird.“ Hier gilt: „[...] Kosten und Nutzen sind auf der Basis der Fitnesskonsequenzen von Verhalten für die gesamte Lebensspanne definiert“ (West et al. 2007, S. 416, übersetzt)

Im Laufe der Forschung wurde zunehmend klar, dass Kooperation – je nach biologischer Art und Umgebungsparametern – über verschiedene Mechanismen erklärt werden muss: Kooperative Interaktionen werden dabei oft nach ihrer theoretischen Erklärung eingeteilt. So kann man Mutualismus (Sachs & Simms 2006; Clutton-Brock 2002), biologische Märkte (Noë et al. 2001), reziproken Altruismus (Trivers 1971; Schino 2007), indirekte Reziprozität (Leimar & Hammerstein 2001; Nowak & Sigmund 1998), Handicap-Altruismus (Zahavi & Zahavi 1997; Johnstone 1995) und nepotistischen Altruismus (Dawkins 1976; Lehmann & Keller 2006) unterscheiden. Für eine Klassifikation und weitere Diskussion siehe Voland 2013, S. 71). Es wurde darauf hingewiesen (West et al. 2011), dass sich alle diese Ansätze unter die inklusive Fitnesstheorie subsumieren lassen (Hamilton 1964a; Hamilton 1964b).

Mit der Vielfalt der Erklärungsversuche und den interdisziplinären Fragestellungen zur Kooperation, die sich zwischen Soziobiologie, Ökologie und Ökonomie bewegen, ist oft auch ein unscharfer Begriffsapparat verbunden. So gibt es in der Tat sowohl Unklarheiten in den Begriffen als auch diverse Vorschläge zur Lösung vorhandener Doppeldeutigkeiten und Unschärfen.

Daher ist es wichtig, die verwendeten Begriffe klar zu definieren: Bei kooperativem Verhalten lassen sich als grundsätzliche Strategieoptionen *Altruismus* (Nutzen für andere bei eigenen Kosten) und *Mutualismus* (Nutzen für sich und Nutzen für andere) unterscheiden. Dem steht das Verhalten *Egoismus* (Nutzen für sich selbst mit möglichen Kosten für andere) gegenüber.

Da zwischen *proximaten* (kausale Wirkursachen, „Wie funktioniert es?“) bzw. *ultimaten* (funktionelle Zweckursachen, „Woher kommt es?“) Tinbergen 1963; Voland 2013) unterschieden werden muss, kommt eine Doppeldeutigkeit ins Spiel, wenn nicht explizit die Ebene genannt wird. Ein Beispiel: So erwartet man aus evolutionärer Sicht bei altruistischem Verhalten auf der proximatn Ebene (z. B. kooperativem Brüten) immer eine egoistische Begründung auf der ultimatn Ebene zu finden.

Doppeldeutig ist – zwischen den Disziplingrenzen – auch beispielsweise der Begriff Mutualismus (*mutualism*). Er wird in der ökologischen Literatur für Kooperation zwischen Arten verwendet (West et al. 2007), in der Soziobiologie jedoch im Sinne von wechselseitigem Nutzen (Voland 2013). Deshalb sollte zwischen zwischenartlichem Mutualismus (*interspecies mutualism*, also Kooperation zwischen Arten, z. B. Putzerfisch und Klient) und gegenseitigem Nutzen (*mutual benefit*) unterschieden werden (West et al. 2007). Zwischenartlicher Mutualismus beschreibt damit z. B. Interaktionen in biologischen Märkten, wo Betrug (*cheating*) eine Strategieoption darstellt, während bei Mutualismus im Sinne von wechselseitigem Nutzen Betrug sinnlos ist, da alle Beteiligten Nutznießer sind. Eine weitere Unterscheidung muss zwischen zeitlich nahen Auszahlungen (*payoffs*) und Fitnesskonsequenzen für das gesamte Leben gemacht werden (*lifetime effects on fitness*) (Bshary & Bergmüller 2008). Dies ist sowohl technisch als auch theoretisch anspruchsvoll (West et al. 2007): so kann ein Verhalten kurzfristig Nachteile einbringen, sich langfristig jedoch als vorteilhaft für ein Individuum erweisen (und anders herum). Evolutionäre Modelle (z. B. Trivers 1971; Hamilton 1964a; Hamilton 1964b; West et al. 2011) sprechen jedoch immer von den *lifetime effects on fitness*, also Fitness-Auswirkungen, die sich auf die gesamte Lebensspanne beziehen.

In der vorliegenden Arbeit wird der (übersetzte) Begriffsapparat nach West et al. (2007) verwendet. Laut diesen Definitionen ist

- Altruismus (*altruism*) „ein Verhalten, das Kosten für den Akteur beinhaltet und Nutzen für den Empfänger; hier, und weiter unten sind Kosten und Nutzen auf der Basis der direkten Fitnesskonsequenzen eines Verhaltens für die gesamte Lebensspanne gemeint.“,
- Egoismus (*selfishness*) „ein Verhalten, das Nutzen für den Akteur und Kosten für den Empfänger bringt.“,
- Gegenseitiger Nutzen (*mutual benefit*) „ein Verhalten, das sowohl für den Akteur als

auch für den Empfänger Nutzen bringt.“,

- Mutualismus (*mutualism*) „Kooperation zwischen biologischen Arten“,
- Direkte Fitness (*direct fitness*) „diejenige Fitnesskomponente, die durch das Verhalten eines Individuums entsteht, welches die Produktion von Nachkommen betrifft.“ Die Produktion von Nachkommen bezieht sich auf die lebenslangen Fitnesskonsequenzen und beinhaltet deshalb auch die soziale Platzierung, den Aufzuchterfolg und andere Faktoren. Schließlich ist
- Indirekte Fitness (*indirect fitness*) „diejenige Fitnesskomponente, die durch das Helfen bei der Reproduktion von Verwandten entsteht (West et al. 2007, S. 416).

Da sich dieser Begriffsapparat (West et al. 2011) der Problematik unterschiedlicher Begriffsverwendung in der Ökologie und der Soziobiologie bewusst ist, dafür Lösungen (oben angegeben) vorschlägt, Fehlverwendungen kritisch in einer Synthese beleuchtet und darüber hinaus eine Systematik für die Verbindung dieser Konzepte liefert, verwende ich diesen Begriffsapparat in der folgenden Arbeit. Als Trittbrettfahrer werden damit Individuen bezeichnet, die sich *selfish* verhalten; wenn wir von Vorteilen etwa von Kooperierenden gegenüber Trittbrettfahrern sprechen, ist die *direct fitness* angesprochen; Kooperierende sind demnach *altruists* oder *mutualists*; auch wenn im im Verlauf der Analysen festgestellt werden kann, dass ihr Verhalten auch sich selbst zu Gute kommt. Da sich diese Arbeit auf nur eine biologische Art (Menschen) konzentriert, ist immer die Rede von gegenseitigem Nutzen (*mutual benefit*) und nicht von Kooperation zwischen Arten (*interspecies mutualism*).

Weiterhin ist zu beachten, dass sich mögliche evolutionäre Wirkmechanismen ergänzen oder überlappen können – in der Praxis ist es oft schwierig, diese zu trennen. Oft ist es aber möglich, einige Erklärungen *auszuschließen*: So können weder inklusive Fitnesstheorie noch die Theorie der teuren Signale (Zahavi & Zahavi 1997) kooperatives Verhalten dann erklären, wenn es bei Nicht-Verwandten oder in anonymen Situationen zu Tage tritt. Deshalb wähle ich in dieser Arbeit Daten, die besonders dafür geeignet sind, einige theoretische Erklärungen auszuschließen. Gleichzeitig sind sie von hoher Bedeutung für viele Kooperationsprobleme, da sie prototypisch für die bedeutende, übergreifende Problemklasse der Gemeingutdilemmata stehen.

Aufgrund der Vielzahl an möglichen evolutionären Wirkmechanismen und der unklaren Befundlage, welcher Mechanismus welchem kooperativen Verhalten zu Grunde liegt, wird ein

theoretischer Rahmen benötigt, um Arbeitshypothesen und Befunde zu strukturieren. Deshalb stelle ich zunächst zwei konzeptionelle und begriffliche Systeme vor, in die ich meine Befunde theoretisch einordnen werde.

In der verwendeten Klassifizierung (Abbildung 1) wird zunächst nach altruistischen (*indirect benefits*) und mutualistischen Interaktionen (*direct benefits*) unterschieden. Sie ist deshalb besonders nützlich, weil sie präzise zwischen Mechanismen trennt, die sowohl für Geber und Empfänger nützlich sind und denjenigen, die einen Fitnessverlust für den Geber und Fitnessgewinn für den Empfänger bedeuten.

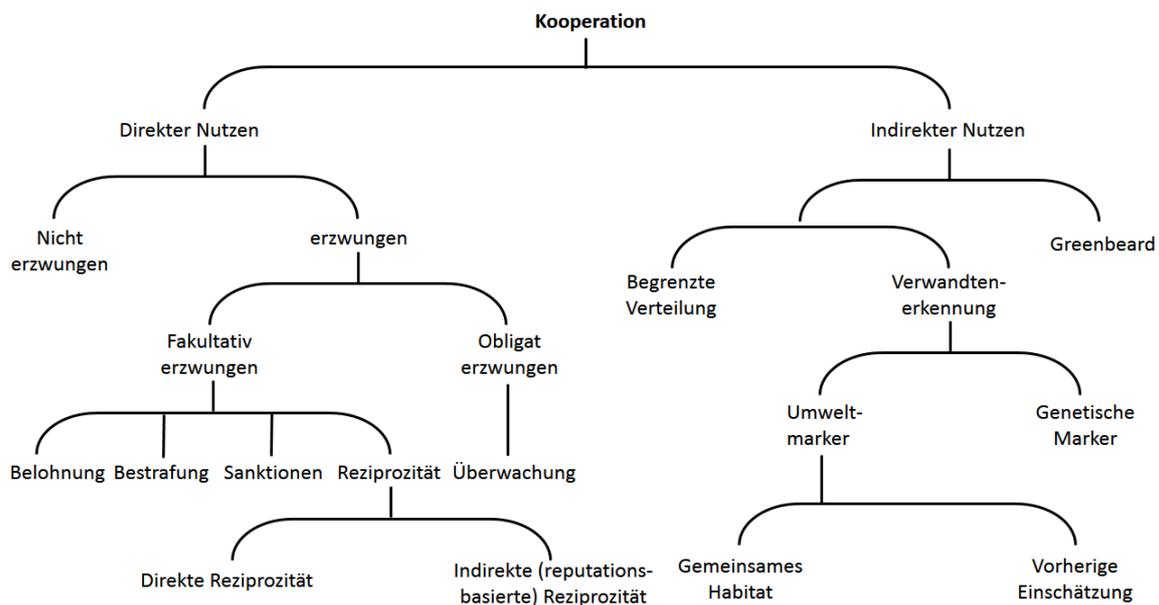


Abbildung 1. Eine Typologie evolutionärer Erklärungen für Kooperation (West et al. 2011, S. 236, übersetzt)

Eine weitere Klassifikation (Voland 2013, S. 71) trennt Kooperation und Altruismus in evolutionäre Szenarien auf, was die theoretische Handhabbarkeit erleichtert. Ich verwende letztere Klassifikation vor allem als Basis für Tabelle 1, um evolutionäre Mechanismen auf ihre Erklärungsplausibilität hinsichtlich der beobachteten Investitionen in Gemeingüter zu prüfen.

Bezüglich beider theoretischer Rahmen ist eine Anmerkung nötig, denn sie gehen davon aus, dass kooperatives Verhalten auf ultimativer Ebene im Mittel gen-egoistischen Interessen zuarbeitet. Das bedeutet freilich nicht, dass in allen Situationen fitness-maximierendes Verhalten zu erwarten ist. Möglicherweise ist menschliches Kooperationsverhalten in einer solchen Weise an einen Kontext aus kleinen, bekannten Gruppen mit spezifischen Kooperationspro-

blemen angepasst, dass es in großen, anonymen oder besonders künstlichen Umgebungen ein nicht funktionales Verhalten zeigt (*adaptation executing vs. fitness-maximizing behavior*). Das bedeutet allerdings weder, dass solches Verhalten nicht evolutionär erklärbar wäre, noch dass bestimmte Erklärungsmodelle ausgeschlossen werden können.

Auch wenn Verhalten demnach nicht immer die Fitness maximiert, so ist dennoch in der Auffassung von Kooperation als *Anpassung*, wie ich es in dieser Arbeit tue, eine genaue Passung auf ein bestimmtes evolutionäres Problem impliziert. So besteht beispielsweise die Passung bei nepotistischem Altruismus darin, dass im Mittel tatsächlich vor allem Verwandten geholfen wird in Abstufung des Verwandtschaftsgrades, und eben nicht Nicht-Verwandten (Hamilton 1964a; Hamilton 1964b). Dafür ist eine wie auch immer geartete Unterscheidungsfähigkeit (etwa über genetische oder Umweltmarker) notwendig. Dies gilt analog für Kooperation in biologischen Märkten, wo es eine Passung für die Wahl der Interaktionspartner geben muss (Noë et al. 2001), andernfalls kann dieses Erklärungsmodell nicht zutreffen.

Anhand des theoretischen Rahmens von West et al. (2007) wird im Folgenden jeder Kooperationsmechanismus, also jeder Typ evolutionärer Erklärung, nach der Systematik von Abbildung 1 an Beispielen aus der Tier- und Pflanzenwelt erläutert. Damit wird das Phänomen „biologische Kooperation“ theoretisch systematisch aufgeschlüsselt mit empirischen Beispielen belegt. Zunächst wird der linke Zweig von Abbildung 1 diskutiert, der Kooperation über den *direkten Nutzen* für ein Individuum erklärt. Weitere Versuche, biologische Kooperation theoretisch zu fassen, bieten beispielsweise Axelrod & Hamilton 1981, Bshary & Bronstein 2004, Nowak 2006 oder Bergmüller et al. 2007.

Nicht erzwungener direkter Nutzen:

Hier sind vor allem zwei Untermechanismen zu nennen, zum einen Gruppenvergrößerung (*group augmentation*), bei der es durch kooperatives Verhalten eines Individuums zur Vergrößerung der Gruppe kommt, was dann in der Zukunft zu einer Zunahme der Fitness jedes Individuums in der Gruppe führt, und damit eben auch der Fitness des kooperativen Individuums (Clutton-Brock 2002; West et al. 2011).

Zum anderen ist koordiniertes Jagen hier einzuordnen, wie es etwa bei afrikanischen Wildhunden (Creel & Creel 1995) oder getüpfelten Hyänen der Fall ist (Drea & Carter 2009), da es sich um ein Beiprodukt eines individuellen Verhaltens handelt, das den anderen Gruppenmitgliedern zu Gute kommt. Viele Koalitionen sind ein Beispiel für mutualistische Kooperation, wie etwa einsichtsvolles kooperatives Problemlösen bei Vögeln (Seed et al. 2008), koopera-

tives Problemlösen bei Hyänen (Drea & Carter 2009), Nahrungsteilung bei Affen (de Waal 1989) sowie gemeinsame Jagd bei Schimpansen in freier Wildbahn (Boesch 2001) bzw. deren Kooperation in Experimenten (Melis et al. 2006).

Ein weiteres, häufig auftretendes, Kooperationsproblem im Tierreich ist das gemeinsame Verteidigen von Ressourcen oder Territorien. Dies führt zu komplexen, individuell sehr verschiedenen Strategien beispielsweise bei Löwinnen (Heinsohn & Packer 1995) und ist beispielsweise bei Meerkatzen von Geschlecht und Reproduktionsphase abhängig (Mares et al. 2012). Diese Ergebnisse verdeutlichen das *freerider*-Problem bei gemeinsamen Territorien, da die Kosten ungleich, der Nutzen aber gleich verteilt ist. Die Bewältigung dieser Problematik stellt möglicherweise einen starken selektiven Druck bei sozial lebenden Primaten dar (Willems et al. 2013). Zwischen vielen Organismengruppen kommt es zu einer *tragedy of the commons*, auch wenn die Auswirkungen durch verschiedene Faktoren oft abgeschwächt werden (Rankin et al. 2007).

Erzwungener direkter Nutzen:

Kooperation kann auch durch verschiedene Mechanismen erzwungen werden. Dazu gehören etwa *Belohnung* und *Bestrafung*. Einen theoretischen Überblick über Bestrafung in der Biologie geben Clutton-Brock & Parker (1995).

Ein Beispiel sind Symbiosen zwischen Pflanzen und N₂-fixierenden Bakterien (Neuhauser & Fargione 2004). Hier kommt es zu komplexen Wechselwirkungen zwischen den Symbionten, etwa zu Überwachung seitens der Pflanze (wird das Gut N₂ bereitgestellt?) und Bestrafung (verminderte Sauerstoffzufuhr), falls es nicht produziert wird (Kiers et al. 2003). Diese Bestrafung führt zu drastisch reduzierten Wachstumsraten der nicht-kooperierenden Bakterien.

Bestrafung findet sich auch bei vielen anderen Arten: so bestrafen dominante weibliche Erdmännchen (*Suricata suricatta*) die Reproduktion anderer Weibchen durch temporären Ausschluss aus der Gruppe (Young et al. 2006); es sind viele Mutualismen (nicht *mutual benefit*) zwischen Feigenbäumen und bestäubenden Wespen bekannt, bei denen ein enger Zusammenhang zwischen der erwarteten Bestrafung durch die Feigenbäume und der Bestäubungswahrscheinlichkeit durch die Wespen besteht (Jandér & Herre 2010).

Auch bei Putzerfischen, die nicht putzen, sondern Hautfetzen herausbeißen, sind Sanktionen seitens der Kunden (etwa seltenere Besuche oder Jagd auf die kleineren Putzerfische beobachtet worden (Bshary & Grutter 2005; Bshary & Grutter 2002). Diese Interaktionen zwischen

Spezies lassen sich auch als *biologische Märkte* begreifen (Noë et al. 2001). Hier werden Waren oder Dienstleistungen getauscht. So bieten etwa Putzerfische die Befreiung von Parasiten als Dienstleistung an (Grutter 1999), während die Klienten hierfür zum Teil ihr Revier verlassen und darauf verzichten müssen, den kleineren Putzerfisch zu fressen. Es kommt zu komplexen strategischen Überlegungen, die von einer Vielzahl von Parametern abhängen – etwa Anzahl der Wettbewerber (Soares et al. 2008) und Qualität der Dienstleistung (Bshary & Schäffer 2002). Wichtig sind auch Reputation (Tebbich et al. 2002) und Kosten-Nutzen-Rechnungen, die Wartezeit, Defektionswahrscheinlichkeit durch den Putzerfisch (Bshary & Grutter 2002) sowie Nahrungsangebot für die Putzer (Grutter 1995) und Parasitenbefall der Klienten (Soares et al. 2008) umfassen. Warum solche Mechanismen des zwischenartlichen Mutualismus stabil bleiben können, ist von hohem theoretischem Interesse, weil man theoretisch erwartet, dass die natürliche Selektion dagegen arbeitet, also Instabilität die Folge sein müsste (Leigh 2010; Sachs & Simms 2006; Raihani et al. 2012).

Ein weiterer Mechanismus in dieser Kategorie ist Überwachung (*policing*). Ein Beispiel hierfür ist die unterdrückte Fortpflanzung von Ameisen (*Formica fusca*, Helanterä & Sundström 2007), Honigbienen (*Apis mellifera*) und Wespen (*Vespula Vulgaris*, Wenseleers et al. 2004). Würden sich Arbeiterinnen egoistisch fortpflanzen, wäre die Kooperation eines Stocks in Gefahr. Um dies zu verhindern und das Kooperationsniveau aufrecht zu erhalten, werden von Arbeiterinnen gelegte Eier durch andere Arbeiterinnen oder die Königin selbst entfernt (Wenseleers et al. 2004). Die Betrugswahrscheinlichkeit von Arbeiterinnen hängt dabei wie erwartbar stark vom Grad der Verwandtschaft mit der Königin ab. Dieser Grad ist bei staatenbildenden Insekten unterschiedlich, je nachdem, wie oft sich die Königin verpaart (Ratnieks et al. 2006). Auch bei kooperativ brütenden Vögeln, etwa den Prachtstaffelschwänzen (*Malurus cyraneus*), bestrafen dominante männliche Revierinhaber jüngere Helfer am Nest für ihre (experimentell induzierte) Abwesenheit mit aggressiver Jagd (Mulder & Langmore 1993).

Schließlich sind noch Partnerwahl und Partnerwechsel als Möglichkeiten der Bestrafung zu nennen, wie sie etwa in biologischen Märkten gang und gäbe sind (Noë 2006).

Direkte und indirekte Reziprozität:

Während es für direkte und indirekte Reziprozität (auch Reputation oder *image scoring* genannt Nowak & Sigmund 1998) bei Menschen viele Beispiele gibt (Wedekind & Milinski 2000; Trivers 1971), sind die Belege für Tiere dünn gesät (West et al. 2011). Ausnahmen machen möglicherweise Schimpansen (Russell et al. 2008) und Fische (Bshary & Grutter

2006) bei denen eine Art *image scoring* beobachtet wird.

Im Folgenden wird nun der rechte Zweig der Abbildung diskutiert, der Kooperation wegen des *indirekten Nutzens* für ein Individuum erklären kann. Indirekter Nutzen entsteht, wenn genverwandten Individuen bei der Reproduktion geholfen wird, etwa durch Kooperation.

Begrenzte Verteilung:

Kooperation kann durch ein begrenztes Habitat begünstigt werden. Nachweise dafür findet man bereits bei sehr einfachen Lebewesen, etwa Bakterien (Brockhurst et al. 2006; Kreft 2004). Hier ist der Verwandtschaftsgrad oft sehr hoch und die Möglichkeit, das Habitat zu wechseln, oft nicht vorhanden. Entscheidend ist aber oft die räumliche Trennung zu Trittbrettfahrern (egoistischen Mutanten), die dann den Nutzen aus der Kooperation, den Biofilm, nicht abgreifen können (Donlan & Costerton 2002; Rainey & Rainey 2003; Brockhurst et al. 2008). Aber auch Territorialität, etwa bei Vögeln, kann Kooperation fördern, wenn die Promiskuität niedrig genug ist (Cornwallis et al. 2010).

Greenbeard:

Kooperation ist in einer gen-zentrierten Sichtweise (Dawkins 1976; Dawkins 1982/2008) auch bei *Nicht-Verwandten* dadurch denkbar, dass es primär um die Identität von Gen-loci geht (Hamilton 1964a; Hamilton 1964b), auch wenn Verwandtschaft derjenige Mechanismus ist, der dafür meist verantwortlich ist. Der sogenannte Greenbeard-Mechanismus koppelt ein phänotypisch für andere erkennbares Merkmal (der hypothetische grüne Bart) mit kooperativem Verhalten bei Nicht-Verwandten. Hier ist ein Gen für die Ausprägung zweier Merkmale verantwortlich (Pleiotropie). Bislang wurden nur einige wenige *Greenbeards* gefunden: vier in Bakterien und einer in Ameisen (Gardner & West 2010).

Genetische Marker:

So wie ähnliche Gene durch bestimmte Merkmale im Phänotyp (*Greenbeard*) erkannt werden können, so ist Ähnlichkeitserkennung auch durch genetische Marker möglich. Ein mögliches Beispiel ist gemeinsame Ansiedlung und Kooperation auf Grund genetischer Ähnlichkeit im Larvenstadium von Plankton (*Botryllus schlosseri*, Grafen 1990). Allerdings sind die Belege, ähnlich wie bei Greenbeard-Effekten leicht missdeutbar und nur wenige Beispiele bekannt. Zwei Beispiele sind Erkennung über Kohlenwasserstoffmarker bei Ameisen (*Formica truncorum*, Boomsma et al. 2003) und Duftmarken bei Eichhörnchen (*Spermophilus beldingi*, Mateo 2002). Interessanterweise kann eine eng verwandte Eichhörnchenart (*Spermophilus*

lateralis) zwar ebenfalls zwischen Verwandten unterscheiden, zeigt aber kein nepotistisches Verhalten. Möglicherweise werden daher bereits vorhandene Mechanismen für die Verwandtenerkennung ko-optiert, was der Forderung nach der spezifischen Entwicklung für diesen Mechanismus (Grafen 1990) widerspricht.

Umweltmarker – vorheriger Umgang bzw. gemeinsames Habitat von Individuen:

Nach Einschätzung von West et al. (2011) sind (oft in der Kindheit) gelernte Indizien aus der Umwelt der häufigste Mechanismus für Individuen, um genetische Verwandtschaft festzustellen. Dafür gibt es viele Beispiele: chemische Marker, die Ameisen (*Formica fusca*) erlauben, Eier einer spezifischen Königin zuzuordnen (Helanterä & Sundström 2007); Assoziationslernen bei kooperativen Brütern (z. B. Schwanzmeisen, *Aegithalos caudatus*, Russell & Hatchwell 2001) und andere Mechanismen bei diversen Vogelarten (Komdeur & Hatchwell 1999). Auch Alarmrufe, die vor allem genetischen Verwandten nützen, sind in diese Kategorie einzuordnen (Seyfarth & Cheney 2003; Fischer 2010). Allerdings sind die Erkennungsraten bei einigen Wirbeltierarten nicht optimal und auch die Erkennungskosten sind oft so hoch, dass sich eine Unterscheidung nicht immer lohnt (Keller 1997).

2 Soziale Dilemmasituationen

Nach diesem Überblick über evolutionäre Mechanismen zur Erklärung von Kooperation im Tierreich wendet sich der folgende Abschnitt nun Kooperationsfragen bei Menschen zu. Im Bereich der Kooperation zeigt der Mensch ein im Tierreich einzigartiges Verhalten, obwohl wir uns davor hüten müssen, ihm eine Sonderstellung zuzuweisen (Frey 2010). Tatsächlich kennen wir kein anderes Tier, das globales, stark vernetztes und über lange Zeiträume andauerndes Kooperationsverhalten zeigt, wie es der weltweite Handel oder global agierende Hilfsorganisationen zeigen. Evolutionäre Studien über gemeinsame Großwildjagd und Fleischteilung in der Gruppe bei Jägern und Sammlern können auch die Ursprünge und Stabilität dieses Verhaltens demonstrieren (Hill 2001; Hawkes 1993).

Diese hohe Kooperationsbereitschaft von Menschen wird besonders kontrovers diskutiert (Marlowe et al. 2011): wie kann Kooperation entstehen (West et al. 2007), wie kann sie bei ständiger Gefahr durch Defektoren stabil bleiben (Fehr et al. 2002) und welche evolutionären Mechanismen können sie erklären? Prinzipiell kommen alle Mechanismen in verschiedenen Kontexten zum Tragen (Nowak 2006): direkte Reziprozität (Guala 2012; Burnham & Johnson 2005; Gächter & Herrmann 2009); indirekte Reziprozität (Leimar & Hammerstein 2001; Rockenbach & Milinski 2006); räumliche Verteilungen (Helbing & Yu 2009) und kulturelle

Gruppenselektion (Gintis et al. 2003). Einige dieser evolutionären Mechanismen verweisen auf die Bedeutung der sozialen Lebensweise (Wilson 2013), die Kooperation über Gruppenkonkurrenz fördert. Andere sehen die Grundlagen dagegen eher in Vorteilen durch Reziprozität (Trivers 1971). Kritik wird dabei vor allem an der neuen Idee der „starken Reziprozität“ geübt, der als weiterer grundsätzlicher Mechanismus gehandelt wird (Hagen & Hammerstein 2006).

Bei Menschen besteht im Gegensatz etwa zu Ameisen oder Termiten eine Besonderheit darin, dass Kooperation auch über genetische Unterschiede und Gruppengrenzen hinweg aufrechterhalten werden kann. Zusätzlich kann sich Vertrauen über wiederholte Interaktionen entwickeln und über sprachliche Mittel (wie Verträge) formalisiert festgehalten werden. Gerade die Möglichkeit der *Partnerwahl* für kooperative Unternehmungen plausibilisiert, dass es von Vorteil ist, als fairer, kooperationsbereiter Partner von anderen wahrgenommen zu werden; dies ist in den meisten Fällen nur dadurch möglich, dass man dies auch durchgängig ist, was ein teures Signal darstellt (Fetchenhauer & Bierhoff 2004).

Möglicherweise lässt sich menschliche Kooperation auch über den Mechanismus Reziprozität erklären. Allerdings ist trotz vieler Arbeiten zur Reziprozität bei Menschen (Trivers 1971; Binmore 2006; Wedekind & Milinski 2000) deren Rolle noch unklar (West et al. 2011).

Menschliche Kooperation kann jedoch unter Umständen aufrechterhalten werden, bei denen im übrigen Tierreich keine Kooperation zu finden ist; so gibt es viele Beispiele für globale und langfristige Kooperation mit einer großen Zahl sehr heterogener Beteiligter, die nicht miteinander verwandt sind. Beispiele sind Blutspenden oder Transaktionen auf globalen Märkten. Aus diesem Grund scheidet die inklusive Fitnessstheorie als Erklärung für solche Phänomene aus. Übrig scheint nur die Theorie der teuren Signale (Zahavi & Zahavi 1997) und die Erklärung über Reziprozität zu bleiben. Im Gegensatz dazu haben sich viele Kandidaten aus dem Tierreich bei näherer Analyse entweder als nepotistischer Altruismus, biologische Märkte oder Situationen mit wechselseitigem Nutzen herausgestellt und eben nicht als Reziprozität.

Erschwerend für Erklärungsversuche menschlicher Kooperation ist die grundsätzliche Problematik, evolutionäre Mechanismen voneinander zu trennen. Als Beispiel kann die Frage nach der erstmals auftretenden Entstehung kooperativen Verhaltens bei Menschen dienen, da hier möglicherweise – zeitlich versetzt und abhängig voneinander – verschiedene Mechanismen am Werk waren:

„In Fällen, in denen diese Prozesse [direkte und indirekte Reziprozität, Bestrafung, Gruppenvergrößerung und kostspielige Signale] zum Tragen kommen, ist es wahrscheinlich, dass Kooperation anfänglich wegen indirekter Fitnessvorteile oder gemeinsamer Interessen entstehen konnte. Erst danach selektieren Mechanismen wie Reziprozität oder Bestrafung auf höhere Kooperationslevel, selbst wenn keine Verwandtschaft gegeben ist. Deshalb könnte es unnötig sein für Menschen beweisen zu wollen, wie Kooperation in nicht-verwandten Populationen de novo entstehen konnte, wenn sie in einer Gruppe verwandter Hominiden entstand.“ (West et al. 2011, S. 239, übersetzt)

Unabhängig vom Mechanismus steht fest, dass kooperatives menschliches Verhalten grundlegenden evolutionären Selektionsdrücken ausgesetzt ist, die von Konflikten zwischen Eltern und Kindern (Trivers 1974) über psychische Mechanismen etwa zur Partnerwahl (Buss 2004) bis hin grundlegenden sozialen Verhaltensweisen reichen (Alexander 1987). Fest steht auch, dass die grundsätzliche Perspektive für menschliche Kooperation eine evolutionäre sein sollte, um langfristige Fitnesskonsequenzen mit einbeziehen zu können.

Wie eben demonstriert, ist das Thema menschliche Kooperation sehr breit. Ich führe im Folgenden dieses Thema eng, indem ich mich auf *soziale Dilemmasituationen* beschränke. Dies hat vor allem zwei Gründe. Zum einen sind sie besonders geeignete Untersuchungsobjekte für Kooperationsprobleme (vgl. Axelrod 1984/2000), da sie Kooperationsentscheidungen auf ein Dilemma und einige wenige gleichartige Entscheidungen zuspitzen. Dadurch können grundlegende Merkmale von Kooperation einfacher herausgearbeitet und theoretisch bearbeitet werden (Fehr & Schmidt 1999). Zum anderen kann mit dieser Klasse von Situationen eine breite Palette an wichtigen menschlichen Kooperationsproblemen charakterisiert werden. Soziale Dilemmata, wie sie etwa in (iterierten) Gefangenendilemmata (public goods game, im Folgenden PGG abgekürzt), Trust-, Ultimatum- oder Diktatorspielen vorkommen, sind demnach sowohl experimentell einfach zu handhaben als auch theoretisch gut zu beschreiben und decken zudem eine Vielzahl an realen Situationen in ihrer Problemstruktur gut ab. Die Grundstruktur von sozialen Dilemmasituationen wird deshalb im Folgenden kurz charakterisiert.

Soziale Dilemmata zeichnen sich dadurch aus, dass es für jeden Einzelnen rationaler ist, egoistisch zu sein, obwohl die Kooperationsgewinne, sofern alle Beteiligten kooperieren, die Auszahlungen für beiderseitigen Egoismus übersteigen. Um die Komplexität zu vereinfachen, sind in der experimentellen Form oft nur zwei Strategieoptionen erlaubt. Dies wird oft im

(iterierten) Gefangenendilemma dargestellt, das durch folgende Auszahlungen charakterisiert wird (Axelrod 1984/2000):

$$T (\textit{temptation}) > R (\textit{reward}) > P (\textit{punishment}) > S (\textit{sucker's payoff})$$

Zusätzlich gilt:

$$2 R > T + S$$

T ist dabei die Belohnung für Egoismus, der auf einen Kooperationsversuch trifft (dieser erhält S). R erhalten die Akteure, wenn beide kooperieren, P, wenn beide defektieren (Egoismus zeigen).

Wird das Gefangenendilemma iteriert, so spricht man von einem *public goods game*. Bei dieser Untersuchungsmethode zeigen sich vor allem dann signifikante Unterschiede im Kooperationsniveau, wenn *externe* Umgebungsparameter variiert werden, z. B. durch Bestrafung (Fehr & Fischbacher 2003), Reputation (Rockenbach & Milinski 2006) oder Kommunikation (Ostrom et al. 1994). Dies wird in Teil B 4 weiter ausgeführt.

Eine Variation dieser Untersuchungsmethode sind Feldexperimente (Cardenas et al. 2000). Sie können – auf Kosten von Komplexitätsverlust und Kontrollierbarkeit – eine höhere ökologische Validität für sich beanspruchen, da sie das Problem öffentlicher Güter meist kontextspezifisch einkleiden (*framing*). Weiterhin untersuchen sie Populationen, die heterogener als die typischerweise an Experimenten teilnehmenden amerikanischen Studenten ohne Abschluss (*undergraduates*) sind. Zudem hinterlassen sie teilweise auch (positive) Effekte auf das reale Verhalten (Kimmich 2012; Cardenas 2000). Die Übertragbarkeit von Labor- zu Feldergebnissen ist grundsätzlich gegeben, aber mit diversen Vorbehalten versehen (Mitchell 2012). Von besonderem Interesse sind dabei Daten, die auf natürliche Weise entstehen, aber ungewollt kontrolliert sind. Ein Beispiel ist eine Bitte der ETH Zürich an alle Studierenden, eine Geldsumme ihrer Wahl in einen sozialen Fonds für andere Studierende einzuzahlen. Durch einen Wechsel in der Administration änderte sich die Art der Bitte (vorher: Rechnung über Studiengebühren mit oder ohne des Betrags für den Sozialfonds; nachher: Ankreuzen der Wahl, ob man in den Sozialfonds einzahlt oder nicht). So entstehen hier zwei unterschiedliche Bedingungen (vergleichbare Treatments) mit unterschiedlicher Bereitschaft, in das öffentliche Gut (den Sozialfonds) einzuzahlen, die allerdings aus ökonomischer Sicht völlig identisch sind (Frey & Meier 2004).

Ein weiteres Feld sind Fallstudien zu Allmenderessourcen. Hier werden reale Fälle, etwa

Bewässerungssysteme, Waldmanagement oder lokale Fischereien untersucht, bei denen eine Vielzahl potentiell einflussreicher Variablen wie etwa Anzahl der Akteure oder Fairness des Regelsystems zu berücksichtigen sind (Tang 1989; van Laerhoven 2010). Als einflussreich haben sich Führungsqualität und Sozialkapital (Gutiérrez et al. 2011), Regenerationsfähigkeit der Ressource und Mitbestimmung der Beteiligten (Frey & Rusch 2014), sowie klare Grenzen und geeignete Bestrafungsmechanismen (Ostrom 1990) erwiesen, auch wenn dies nach Kontext, Land und Sektor stark differenziert werden muss.

Öffentliche Güter sind Paradebeispiele für soziale Dilemmata. Sie sind dadurch gekennzeichnet, dass sie schwer aufzuteilen, aber leicht zugänglich sind (Hess & Ostrom 2003). Beispiele sind Sonnenuntergänge, saubere Luft oder allgemeines Wissen. Es ist praktisch unmöglich, bzw. mit hohem Aufwand verbunden, Personen den Zugang zu einem Sonnenuntergang zu verwehren. Gleichzeitig lässt sich ein Sonnenuntergang auch nicht aufteilen. Ein öffentliches Gut kann theoretisch auch von einer Person alleine erzeugt werden kann (z. B. Bau eines Leuchtturms), während Kooperation trivialerweise mindestens zwei Beteiligte benötigt.

Im Unterschied dazu sind Allmendegüter (*common pool resources*) bei gleichbleibender schwieriger Ausschließbarkeit vergleichsweise leichter aufzuteilen (*divisibility*). Ihr Hauptunterschied zu öffentlichen Gütern besteht darin, dass sie durch Übernutzung oder Verschmutzung zerstört werden können (*subtractability*). Dies ist z. B. bei sauberer Luft, Parks oder den weltweiten Fischbeständen der Fall, während Sonnenuntergänge nicht übernutzt werden können.

Wenn dagegen in der Literatur der experimentellen Ökonomie die Rede von öffentlichen Gütern (*public goods*) ist, dann besteht das Dilemma in der Regel darin, ein öffentliches Gut durch Einzahlungen der Beteiligten zu *erzeugen*. Dies geschieht in der Regel dadurch, dass vom Experimentleiter bereitgestellte Ressourcen (meist Geld) in ein „öffentliches Gut“ eingezahlt werden und die Effizienz einer solchen Kooperation durch die Verdoppelung der Einzahlungen simuliert wird. Im Gegensatz zu dem Nicht-zustande-Kommen des öffentlichen Gutes in *public goods games* ist in den selteneren *common pool resource*-Spielen (Anderies & Janssen 2013), die Übernutzung simuliert und führt bei zu hoher Extraktion zum Zusammenbruch der Ressource, die sich in einem solchen Fall auch nicht wieder erholt.

Die eben genannten Modelle für soziale Dilemmata (z. B. PGG) beschreiben eine extrem breite Problemklasse. Viele wichtige Kooperationsprobleme können damit simuliert werden.

Einige Beispiele – systematisiert nach Problemklasse – sollen die Wichtigkeit dieser Art von Kooperationsproblemen verdeutlichen.

Eine erste Klasse von Kooperationsproblemen betrifft Güter, die schwer zu *erzeugen* sind. Beispiele sind Blut- oder Organspenden. Das öffentliche Gut ist hier die Verfügbarkeit von Blut bzw. Organen im Falle eines Unfalls für jedes Mitglied der Solidargemeinschaft. Nur ein Bruchteil der Bevölkerung spendet jedoch, da dies mit Aufwand verbunden ist. Wie wichtig institutionelle Rahmenbedingungen sind, kann an diesem Beispiel gut belegt werden, da in Österreich, Frankreich oder Portugal mehr als 99,9% der Bevölkerung ihre Zustimmung zur Organspende geben, in Deutschland jedoch nur 12% und in Dänemark sogar nur 4,3%. Die untenstehende Abbildung 2 zeigt diese Unterschiede deutlich auf.

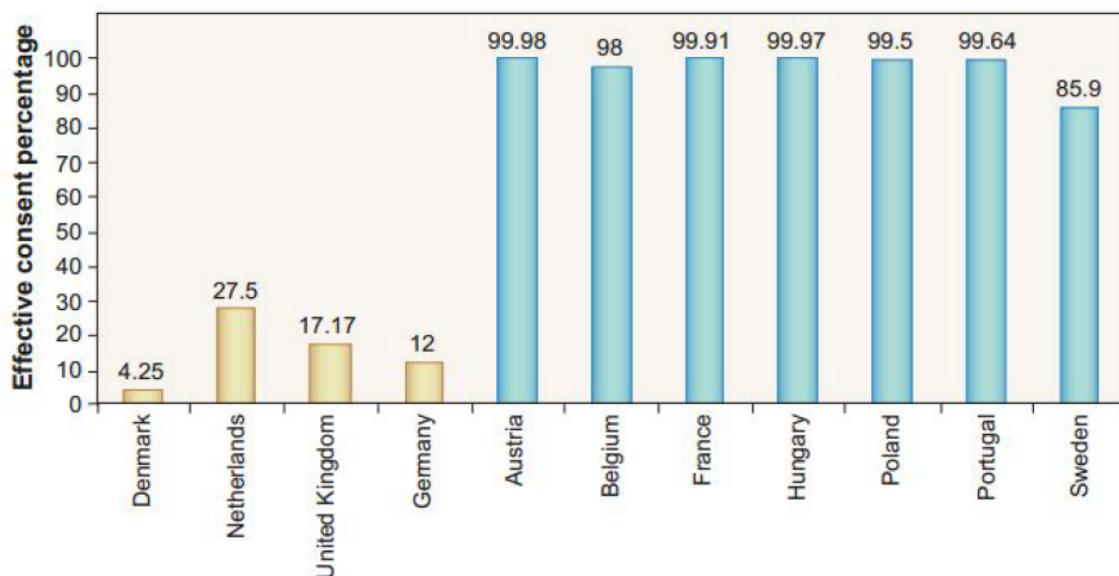


Figure 12.2: Effective consent rates, by country. Explicit consent (*opt-in*, gold) and presumed consent (*opt-out*, blue). (Source: Johnson and Goldstein, 2003).

Abbildung 2: Unterschiede in der Organspendebereitschaft je nach *opt-in* oder *opt-out* Verfahren

Der Grund für diesen gewaltigen Unterschied findet sich nicht in grundsätzlichen kulturellen oder moralischen Unterschieden, sondern liegt schlicht darin, dass in allen Ländern mit hohen Zustimmungsquoten zur Organspende die Zustimmung ohne expliziten Widerruf gilt (*opt-out*), während man in allen Ländern mit niedrigen Quoten der Spende explizit zustimmen muss (*opt-in*). Das Problem besteht also in der Herstellung der Zustimmung zur Kooperation, die dann wiederum das Gut erzeugt.

Eine zweite Klasse von Kooperationsproblemen betrifft Güter, die leicht zu *zerstören* sind. Prominentestes Beispiel ist die Vermüllung der Umwelt auf individueller, lokaler, nationaler und globaler Ebene. So ist es für Firmen kostengünstiger, Giftmüll auf Meeren illegal zu verklappen anstatt zu entsorgen. Weitere bekannte Problematiken sind Klimaschutz (CO₂-Minderung) und die Übernutzung natürlicher Ressourcen wie Wälder, Fischbestände oder anderer Tierarten. Natürliche Ressourcen, die übernutzt werden können und auf die meist nur eine bestimmte Gruppe Zugriff hat, werden als Allmendeprobleme (*common pool resources*) bezeichnet (Ostrom 1990).

Wenn man nach dem *Ergebnis* von Kooperation fragt, wird oft das Abkommen von Montreal zur Ächtung ozonschädlicher Fluorkohlenwasserstoffen als *Erfolg* genannt. Hier gelang es innerhalb kurzer Zeit (zwei Jahre, von 1985-1987) die genannten Schadstoffe weltweit zu ächten. Viele Staaten (197) konnten sich in einem Vertrag einigen. Als Beispiel für *Misserfolg*, also gescheiterter Kooperation auf globaler Basis, lassen sich die Klimaschutzkonferenzen zur CO₂-Minderung anführen.

An diesen sehr komplexen Beispielen wird auch klar, dass mögliche Lösungsversuche sozialer Dilemmata auf verschiedenen *Zugängen* beruhen können: so sieht die Institutionenökonomie hier vor allem ein Designproblem der Regelsysteme (Ostrom 1990), die traditionelle Ökonomie hingegen eher die Anreizoptimierung für nutzenmaximierende Individuen (Bolton & Ockenfels 2000), während Ansätze aus der Biologie eher die evolutionäre Herkunft unserer Präferenzen in den Vordergrund stellen (Volland 2003). Gemeinsam ist allen Ansätzen die Frage, wie sich die Entscheidungen der Akteure beeinflussen lassen – sei es über externe oder interne Parameter.

Allerdings ist für Kooperation in sozialen Dilemmasituationen weder ein Konsens auf Datenebene (Konvergenz der experimentellen Ergebnisse) noch auf Theorie-Ebene (Fehr & Schmidt 1999; Bolton & Ockenfels 2000) in Sicht. Es ist weder klar, welchen Stellenwert einzelne Parameter wie Bestrafung tatsächlich haben (Balliet et al. 2011; Chaudhuri 2011), noch wie sie in theoretische Modelle eingebaut werden könnten. Tatsächlich gibt es nur wenige Versuche in der Ökonomik, Experimentalergebnisse zu modellieren (z. B. Bolton & Ockenfels 2000). Diese scheitern an der Breite der berichteten Ergebnisse bzw. werden bereits wegen ihrer Annahmen scharf kritisiert (Hagen & Hammerstein 2006). Natürlich lassen sich die Ergebnisse vor einem evolutionären Hintergrund verstehen (Volland 2013), aber es ist weitgehend unklar, wie sich die allgemeinen theoretischen Modelle auf die Experimentalergebnisse

anwenden lassen. Andererseits gibt es aber auch einige robuste Befunde, auf die ich unten zu sprechen komme (Abschnitte B 1 und 2).

Besonders der Zugang aus der Biologie erweist sich als hilfreich, um moderne Kooperationsprobleme verstehen zu können; und hier vor allem die Relevanz unserer Vergangenheit als Jäger und Sammler. Die – gut abgesicherte – Hypothese lautet, dass der Großteil unserer Anpassungen an Bedingungen stattfand, die durch verwandte Kleingruppen (Hill et al. 2011) und andere Merkmale gekennzeichnet waren, die im so genannten *environment of evolutionary adaptedness* (dem EEA, einer statistischen Zusammenstellung von Selektionsdrücken) zusammengefasst werden können (vgl. Barkow et al. 1992). Hier scheinen vor allem soziale Kontextparameter über Kooperation bei öffentlichen Gütern (wie Teilung der Jagdbeute) zu entscheiden (Hill 2001), auch wenn Hypothesen wie das *Handicap*-Prinzip (Glaubwürdigkeit von Signalen durch Fälschungssicherheit, siehe Zahavi & Zahavi 1997) nicht ausgeschlossen werden können (Hawkes & Bird 2002).

Damit sind die die theoretischen Hintergründe von Kooperation allgemein für die Biologie (Abschnitt A 1) und speziell für den Menschen, im besonderen die Charakteristika verschiedener sozialer Dilemmasituationen, dargestellt. Dabei wurden auch verschiedene, interdisziplinäre Zugänge und Analysemöglichkeiten skizziert. Der folgende Abschnitt A 3 gibt einen kurzen Überblick über diese Zugänge. Zusätzlich gewinne ich drei Argumente aus dieser Zusammenschau: erstens, das Verständnis für diese Problematik gewinnt durch unterschiedliche Zugänge. Zweitens, größere Datensätze (und Meta-Analysen) aus verschiedenen Kulturen und Populationen sind notwendig, um übergeordnete und stabile Muster zu finden. Drittens, ohne den Rückgriff auf unsere Vergangenheit als Jäger und Sammler sind diese Muster nicht zu verstehen. Alle drei Argumente werden für den Ergebnis- und Diskussionsteil D eine wichtige Rolle spielen.

3 Fazit für die Methodik

Die obige Darstellung (A 2) zeigt, wie vielschichtig sowohl das untersuchte Phänomen als auch die Zugänge mit ihrerseits diversen Methoden sind. Gerade weil meines Wissens nach zur Zeit keine Theorie verfügbar ist, welche die (Experimental)-Ergebnisse in ihrer Vielfalt zu bündeln vermag, will ich aus der bisherigen Befundlage heraus für die im vorigen Abschnitt genannten methodischen Vorsichtsmaßnahmen plädieren.

So kann, *erstens*, die Kenntnis der oben dargestellten verschiedenen Zugänge (etwa aus der Biologie, der Ökonomie und der Politikwissenschaft) im schlechtesten Fall Betriebsblindheit

verhindern, im besten Fall zur Verbesserung der eigenen Methodik oder Verbreiterung des Untersuchungsspektrums führen. So wird aus der Forschung zu Bakterien, Symbiosen und biologischen Märkten unmittelbar klar, dass institutionelles Design kein Allheilmittel sein und auch nicht die einzige Lösung darstellen kann, wie es etwa in der etwas einseitig geführten *Governance*-Debatte in der Politikwissenschaft behauptet wird (Gruber 2008).

Dies gilt auch für die verwendete *Methodik*: auch Feldstudien und Laborexperimente können sich gegenseitig befruchten – so zeigt Wiessner 2009, wie drastisch sich Verhalten im Experiment im Gegensatz zum Verhalten in der gewohnten sozialen Umgebung ändert. So zeigten Kung Buschleute in Diktator- und Ultimatumspielen ein sehr geiziges Verhalten und strafte nicht – mit dem Gewinn aus den Spielen gingen sie in ihrem Sozialleben allerdings großzügig um und verwendeten Strafen. Ein weiteres Bindeglied besteht darin, dass Menschen, die täglich mit Allmendeproblemen zu tun haben, ihre Einsichten aus Feldexperimenten auf ihre realen Probleme übertragen und profitieren können (Kimmich 2012). Ein Beispiel sind indische Bauern, die sich an Feldexperimenten zu wiederholten Gefangenendilemmata beteiligen und sich daraufhin in der realen Dilemmasituation kooperativer verhalten. Daraus können wiederum die Forscher lernen.

Weiterhin können Fallstudien generell aus *Experimenten* Nutzen ziehen, da deren Ergebnisse wichtig für die Hypothesengenerierung bei komplexen und unübersichtlichen Realweltfällen sind. So überträgt Rustagi et al. 2010 Ergebnisse zu verschiedenen Strategietypen im Labor auf Waldmanagement-Gruppen in Äthiopien und zeigt, dass Gruppen mit einem höheren Anteil an konditional Kooperierenden einen höheren Erfolg bei der Bewirtschaftung erzielen. Und schließlich werfen Befunde aus unserer evolutionären Vergangenheit ein oft völlig anderes Licht selbst auf jene Experimentalergebnisse, die vielfach repliziert und für robust gehalten wurden, etwa die Effizienz von Bestrafung (siehe Abschnitte D 2 und 3).

Ebenso wichtig sind, *zweitens*, kulturübergreifende Studien mit einer größeren Anzahl an Versuchspersonen. Die bahnbrechenden Studien von Henrich et al. 2010 und Henrich et al. 2001 belegen, dass die Übertragbarkeit zwischen Kulturen eben nicht ohne weiteres gegeben ist. Im Vergleich mit nicht-westlichen, nicht-studentischen Populationen, insbesondere auch Jägern und Sammlern zeigen sich deutliche Unterschiede etwa bei Kooperationsentscheidungen im Ultimatum- und Diktatorspiel (Abbildung 3, aus Henrich et al. 2010).

So bieten US-Amerikaner im Durchschnitt etwa 47% der zur Verfügung gestellten Summe dem Partner an, während dies bei den Hadza – Jägern und Sammlern aus Tansania – nur etwa

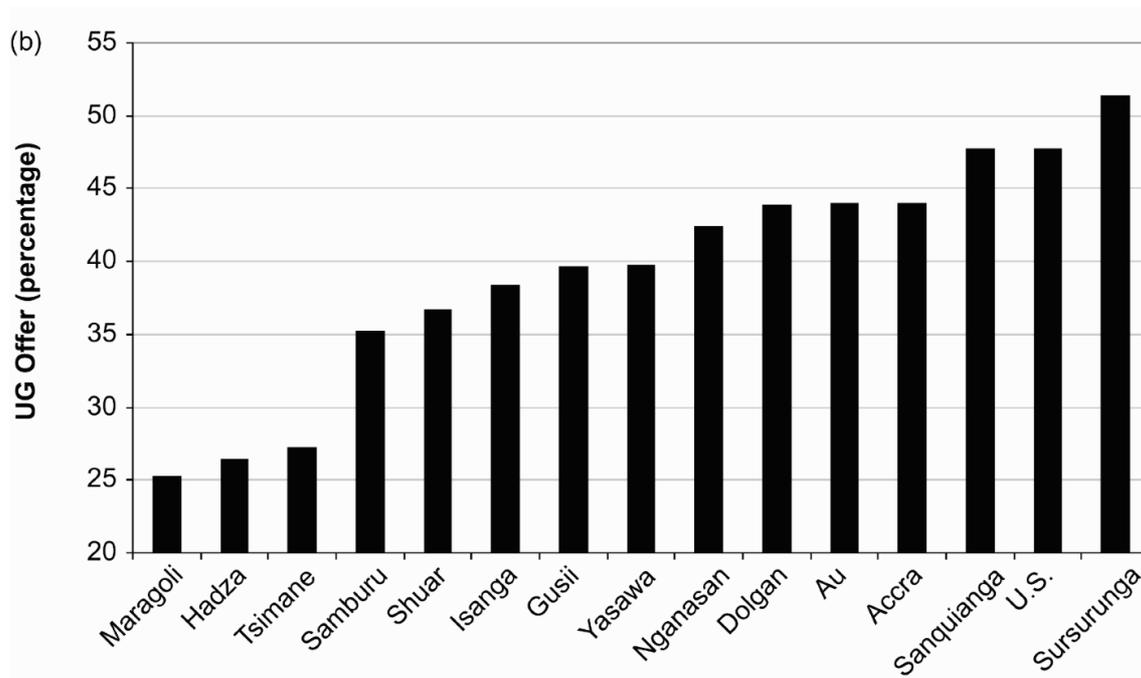


Abbildung 3: Durchschnittliches Angebot im Ultimatumspiel in 15 Kulturen in Prozent (Henrich et al. 2010)

26% der Summe sind. Allerdings wird oft der westliche Wert von etwa 50 % als universal angenommen. Welches Angebot dagegen gerade noch angenommen wird, variiert in dieser Studie ebenfalls stark – d.h. es gibt keine sehr gute Übereinstimmung zwischen dem Angebot, das man machen sollte, um das Einkommen zu maximieren, und der Summe, die gerade noch akzeptiert wird (*income-maximising offer*). Diese Diskrepanz zwischen Erwartungen und tatsächlichem Verhalten der Gegenperson wird in komplexer Form beispielsweise auch in *trust games* beobachtet (Fetchenhauer & Dunning 2009).

Diese Vielfalt bezüglich Strategien, die oft als elementar angesehen werden, zeigt sich auch in anderen Studien, etwa bei Bestrafung in PGG (Herrmann et al. 2008, siehe Abbildung 4 auf S. 49), wo ebenfalls große kulturelle Unterschiede deutlich werden. Auch individuelle Präferenzen bezüglich Kooperation (Kocher et al. 2008), also wie viele Individuen in ein Gemeingut investieren, unterscheiden sich zwischen Ländern scheinbar deutlich. Besonders problematisch ist dabei, dass 96% der publizierten psychologischen Studien aus westlichen Ländern kommen und davon wiederum 67-80% der Versuchsteilnehmer ausschließlich aus Psychologie-*undergraduates* bestehen. Daraus folgt, dass die kulturelle Übertragbarkeit nicht ohne weiteres gegeben ist und einige „robuste“ Experimentalergebnisse nur für die Gruppe westlicher Studenten gelten, die ein Ausreißer sind bezüglich einiger Aspekte von Kooperation (Henrich et al. 2010).

Wenn diese Belege und Argumente stimmen, müssten viele Experimente mit Beachtung von Kontexteffekten mit einer größeren Zahl an Versuchspersonen über mehrere Kulturkreise hinweg erneut durchgeführt werden, um deren Robustheit besser einschätzen zu können. PGG-Experimente müssten dieser Logik zufolge auch zwingend die Erkenntnisse zu Spielertypen (siehe Abschnitt D 3) berücksichtigen – indem man etwa vor dem Experiment die Spieler nach ihren Strategien trennt. Obwohl die verschiedenen Typen seit etwa 2001 gut bekannt sind, gibt es praktisch keine Experimente, die zwischen den Spielertypen unterscheiden. Eine Ausnahme sind hier Burlando & Guala 2005, die verschiedene Spielertypen identifizieren und mit ihnen homogene Gruppen bilden (d.h. beispielsweise Gruppen, die ausschließlich aus Trittbrettfahrern bestehen). Erst eine solche Unterscheidung und Unterteilung in verschiedene Gruppen, die identische Treatments durchlaufen müssten, macht eine Differenzierung der Ergebnisse möglich – und stellt wiederum vielleicht einige bislang für unumstößlich gehaltene Wahrheiten in Frage.

Diese Unsicherheit bezüglich der Übertragbarkeit führt zu einem weiteren *caveat*: erst eine erhebliche Anzahl von Versuchspersonen führt zu relativ robusten Ergebnissen. Ein Beispiel: Bestrafung gehört zu den bestuntersuchten Parametern in PGG überhaupt und wird oft mit Belohnung hinsichtlich Wirksamkeit verglichen (z. B. Andreoni et al. 2003; Sutter et al. 2008). Meist wird argumentiert, dass Bestrafung effektiver sei. Eine Metastudie, die 187 Studien zu Bestrafung und Belohnung in ihrer Effektstärke vergleicht, findet jedoch keinen signifikanten Unterschied (Balliet et al. 2011). Dies gilt ebenso für Fallstudien mit öffentlicher Gut-Problematik: hier existieren extrem viele Einzelstudien, die jedoch keine belastbaren und generalisierbaren Resultate liefern können, weil unterschiedliche Methodik und Variablenwahl die Vergleichbarkeit extrem erschwert (Poteete & Ostrom 2008). Erst die wenigen existierenden Meta-Analysen (z. B. Cinner et al. 2012; Gutiérrez et al. 2011), die viele Fälle des Managements natürlicher Ressourcen wie z. B. Fischerei vergleichen und auf Erfolgsfaktoren prüfen oder Studien mit großen Fallzahlen (Gibson et al. 2005; Frey & Rusch 2013) zeigen Muster auf, die als relativ unabhängig von lokalen Bedingungen gelten können. So zeigt sich etwa, dass Mitwirkungsmöglichkeiten und Führungsqualität in vielen Situationen von entscheidender Bedeutung für Gemeingutproblematiken bei natürlichen Ressourcen sind.

Und schließlich, drittens, ist es wichtig, unsere Vergangenheit als Jäger und Sammler zu berücksichtigen; dies gilt insbesondere für die Hypothesenbildung und Interpretation von Resultaten: So waren in der Psychologie über Jahrzehnte die Gründe für die anscheinend unsystematisch variierenden Ergebnisse des *Wason Selection Task* zur Falsifikation (Wason

1968) heftig umstritten. Versuchspersonen zeigen große Schwierigkeiten, Sachverhalte korrekt zu falsifizieren. Die zu lösende Aufgabe hat die Form des *Modus Tollens* (aus „Wenn A, dann B“ und „nicht B“ kann auf „nicht A“ geschlossen werden) und wird in abstrakter Formulierung von weniger als 5% der Versuchspersonen gelöst (Cosmides & Tooby 1992). In anderen Formulierungen sind die Ergebnisse zwar besser, aber folgen scheinbar keinem Muster, was zu teilweise absurden Erklärungsversuchen führte, etwa dass Personen, die Formulierungen mit Autos und Städten dann besser lösen, wenn sie Auto fahren (vgl. etwa Griggs & Cox 1982). Erst eine evolutionspsychologische Erklärung konnte die große Varianz der richtigen Antworten erklären und systematisieren (Cosmides & Tooby 1992).

Ein weiteres Beispiel für den Zugewinn an Verständnis, wenn man unsere evolutionäre Vergangenheit einbezieht, sind die starken *framing*-Effekte, denen Menschen unterliegen – je nach Kontext werden Aufgaben sehr schlecht bis sehr gut gelöst (Kühberger 1998) – auch das ein Hinweis auf stark kontextsensitive kognitive Module, die Anpassungen darstellen. So reagieren Menschen unterschiedlich auf Gewinn- (durch mehr Sport hat man eine bessere Kondition) bzw. Verlustrahmen (durch mehr Sport hat man weniger Herzinfarkte). Für weitere Beispiele siehe Frey (2007).

Das bedeutet auch, dass experimentelle Paradigmen (wie etwa das typischerweise zehn Runden dauernde PGG, das in hunderten von Varianten durchgeführt wird) eine zumindest lose Anbindung an unsere stark kontextsensitiven Anpassungen besitzen müssen. Ist dies nicht der Fall, greift vermutlich die sogenannte *mismatch theory* (vgl. Zefferman 2014), die postuliert, dass einige Anpassungen unserer evolutionären Vergangenheit heute ihre Funktion verloren haben. Verliert man also das *environment of evolutionary adaptedness* aus den Augen, laufen unsere Heuristiken buchstäblich ins Leere – mögliche Resultate sind in solchen Fällen irreführend oder falsch. Meiner Meinung nach sind einige angeblich rätselhafte Effekte aus der Ökonomie Beispiele für Heuristiken, die zwar evolutionär gut erklärbar sind und wohl auch Anpassungen darstellen, aber gleichzeitig in einer solchen künstlichen Umgebung ins Leere laufen (siehe Abschnitt E 2). Dazu gehört etwa der *Restart*-Effekt (Andreoni 1988), bei dem Versuchspersonen nach einigen Runden mitgeteilt bekommen, dass das Spiel erneut beginnt, und oft entsprechend ihrem Verhalten in der ersten Runde mit den Einzahlungen erneut nach oben gehen. Ein weiterer möglicher Kandidat ist die Bestrafung in der letzten Runde von PGG, die aus Sicht der *rational choice*-Theorie unverständlich bleibt, da in der letzten Runde keine Verhaltensänderung mehr bewirkt werden kann. Und schließlich findet man Unterschiede zwischen einer Versuchsanordnung, die zwischen *partner* und *stranger* unterscheidet (Fehr &

Fischbacher 2003), was ebenfalls für die Strategie irrelevant sein sollte.

Ich plädiere deshalb dafür, Laborexperimente (etwa PGG oder Ultimatumspiele) im Labor in Hinblick auf evolutionäre Hypothesen zu entwerfen. Einerseits würde sich das Design einiger Experimente vermutlich stark verändern müssen, andererseits könnten bisherige Resultate in Verbindung mit evolutionären Mechanismen und Erklärungen entlang proximaler sowie ultimativer Motivation zu produktiven Reinterpretationen führen.

Aus diesen Gründen trifft die folgende Arbeit diverse Vorsichtsmaßnahmen:

- die Daten stammen aus unterschiedlichen Ländern
(adressiert das Problem der mangelnden Robustheit)
- die Daten stammen aus verschiedenen Studien
(adressiert das Problem der mangelnden Robustheit)
- die Daten stammen aus einer heterogenen Stichprobe in Bezug auf Alter, Ausbildung, Land, usw. (adressiert das Problem der mangelnden Robustheit)
- die Daten stammen aus einer sehr großen Stichprobe
(adressiert das Problem der mangelnden Robustheit)
- die Daten stammen nicht ausschließlich aus dem Labor
(adressiert das Problem der Artifizialität)
- die Daten laufen über eine längere Zeitspanne
(adressiert das Problem der Artifizialität)
- die Ergebnisse werden mit evolutionären Erklärungen verknüpft
(adressiert das Problem der Artifizialität)

Aus den eben diskutierten Problematiken und den bisherigen Überlegungen werden in den folgenden Abschnitten die Arbeitshypothesen entwickelt.

4 Ziel und Fragestellung

Übergeordnetes Ziel der vorliegenden Arbeit ist es, besser zu verstehen, welche Umstände Kooperation bei Menschen beeinflussen und welche evolutionären Mechanismen (siehe Abschnitt A 1) hierbei eine Rolle spielen.

Dafür sind neben den bereits diskutierten Erklärungsrahmen weitere Synthesen theoretischer

Erklärungen verfügbar (z. B. Bergmüller et al. 2007; Bshary & Bergmüller 2008; Nowak 2006). *Beobachtungsgrundlage* ist die Tatsache, dass in beiden verwendeten Datensätzen (siehe Abschnitt C) – Versuchspersonen in *public goods games* im Labor sowie Spieler im Online-Browser-Spiel Ikariam – ein hohes Niveau an Bereitschaft zeigen, in das jeweilige Gemeingut zu investieren. Auffällig ist dabei, dass sich dieses hohe Kooperationsniveau bislang nur schwer unter einen einheitlichen theoretischen Rahmen fassen lässt (Fehr & Schmidt 1999; Bolton & Ockenfels 2000) und sich auch die Befunde, je nach den beeinflussenden Parametern zwar vielfältig darstellen, sich aber teilweise sogar widersprechen (wie sie etwa Balliet et al. 2011 aufzählen).

Die vorliegende Arbeit nimmt dabei drei eng umrissene und miteinander verbundene Probleme innerhalb dieser übergeordneten Fragestellung auf, die zentral für das Verständnis von Kooperation sind. Untersuchungsgegenstand ist jeweils die soziale Dilemmastruktur bei öffentlichen Gütern, d.h., wenn individuelle Nutzenmaximierung eine optimale Lösung für das Allgemeinwohl verhindert.

Erstens ist aus der experimentellen Ökonomik bekannt, dass es individuell verschiedene Kooperationsstrategien gibt. Lassen sie sich auch außerhalb des Laborkontextes bestätigen? Falls ja, welche Strategien im Kooperationsverhalten lassen sich nachweisen (Abschnitt D 1)? Und wie differenzieren sich diese unterschiedlichen Strategietypen aus? Ich untersuche dabei, welche Strategien überhaupt existieren, welche Eigenschaften sie aufweisen, wie robust sie gegenüber Kontextveränderungen sind und wie stark sie sich modifizieren lassen. In den Abschnitten D 3 und E wird darüber hinaus auch diskutiert, warum diese Strategien existieren (ihr proximaler Grund), wie sie miteinander wechselwirken und wie sich mit diesem Wissen möglicherweise Kooperation auf einem hohen Niveau stabilisieren lässt.

Zweitens wird untersucht, wie Menschen mit Trittbrettfahrern (Egoisten, die nichts oder sehr wenig in Gemeingüter investieren, sie aber dennoch nutzen) umgehen, einer evolutionär stabilen und ubiquitären Strategie, mit der sich kooperierende Strategien zwangsläufig auseinandersetzen müssen. Hier hat *Bestrafung* von nicht kooperierenden Individuen die stärkste Beachtung in der Literatur gefunden. Ich versuche, Strafen vor unserem evolutionären Hintergrund und hinsichtlich der Unterscheidung von Spielertypen zu verstehen (Abschnitt D 2).

Drittens greife ich eine unmittelbare Folgerung aus diesem Befund auf, die bislang jedoch wenig Beachtung in der Literatur gefunden hat: Wie schneiden Strategien auf dem Kontinuum von egoistisch bis hoch kooperativ im Vergleich ab (Abschnitt D 3)? Dies lässt Rückschlüsse

auf ihren evolutionären Erfolg zu. Der bisherige Stand des Wissens lässt erwarten, dass egoistische Strategien erfolgreicher sind als kooperative – daher untersucht diese Arbeit, ob dies wirklich der Fall ist. Erfolg beinhaltet dabei nicht nur die Häufigkeit ihres Auftretens im Vergleich zu anderen Strategien, sondern auch ihre Leistungsfähigkeit im direkten Vergleich.

Damit werden drei Forschungslücken adressiert:

Erstens beschäftigt sich ein großer Teil der Forschung zu Kooperationen (zumindest in der experimentellen Ökonomik) nur mit *externen* (d.h. von außen vorgegebenen Rahmenbedingungen oder Institutionen) Einflüssen auf das Kooperationsverhalten – Beispiele sind Bestrafung (Gürek et al. 2006; Fehr & Fischbacher 2003), Reputation (Milinski et al. 2002), Kommunikation (Ostrom et al. 1994) und weitere. Für einen Überblick siehe Chaudhuri 2011.

Ich bin hingegen der Meinung, dass *endogene* Faktoren eine ebenso große Rolle spielen, wie etwa die selbstständige Sortierung von Gruppen gemäß der Beiträge der jeweiligen Versuchspersonen zum öffentlichen Gut. Dies kann zu sehr hohen Beiträgen in altruistischen Gruppen und sehr niedrigen Beiträgen in Trittbrettfahrergruppen führen (Page et al. 2005; Burlando & Guala 2005).

Zweitens sehe ich eine Forschungslücke in einem Mangel an plausiblen evolutionären Erklärungen für viele Verhaltensexperimente bei Menschen. Diese Studie verbindet daher die Ergebnisse dieser Studien eng mit evolutionären Hypothesen (Abschnitte D 1.2.2, B 4.4 und E).

Drittens vermisste ich Analysen, die untersuchen, unter welchen Umständen Kooperierende erfolgreicher als Trittbrettfahrer sein können. Ich vermute, dass es typische Situationen gibt, in denen das der Fall sein könnte – wenn etwa Kooperierende nur untereinander interagieren und den vollen Nutzen aus erzeugten Gemeingütern ziehen können. Diese Arbeit trennt deshalb Spieler nach Strategien und Kontexten und analysiert, welche Parameter Erfolg und Misserfolg beeinflussen.

Die vorliegenden Analysen bedeuten deshalb einen Fortschritt gegenüber dem bisherigen Wissensstand. Derzeit gibt es keine Analyse von Daten zu *public goods games* (siehe C) in diesem Umfang (18 000 Personen), in dieser kulturellen Breite (fünf Länder) und über einen derart langen Zeitraum (zehn Monate). Diese sehr reichhaltigen Daten sind sowohl dafür geeignet, bisherige Befunde aus herkömmlichen Laborexperimenten zu bestätigen oder zu widerlegen – also die Robustheit zu testen –, als auch Detailanalysen durchzuführen (etwa

langfristige Veränderungen in der Kooperationsbereitschaft), die bisher nicht möglich waren. Anhand dieser Daten testet diese Arbeit auch die Behauptung, ob verschiedene Kooperationsstrategien existieren, d.h., ob es voneinander unterscheidbare Spielertypen (Egoisten, Reziprokatoren, Kooperierende) gibt. Diese Behauptung ist fundamental für viele Studien zur Reziprozität und für konditionale Kooperation.

B Stand der Forschung und Arbeitshypothesen

1 Allgemeine Arbeitshypothesen

Zunächst ist zu klären, welche evolutionären Erklärungsmodelle überhaupt in Anschlag gebracht werden können. Einige Erklärungsmodelle sind – wie eben diskutiert – beispielsweise nicht zufriedenstellend für die Erklärung in den nachfolgend untersuchten Kontexten, da sie bestimmte Voraussetzungen fordern, um zur Anwendbarkeit zu kommen (z. B. Verwandtschaft bei nepotistischem Altruismus, was bei Computerspielen nicht gegeben ist). Ich nehme diesen Gedanken in wenigen Absätzen mit der Diskussion der *mismatch theory* wieder auf.

In beiden untersuchten Datensätzen (Teil C) sind die Versuchspersonen nicht miteinander verwandt und wissen dies auch. Auch die hohe Zahl der beteiligten Personen und die Anonymität unterstützen dies. Damit ist eine Erklärung der Kooperationsbereitschaft über *nepotistischen Altruismus* unwahrscheinlich, wenn die Passung als gegeben angenommen wird.

Auch für eine Erklärung über den *Handicap-Altruismus* sind einige Grundprämissen für die Funktionalität nicht gegeben, etwa die zufällige Gruppenzusammensetzung, die räumliche Trennung und das Fehlen geeigneter anderer Kontexte (vor allem Sexualität), die durch teure Signale kommuniziert werden könnten (Uhl & Voland 2002).

Biologische Märkte sind ebenfalls problematisch für eine Erklärung, da zum einen der Markt fehlt, d.h. es keinerlei Tauschmöglichkeiten gibt, und zum anderen der persönliche Nutzen durch erworbene Güter oder Dienstleistungen besonders in der virtuellen Umgebung nicht unmittelbar sichtbar ist. Zumindest in PGG ist auch die Möglichkeit der Wahl des Interaktionspartners – ein unverzichtbarer Bestandteil für Märkte – nicht gegeben (Noë et al. 2001).

Da es durch die Einzahlungen in das Gemeingut zu altruistischen Vorleistungen kommt, liegt ein echtes soziales Dilemma vor. Damit ist es ebenfalls schwierig, *Mutualismus* als Erklärung in Stellung zu bringen.

Damit bleiben nach der Klassifikation von Voland (2013) nur noch *direkte und indirekte Reziprozität* übrig. Zunächst zur indirekten Reziprozität: Voraussetzung ist der Reputationsaufbau in sozialen Netzwerken, was in den Labor-PGG sowohl durch die strikte Anonymität als auch durch die engen Gruppengrenzen (3 bzw. 4 Spieler) und in manchen *Treatments* auch durch das *stranger-design* (einmalige Interaktion in einer Gruppe) unmöglich wird. Auch in Ikariam ist der Reputationsaufbau stark begrenzt, wenn auch nicht unmöglich: man interagiert nur mit wenigen Spielern (max. 20-30), kennt diese nur unter Pseudonym und kann die Reputation Fremder auch nicht prüfen, da diese Informationen auf neuen Inseln nur für deren Bewohner sichtbar sind. Trotz dieser Beschränkungen gibt es soziale Kontakte per e-mail, möglicherweise auch unter echtem Namen, d.h. diese Erklärung kommt durchaus in Frage.

Eine Erklärung über direkte Reziprozität würde implizieren, dass Spieler in das Gemeingut investieren, in der Erwartung, dass andere dies ebenfalls tun oder um andere Spieler dazu zu bringen, auch zu investieren. Die Situation und die Befunde aus den Labor-PGG eignen sich für Erklärungen über Reziprozität; dies wird auch durch Untersuchungen zu Spielertypen gestützt (siehe Abschnitt D 1 und 1.2), bei denen konditionale Kooperierende die Mehrheit stellen, also Personen, die reziprok agieren. Dies wird allerdings durch die Mehrheit der Aussagen in einer Umfrage für Ikariam stark relativiert. Hier *dominieren* Aussagen, die rein egoistische Überlegungen in den Vordergrund stellen. Ein typisches Beispiel ist die folgende Antwort auf die Frage „Warum spendest du mehr bzw. weniger als andere?“: „Höhere Minen bringen einen höheren Ertrag, desweiteren steige ich somit im Spendenhighscore.“ Daneben existieren allerdings auch Aussagen, bei denen Spieler auf den kollektiven Charakter abzielen, wie etwa „Es ist der Schlüssel zum Erfolg. Organisiert man die Spenden, wird man automatisch zum Topspieler, weil genug andere Spieler das nicht hinbekommen.“ Einige wenige Aussagen passen auch gut zu Erklärungen über Reziprozität, wie etwa „Ich will angemessen viel spenden, um keine Sozialsau zu sein.“ (Ikariam-Umfrage 2014, siehe Anhang G)

Zwei weitere Aspekte werfen ebenfalls Zweifel auf eine Erklärung über direkte Reziprozität. Zum einen die extreme Seltenheit dieses evolutionären Szenarios zumindest im Tierreich (Voland 2013; West et al. 2011) und zum anderen die problematische Anwendung bei Gemeingütern (Smith & Bliege Bird 2005). Ich komme darauf zurück.

Da sich Spieler in beiden untersuchten Datensätzen in der Kooperationsbereitschaft ähneln, was zudem in Einklang mit vielen weiteren Experimentalergebnissen steht (siehe etwa Chaudhuri 2011), liegt der Verdacht nahe, dass eine gemeinsame oder sich mindestens teilweise

überlappende Erklärung herangezogen werden sollte. Dabei können die Daten jeweils nur als *pars pro toto* gelten, da Fitnessbilanzierungen auf das gesamte Leben bezogen sein müssen, also einen weit längeren als den untersuchten Zeitraum umfassen (West et al. 2011). Allerdings geht diese Arbeit auf Grund der starken Konvergenz der Befunde zum menschlichen Kooperationsverhalten (starke anfängliche Kooperation, höhere Kooperation mit bekannten Partnern, häufiges Nutzen von Bestrafungsoptionen, stark positiver Einfluss von Kommunikation, usw.) davon aus, dass die menschlichen Kooperationsheuristiken in ihren Grundzügen nicht zu stark modifizierbar sind.

Damit haben sich alle oben beschriebenen Erklärungsansätze – bis auf direkte Reziprozität mit Einschränkungen – als in dieser Form ungeeignet erwiesen, zu erklären, warum Menschen in vielen Situationen so großzügig in öffentliche Güter investieren.

Drei andere Möglichkeiten existieren jedoch. Erstens könnte es sich einfach um eine „Fehlzündung“ (*mismatch hypothesis*) handeln, also um Verhalten, das in artifiziellen Laborumgebungen oder virtuellen Welten nicht die Bedingungen vorfindet, an die es angepasst ist (Hagen & Hammerstein 2006; Zefferman 2014; siehe auch Abschnitt D 1 für eine detaillierte Diskussion). Wichtig ist in diesem Zusammenhang auch eine Unterscheidung zwischen evolutionären Erklärungen und Erklärungen aus der *rational-choice*-Theorie: evolutionäre Erklärungen zielen auf *durchschnittliche* Fitnesskonsequenzen ab, d.h. auch scheinbar „irrationales“ Verhalten wie etwa Kooperation und Bestrafung in der letzten Runde von *public goods games* kann dadurch erklärt werden, dass es im Durchschnitt in einem normalen sozialen Kontext angepasstes Verhalten darstellt (West et al. 2011; Hagen & Hammerstein 2006).

Zweitens könnte die Erklärung über direkte Reziprozität durch *Zwang* (*enforcement*) gestützt werden. Hierfür existiert eine breite Unterstützung in der Literatur, die von Reputation (Milinski et al. 2002) über Bestrafung (Fehr & Fischbacher 2003) bis zu Partnerwahl sowie Partnerwechsel (*partner choice*, *partner switching*) reicht (Frank 2003; Bergmüller et al. 2007). In beiden Datensätzen ist Partnerwechsel jedoch entweder gar nicht oder nur schwer möglich.

Es ist klar, dass bei allen Zwangsmaßnahmen ein Gemeingutproblem zweiter Ordnung entsteht, wie es im lateinischen *Quis custodiet ipsos custodes?* (Wer aber bewacht die Wächter selbst?) zum Ausdruck kommt. Zwei mögliche Lösungen wurden dafür vorgeschlagen. Zum einen ist dies der höhere direkte Nutzen des Bestrafenden, etwa durch Verhaltensänderung des bestraften Individuums (Clutton-Brock & Parker 1995) und zum anderen die Ermöglichung

eines schnelleren Wechsels zu kooperativeren Individuen (Frank 2003).

Eine dritte Möglichkeit, die in dieser Arbeit genauer beleuchtet wird, besteht darin, dass die *Erzeugung von Gemeingütern* in einigen Fällen bereits auf der proximativen Ebene durchaus *erfolgreicher* sein kann, als dies nicht zu tun. Mit anderen Worten: Individuen, die in Gemeingüter investieren sind nicht altruistisch, da sie sich dadurch keine negativen Fitnesskonsequenzen einkaufen, sondern sie handeln durchaus im Selbstinteresse, da sie von den Vorteilen des Gemeinguts profitieren. Dies könnte sich, so die Idee, die in Abschnitt D 3 im Detail weiter ausgeführt und belegt wird, im direkten Vergleich zu Individuen, die keine Gemeingüter erzeugen, in einer positiveren Fitnessbilanzierung auswirken. Dies erfordert allerdings einige Qualifikationen. Erstens steckt dahinter die Annahme, dass dieses Verhalten, wie viele andere soziale Verhaltensweisen auch, einer frequenzabhängigen Selektion unterliegt (Volland 2013). Zweitens wird damit kein evolutionär stabiles Gleichgewicht (ESS) postuliert. Drittens muss die Möglichkeit der Invasion durch Trittbrettfahrer zumindest im Ansatz eingeschränkt sein.

Die Überlegung ist wie folgt: In einer ursprünglichen Population von Defektoren entwickeln sich – sei es durch nepotistischen Altruismus oder andere Mechanismen – einige Individuen, die kooperieren. Sie können kleine Inseln der Kooperation aufrechterhalten, indem sie innerhalb der Kooperierenden zusammenarbeiten und dadurch Fitnessvorteile gegenüber Defektoren genießen. Allerdings müssen sie Defektoren mit Defektion begegnen, ein Verhalten, das bei Menschen gut belegt ist (siehe etwa Marlowe et al. 2011; Axelrod 1984/2000). Die Erzeugung eines Allgemeinguts folgt einer ähnlichen Logik: kann die gesamte Gruppe davon profitieren, auch die Defektoren, dann sind die Kooperierenden im Durchschnitt in der Fitness unterlegen. Ist es jedoch möglich, den Zugang zum öffentlichen Gut nur einem Teil der Gruppe, nämlich denjenigen, die sich ebenfalls an der Erzeugung beteiligen, zu ermöglichen, dann würden die Effizienzgewinne durch das öffentliche Gut Fitnessvorteile für diejenigen Individuen mit dieser Strategie bedeuten.

Nun ist es bei einem öffentlichen Gut per Definition praktisch unmöglich, den Zugang zu verhindern. Allerdings ist selbst bei den klassischen Lehrbuchbeispielen (Deich, Leuchtturm, Straßenbeleuchtung) schnell offensichtlich, dass gerade Infrastruktur, gleich welcher Art, ortsgebunden und damit gruppenabhängig ist. Damit ist das Definitionskriterium ad absurdum geführt, da es zwar für jeden Nutzer *prinzipiell* möglich ist, von einem Deich zu profitieren, auch wenn er nichts investiert hat – *in der Praxis* profitiert aber nur eine begrenzte Gruppe.

Diese Überlegung gilt analog für Allmendegüter (*common pool resources*), bei denen sich zeigt, dass die die meisten natürlichen Ressourcen (z. B. Fische, Wälder) von Gruppen ausgebeutet werden, die erfolgreich darin sind, andere von der Nutzung abzuhalten (Ostrom 1990). Der entscheidende Unterschied bei Gemeingütern besteht demnach *nicht* in der Ausschließbarkeit, denn diese kann durch institutionelle Maßnahmen verändert werden. So kann z. B. ein Auto sowohl im Privatbesitz sein (Privatgut, also hohe Ausschließbarkeit), sich in einem *cars-haring-pool* befinden (Clubgut, also begrenzter Zugang) oder gar für die Allgemeinheit freigegeben sein (öffentliches Gut, keine Zugangsbeschränkung). Wichtiger ist das Kriterium der *Rivalität* – einerseits Ressourcen, die durch Verbrauch oder Nutzung weniger werden (alle natürlichen Ressourcen), andererseits Ressourcen, die durch das Teilen mehr werden (z. B. Wissen oder Information).

Wenn diese Überlegungen bis hierher richtig sind, dann muss man sich die Frage stellen, welche öffentlichen Güter bzw. Allmenderessourcen in unserer evolutionären Vergangenheit eine Rolle gespielt haben. Ein ohne Zweifel gut erforschtes Gemeingut ist die un konditionale Fleischteilung, die sich in vielen Jäger und Sammler-Gesellschaften findet. Ob sich hierfür eher Reziprozität als evolutionäres Szenario eignet, da auch erfolgreiche Jäger an den meisten Tagen keinen Jagderfolg haben (Hill & Hurtado 2009) oder die Befunde eher für teure Signale der erfolgreichen Jäger sprechen (Hawkes & Bird 2002, ist nicht abschließend geklärt. Weitere kooperative Aktivitäten bei Jägern und Sammlern (nicht alle folgenden Beispiele sind Gemeingüter) für Männer sind Infrastrukturmaßnahmen (Pfad anlegen), gemeinsame Arbeiten (Baumfällen), Arbeitsteilung (Reparieren von Werkzeugen) und Weitergabe von Informationen. Für Frauen sind es Arbeitsteilung (wechselseitiges Aufpassen auf Kinder), Weitergabe von Informationen sowie Lehren und gemeinsame Arbeiten (Hill 2001).

Es liegt nahe, dass einige dieser Gemeingüter, wie das Anlegen eines Pfades, kaum vor Trittbrettfahrern geschützt werden können. Für andere ist dies hingegen ohne Probleme möglich: ein Beispiel sind komplexere Arbeitsteilungen zwischen mehreren Personen, was, wenn es über längere Zeiträume funktioniert, durchaus als Institution und damit als Gemeingut angesehen werden kann.

Bei diesen Überlegungen zur Erzeugung von Gemeingütern spielt ein weiterer wichtiger Aspekt mit hinein – die Gruppengröße. Die mittlere Größe beträgt bei einer Stichprobe von 32 Jäger und Sammler-Gesellschaften 28 Individuen (Hill et al. 2011). Da Jäger und Sammler in kleinen Gruppen leben, die räumlich deutlich getrennt sind, könnte dies für die Erzeugung von

öffentlichen Gütern, zu denen nur die unmittelbare Gruppe Zugang hatte, eine Rolle gespielt haben. Denn in einem solchen Kontext ist es sehr leicht, Trittbrettfahrer von der Nutznießung auszuschließen, da größere räumliche Entfernungen die Regel waren. Damit ist eine wesentliche Voraussetzung für die obige Hypothese in unserer evolutionären Vergangenheit gegeben. In diesem Sinne ist es denkbar, dass es auf diese Weise zu Wettbewerbsvorteilen kooperativer Gruppen gekommen ist. Anders formuliert: Investition in ein Gemeingut kann sich durchaus als egoistische Strategie in einem proximatem Sinn erweisen.

Diese Überlegung muss, um Missverständnisse zu vermeiden, von einer Erklärung über Gruppenselektion abgegrenzt werden. Der Mechanismus, der hier zum Tragen kommen, bezeichnet man im weiteren Sinne als Gruppenkonkurrenz (Voland 2013; im engeren Sinne als *group augmentation* (wenn kooperatives Verhalten zur Vergrößerung der Gruppe führt, West et al. 2011). Die Funktionslogik ist allerdings eine völlige andere als bei Gruppenselektion, bei der ein genetisch altruistisches Verhalten eines Individuums zu Gunsten der Gruppe postuliert wird; bei Gruppenkonkurrenz führt das „altruistische“ Verhalten stattdessen in gen-egoistischer Weise zu Vorteilen der Gruppe gegenüber anderen Gruppen, von denen das Individuum selbst auch profitiert.

Die folgende Tabelle 1 gibt zusammenfassend eine Übersicht darüber, welches Erklärungsmodell für das beobachtete kooperative Verhalten passen könnte, mehrheitlich in ein Gemeingut zu investieren und erweitert die obige Diskussion. Die dritte Spalte nennt dabei sowohl gut als auch schlecht passende Attribute des jeweiligen Erklärungsmodells für Gemeinguterzeugung. Somit stellt Tabelle 1 das Fazit der bisher diskutierten Überlegungen zu evolutionären Mechanismen dar.

Erklärungsmodell	Kurzbeschreibung der Funktionsweise	Passung für Gemeinguterzeugung
Nepotismus	Verwandte profitieren von altruistischen Handlungen	Keine Verwandtschaftsbeziehungen; Anonymität
Biologischer Markt	Angebot und Nachfrage	Kein Markt vorhanden, keine Tauschmöglichkeiten
Direkte Reziprozität	Vorleistung und Erwidmung	Nicht notwendigerweise Produktion eines Gemeingutes; Instabilität;
Indirekte Reziprozität	Kooperation innerhalb eines Netzwerks	Anonymität und enge Gruppengrenzen in PGG und Ikariam
Mutualismus als Nebenprodukt	Nutzen für andere als Nebenprodukt	Kein Nebenprodukt, da es sich um zielgerichtete Kooperation handelt; Ausbeutung möglich
Mutualismus	Gegenseitiger Nutzen	Ausbeutung möglich
<i>Mismatch Hypothesis</i>	Heuristiken können Kooperation im Labor / online nicht von „normaler“ Situation unterscheiden	Verhalten wird in diversen, auch realistischen Kontexten beobachtet
<i>Greenbeard</i>	Gekoppelte Verbindung zwischen Verhalten im Genotyp und Aussehen im Phänotyp	Extrem selten; keine genetische Verwandtschaft
Teure Signale	Signalisierung verborgener Qualitäten	Zufällige Gruppenzusammensetzung; räumliche Trennung; kein Übertrag in anderen Kontext
<i>Kin discrimination</i>	Erkennen von Verwandten	Keine Verwandtschaft
<i>Limited dispersal</i>	Unkonditionale Kooperation mit räumlich nahen (und daher wahrscheinlich verwandten) Gruppenmitgliedern	Keine räumliche Nähe, keine Verwandtschaft
Gruppenselektion	Altruistische Gruppen haben Vorteile gegenüber egoistischen Gruppen	Aus theoretischer Sicht nicht haltbar
Starke Reziprozität	Prosoziale Präferenzen erfordern neuen proximalen evolutiven Mechanismus	Aus theoretischer Sicht nicht haltbar
Gruppenkonkurrenz	Erfolgreiche Gruppen verdrängen unerfolgreiche	Keine stabilen Gruppen
Direkte Fitnessvorteile durch erhöhte Effizienz der Gemeingüter	Gemeinguterzeugung ist trotz Kosten individuell unter bestimmten Umständen vorteilhafter	Erfordert teilweise räumliche Trennung von Trittbrettfahrern

Tabelle 1. Evolutionäre Erklärungsmodelle für die Gemeinguterzeugung

Zusammenfassend lässt sich sagen, dass keines der Theorieangebote a priori als Erklärung für die hohe Bereitschaft von Menschen, sich an Investitionen in Gemeingüter zu beteiligen, favorisiert werden muss. Im Gegenteil, kein Kandidat scheint recht geeignet zu sein, was die Möglichkeit offen lässt, ob es zu den postulierten direkten Vorteilen durch Gemeingüter kommt. Damit sind die allgemeinen Erklärungsmechanismen diskutiert. Ich wende mich daher den spezielleren Hypothesen zu und nehme die Diskussion der evolutionären Mechanismen in Teil E wieder auf.

2 Stand der Forschung zur Studie 1 (Spielertypen): Kooperationsverhalten in public goods games

Im vorigen Abschnitt (Teil B 1) sind mögliche Zugänge zu Forschungsfragen bezüglich menschlicher Kooperation und evolutionären Mechanismen diskutiert worden, die zur Erklärung herangezogen werden können. Man findet ein weites Feld an Ansätzen vor, die an ihre jeweilige Disziplin eng gebunden sind. Hier bedarf es der interdisziplinären Zusammenarbeit (siehe A 3). Im Folgenden konzentriert sich diese Arbeit auf den Stand der Forschung zu PGG im Labor, da dies die Methodik aller folgenden drei Studien zur Kooperation ist (Abschnitte D 1, 2 und 3). In den jeweiligen Abschnitten verenge ich die Diskussion der Literatur dann ein weiteres Mal auf die jeweiligen Themen „individuelle Strategien im Ländervergleich“ (dieser Abschnitt), „Bestrafung aus evolutionärer Perspektive“ (Abschnitt B 4) und „Erfolg hochkooperativer Strategien“ (Abschnitt B 6).

Menschliches Kooperationsverhalten wurde in den letzten Jahrzehnten mit verschiedenen Ansätzen untersucht. Das wohl am häufigsten eingesetzte Experimentaldesign sind Gemeingutspiele (*public good games*), wie sie etwa Fehr & Fischbacher (2003) analysieren. Dieses Design bildet das soziale Dilemma präzise ab, ohne zu komplex zu sein. Diese Arbeit beschränkt sich auf PGG, weil sie Realweltprobleme unter kontrollierten Bedingungen abbilden. Diese starke Ähnlichkeit mit der Realwelt ist wichtig, weil die ökologische Validität ein Problem darstellt. Die Literatur über PGG ist enorm; für Überblicksartikel siehe Zelmer 2003; Ledyard 1995; Chaudhuri 2011 und Gächter & Herrmann 2009; für Strafen im Besonderen Balliet et al. 2011.

Nach Jahrzehnten der Forschung und tausenden durchgeführten PGG haben sich einige robuste Befunde herauskristallisiert. Versuchspersonen geben weltweit – mit wenigen Ausnahmen – anfangs zwischen 40 und 60% ihres Geldes für das öffentliche Gut. Allerdings beobachtet man eine hohe Varianz zwischen 0 und 100% (Henrich et al. 2001). Dabei macht

es einen Unterschied, ob man in einer Gruppe spielt, die zusammenbleibt (Partner) oder die jede Runde wechselt (Stranger) (Fehr & Gächter 2000). Einsätze fallen im Verlauf auf etwa 10-20%. Nachdem viele Erklärungen dafür versucht wurden (z. B. Lerneffekte (Andreoni 1988)), ist man heute der Meinung, dass Menschen verschiedene Präferenzen haben (Kocher et al. 2008). Der Niedergang der Kooperationsrate ist nach dieser Hypothese das Resultat des Aufeinandertreffens von Trittbrettfahrern und konditionalen Kooperatoren, die sich je nach Investment der anderen mehr oder weniger kooperativ zeigen. Da die Trittbrettfahrer nichts beisteuern, geben die anfangs Kooperierenden ebenfalls weniger, was zu einer Abwärtsspirale führt.

Es gibt einige Mechanismen, die diesen Niedergang zuverlässig aufhalten können. Dazu zählt vor allem direkter Kontakt mit Gesprächsmöglichkeit (*face-to-face communication*). Einzahlungen steigen hier auf über 95% der geldlichen Ausstattung an (Bochet et al. 2006; Ostrom et al. 1994). Auch Bestrafung ist effektiv (Fehr & Gächter 2000; siehe auch D 2). Dazu gehört auch nicht-monetäre, soziale Bestrafung, die lediglich Missbilligung oder Lob zum Ausdruck bringt (Masclet et al. 2003). Reputation hat sich ebenfalls als zuverlässig erwiesen, um Kooperationsraten auf einem hohen Niveau zu halten (Milinski et al. 2006). Entscheidend ist zudem die Auszahlungsrate des öffentlichen Gutes (*marginal per capita return*, MPCR), die ja die paradoxe Wahl zwischen sozialem Optimum und egoistischer, rationaler Strategie determiniert. Deshalb gilt: Je höher der MPCR, desto höher die Einzahlungen in das öffentliche Gut (Isaac et al. 1985; Isaac & Walker 1988). Ein weiterer Faktor ist die Sortierung von Spielern (*assortment*). Teilt man Spieler je nach Höhe ihrer Einzahlungen in verschiedene Gruppen, so sind die Beiträge zum Gemeingut in Gruppen mit hohen Einzahlungen gleichbleibend hoch und in Trittbrettfahrergruppen nahe Null. Es zeigt sich demnach, wie erwartet, ein Gradient von sehr hohen bis hin zu sehr niedrigen Beiträgen je nach Gruppenzusammensetzung (Burlando & Guala 2005; Page et al. 2005). Dies ist gerade in Hinblick auf Spielertypen mit verschiedenen Strategien wichtig.

Einige Faktoren, die in der Literatur als Einflussfaktoren auf das Kooperationsniveau gehandelt wurden, haben sich hingegen als nicht relevant erwiesen bzw. zeigen nur kleine oder instabile Effekte. Dazu gehören z. B. Alter oder sozio-ökonomischer Status (Gächter et al. 2004), Geschlecht (Koopmans & Rebers 2009) oder Gruppengröße (Isaac et al. 1994).

Im Folgenden geht es genauer um die ökonomisch relativ junge Einsicht, dass Menschen individuelle Präferenzen bezüglich Kooperation haben (Fischbacher et al. 2001). Das heißt,

Menschen spielen verschiedene Strategien, ja können eventuell sogar Typen zugeordnet werden (unterschieden werden dabei Kooperierende (*cooperators*), Trittbrettfahrer (*free rider*), konditional Kooperierende (*conditional cooperators*) und Dreiecksformspieler (*triangle-shaped* oder *hump-shaped*)). Kooperierende sind dabei Spieler, die mehr als andere bzw. im Durchschnitt mehr als 50% ihres Einkommens in das öffentliche Gut einzahlen (eine genaue Diskussion der verschiedenen Definitionen findet sich in Abschnitt C 2.1). Trittbrettfahrer sind Spieler, die in jeder Runde 0 (oder sehr wenig) einzahlen und „Dreiecksform“-Spieler geben ansteigend mehr, je mehr andere einzahlen, allerdings nur bis zu einem gewissen Betrag. Ab dann verringern sie ihren Einsatz schrittweise. Konditional Kooperierende passen ihren Beitrag immer dem der Gruppe an.

Die hinter den Strategien stehende Motivation ist vermutlich unterschiedlich, wird allerdings meist nicht weiter thematisiert (z. B. Kocher et al. 2008; Fischbacher et al. 2001). Sie ist aber von einem evolutionären Standpunkt aus von hohem Interesse, weil diese proximativen Strategien verschiedene Ausdrucksweisen sind, Fitness zu maximieren. Auch ökonomisch kommt der Analyse von Strategien eine hohe Bedeutung zu, da die Existenz *verschiedener* Präferenzen nicht nur die Standardannahme des uniformen, nutzenmaximierenden *homo oeconomicus* zu Fall bringt, sondern auch möglicherweise das Grundmuster der Verelendung öffentlicher Güter (der berüchtigten *tragedy of the commons*) einer Erklärung nahe führen kann.

Ich führe den Überblick über die Forschung mit den Verhaltensunterschieden zwischen Kulturen fort. Zu den bereits genannten erheblichen Unterschieden zwischen den USA und marktfernen, nicht-industrialisierten Gesellschaften (Henrich et al. 2001) gesellen sich auch innereuropäische: eine Studie findet einen höheren Anteil an Trittbrettfahrern in England (60%) gegenüber Italien mit 40% (Burlando & Hey 1997). Ebenfalls geringere Beiträge zum öffentlichen Gut findet man bei Ostdeutschen im Vergleich zu Westdeutschen 5 bis 6 Jahre nach der Wende. Hier beginnen Westdeutsche mit 45%, Ostdeutsche mit 21% des Einsatzes, was auf 8% bzw. 1% nach 10 Runden absinkt (Ockenfels & Weimann 1999). Brandts et al. (2004) finden hingegen nur geringe Unterschiede im Verhalten in PGG zwischen Japan, den Niederlanden, Spanien und den USA.

Untersuchungen zu PGG deuten auf eine größere Varianz in verschiedenen Kulturen hin, wie bereits diskutiert (Brandts et al. 2004; Ockenfels & Weimann 1999; Burlando & Hey 1997). Größere Unterschiede werden auch bei Gruppen gefunden, die weder westlich noch studentisch geprägt sind (Buchan et al. 2009; Henrich et al. 2001). Der Befund scheint auch auf

konditionale Kooperation sowie auf Trittbrettfahrerverhalten zuzutreffen (Kocher et al. 2008). Im Widerspruch dazu finden Herrmann & Thöni (2009) stabile konditionale Kooperation bei unterschiedlichen Anteilen von Trittbrettfahrern.

Neben diesen kulturellen Unterschieden scheinen zwei Kennzeichen des Kontextes, in dem sich menschliche Kooperation entwickelt hat, von besonderer Bedeutung für eine stabile Kooperation sowie für die Entwicklung verschiedener kooperativer Strategien zu sein. Erstens: Der Situationsrahmen (einschließlich gemeinschaftlicher Aktionen wie Jagd sowie Arbeitsteilung oder die Verteilung größerer Jagdbeute) war auf *länger dauernde* Interaktionen ausgerichtet, so dass der Gewinn durch Betrug der Partner gering ausfällt und somit stabile Kooperation begünstigt (Frey & Rusch 2012; Gächter et al. 2008). Zweitens: Es standen typischerweise nur wenige hundert soziale Beziehungen zur Verfügung, was offenbar eine Mischung aus nepotistischen und reziproken Beziehungen begünstigte (Hill 2001).

Auch die Trennvariable „Kollektivismus“ spielt eine Rolle, da chinesische, japanische, koreanische und US-amerikanische Studenten signifikant unterschiedliches Verhalten bezüglich Vertrauen und Reziprozitätsverhalten in *trust*-Spielen an den Tag legen (Buchan et al. 2002). In einer weiteren Studie dieser Arbeitsgruppe wird dagegen das Globalisierungsniveau zur Trennung von Gruppen (hier für USA, Italien, Russland, Argentinien, Südafrika und Iran) benutzt (Buchan et al. 2009). Dieser Faktor spielt hier ebenfalls eine signifikante Rolle: Personen aus den USA (höchster Indexwert) zahlen mit einer 77% Wahrscheinlichkeit in das öffentliche Gut ein, Personen aus dem Iran (niedrigster Indexwert) mit der geringsten Wahrscheinlichkeit von 17%. Dieses Resultat ist unmittelbar anschlussfähig an die Variable „Marktintegration“, die auch in der Studie von Henrich et al. (2001) eine bedeutende kulturübergreifende positive Rolle für die Kooperation spielt. Damit wird die Vertrautheit von Personen mit Handel gemessen, also auch der Bereitschaft, Fremden zu Zwecken der Zusammenarbeit zu vertrauen.

Insgesamt findet man viele die Kooperationsbereitschaft beeinflussende Parameter. Darunter sind exogene wie etwa Kommunikation; kulturelle, wie etwa Marktintegration; und in der Persönlichkeit verankerte, wie etwa Spielertypen.

3 Arbeitshypothesen zur Studie 1 (Spielertypen)

Ein wichtiger Vorschlag zur Beantwortung der Frage, wovon die Kooperationsbereitschaft wesentlich abhängt, weist darauf hin, dass Interaktionen zwischen altruistischen und reziprok agierenden Individuen auf lange Sicht gesehen relativ stabil und profitabel bleiben können,

wohingegen Trittbrettfahrer diese stabilen Kooperationen schädigen, indem sie die Bereitschaft zum kollektiven Handeln bei konditional Kooperierenden zerstören (Burlando & Guala 2005; Gächter 2007). Das impliziert die Annahme, dass Individuen über verschiedene Strategien wie Kooperation und Defektion verfügen oder die Möglichkeit besitzen, zwischen diesen zu wechseln. Ob und wie Individuen ihre Strategien tatsächlich wechseln, ist nicht geklärt. Es gibt jedoch deutliche Hinweise auf unterschiedliche *Typen* von Individuen (Fischbacher & Gächter 2010).

Problematisch sind dabei die vielen unterschiedlichen Definitionen; je nach Definition wird ein Spieler als anderer Spielertyp klassifiziert. In einigen Studien werden beispielsweise Trittbrettfahrer nur dann als solche eingestuft, wenn sie immer 0 einzahlen (z. B. Fischbacher et al. 2001), in anderen Studien ist ein Trittbrettfahrer jemand, der immer weniger als 50% seines Einkommens gibt (Kurzban & Houser 2005). Darunter leidet die Vergleichbarkeit enorm.

Trotz dieser technischen Schwierigkeiten rücken für das Verständnis der Stabilität menschlicher Kooperation damit die unterschiedlichen Strategien in den Mittelpunkt des Interesses. Daher ist es notwendig, festzustellen, welche Strategien es gibt und wie stabil solche Spielertypen in verschiedenen Umfeldern sind. Das schließt die Abhängigkeit von sozioökonomischen Variablen, den spezifischen Parametern des Umfelds sowie von Persönlichkeitsmerkmalen mit ein.

Besonders interessant sind in diesem Zusammenhang deshalb Befunde zu Länderunterschieden bei Spielertypen. Hier gibt es unseres Wissens nur zwei Studien (Kocher et al. 2008; Herrmann & Thöni 2009). Beide Studien finden signifikante Unterschiede in der Häufigkeit des Auftretens der Spielertypen im Vergleich zu Fischbacher et al. 2001, der den Anteil konditional Kooperierender mit 50% und den der Trittbrettfahrer mit 30% angibt. Kocher et al. (2008) berichten dagegen 81%, 44% und 42% für konditionale Kooperierende sowie 8%, 22% und 36% für Trittbrettfahrer jeweils in den USA, Österreich und Japan. Herrmann & Thöni 2009 finden für Russland zwar vergleichbare 56% an konditionaler Kooperation, aber nur 6% Trittbrettfahrer. Eine Übersicht über die Ergebnisse der verschiedenen Studien gibt Tabelle 10.

Unterschiedlich große Anteile der vertretenen Typen in einzelnen Gruppen könnten beispielsweise erklären, warum sich die beobachteten Kooperationsniveaus zwischen verschiedenen Gruppen so stark unterscheiden. Schließlich könnten uns Erkenntnisse aus diesem Bereich auch ermöglichen, Kooperationsraten gezielt zu steigern und optimal angepasste Institutionen

zu entwickeln (Güerck et al. 2006). Untersuchungen belegen, dass Mechanismen, die Trittbrettfahrer von anderen Spielern trennen können, zu einer deutlichen und signifikanten Beitragssteigerung führen (Page et al. 2005). Die genannten Unterschiede der bisherigen Studien führen zu den spezifischen Arbeitshypothesen zu möglichen Spielertypen.

Hypothese 1: Die in ökonomischen Experimentallaboren nachgewiesenen Spielertypen existieren auch außerhalb dieser artifiziellen Laborbedingungen.

Aufgrund der hohen Varianz menschlicher Kooperationsentscheidungen (Abschnitte A 3 und B 2) und der bekannten Kontextabhängigkeit (Cosmides 1989), ist ein Übertragbarkeit aus dem Labor nicht ohne weiteres gegeben. Dafür ist es nötig, einen Kontext zu finden, der zwar vergleichbar ist, also dieselbe Dilemmastruktur aufweist, aber z. B. in Motivation, thematischer Einkleidung und den Anreizen Unterschiede zeigt. Diese Bedingungen sind im Online-Spiel Ikariam erfüllt (siehe Teile C und D).

Hypothese 2: Spielertypen bleiben über unterschiedliche Studien, Versuchsbedingungen, Definitionen und Länder in ihren Häufigkeiten relativ stabil.

Man kann nur dann sinnvoll von Spielertypen sprechen, wenn eine gewisse Stabilität und Unabhängigkeit von externen Faktoren gegeben ist. Zudem gehe ich davon aus, dass viele Aspekte von Kooperation von entscheidender Wichtigkeit im menschlichen Evolutionsprozess sind. Daraus folgt, dass unsere Heuristiken für Kooperationsentscheidungen Anpassungen darstellen. Damit lässt sich erwarten, dass diese Heuristiken robust gegen einzelne Änderungen in Versuchsbedingungen und kulturellen Unterschieden sind. Auf der anderen Seite legt die Modularitätsthese nahe (Barkow et al. 1992), dass solche Heuristiken durchaus nur innerhalb gewisser spezifischer Problemlagen funktionieren. Die Tests unter verschiedenen Bedingungen sollen zeigen, welche Parameter für die Stabilität der Heuristik eine Rolle spielen.

4 Stand der Forschung zur Studie 2 (Effiziente Bestrafung) – Sanktionen in public goods games, in sozial-ökologischen Systemen und bei Jägern und Sammlern

Das Thema Bestrafung wird in vielen Disziplinen intensiv erforscht, und spielt bei Gemeingutproblemen eine zentrale Rolle. Vor allem innerhalb der PGG-Literatur sind Studien zur Bestrafung kaum noch zu überblicken. Ich fasse die wichtigsten Ergebnisse zusammen. Zunächst ist festzuhalten, dass die theoretische Erwartung zwar darin besteht, dass niemand bestraft, da es mit Kosten verbunden ist (Fehr & Fischbacher 2003; Voland 2013), dies doch

ein substantieller Anteil an Versuchspersonen tut.

Stehen Menschen vor der freien Wahl, sich für eine Institution mit oder ohne Strafen zu entscheiden, so wählen zunächst ca. 70% keine Bestrafung. Nach kurzer Zeit (20 Runden) hat sich das Bild umgedreht – jetzt verbleiben nur noch ca. 5% in dieser Institution (Güerck et al. 2006). Dabei ist wechselnden Spielern klar, dass Trittbrettfahrer nicht weit kommen werden und erhöhen ihre Einsätze in das Gemeingut drastisch.

Nicht überraschend gibt es eine Korrelation zwischen der *Härte und Häufigkeit der Bestrafung* und der *Höhe der Normabweichung* in den Einzahlungen (Egas & Riedl 2008; siehe auch Abbildung 4, grüne Balken). Je stärker ein anderer Spieler von der Investition des Bestrafers abweicht – im negativen Sinne, also weniger einzahlt (grüne Balken) – desto härter wird er bestraft. Allerdings gibt es auch zwei weitere Arten von Bestrafung: Gegenbestrafung (Nikiforakis 2008) und anti-soziale Bestrafung (*anti-social punishment*). Bei letzterer wird gerade nicht nach der sozialen Abweichung bestraft (rote Balken in Abbildung 4, Herrmann et al. 2008). Anti-sozial wird sie deshalb genannt, weil der Bestrafte mehr als der Strafende einzahlt, also mehr für die Gemeinschaft tut. Die Höhe der (positiven) Abweichung sieht man in den roten Balken.

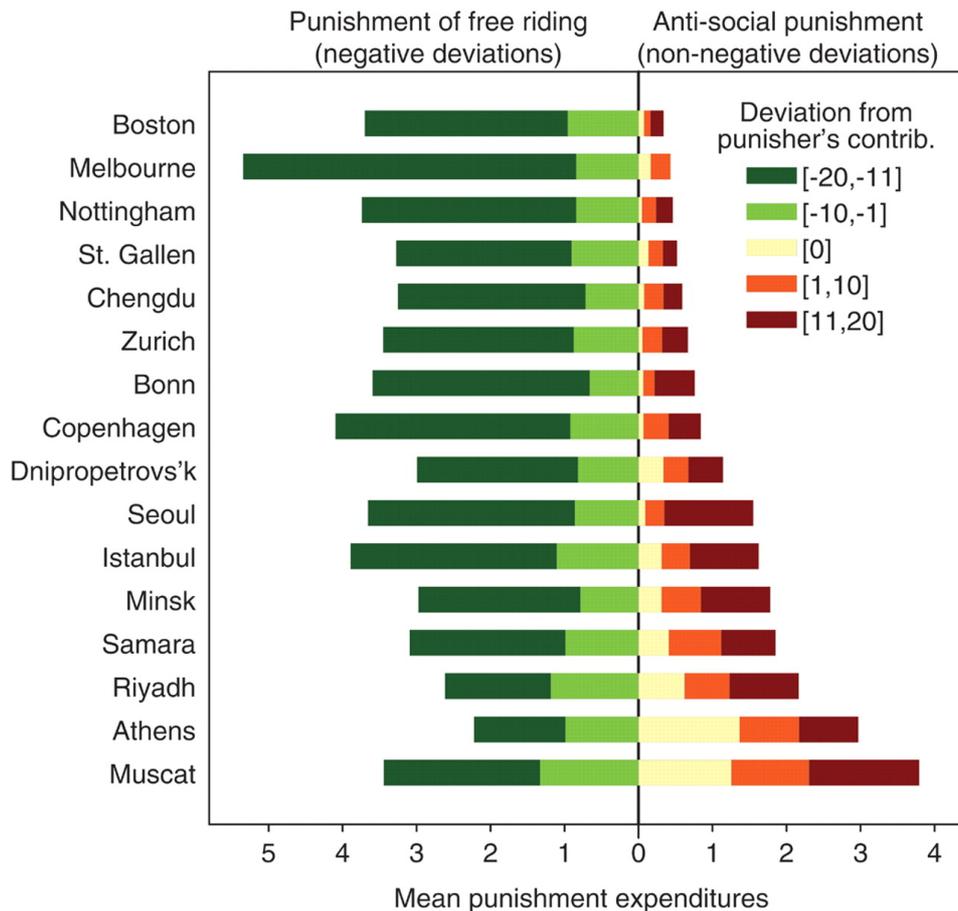


Abbildung 4. Mittlere Höhe der Bestrafung in PGG in 16 Städten (Herrmann et al. 2008)

Nimmt man – aus der Anwendungsperspektive – die oben genannten (B 2) kooperationsfördernden Maßnahmen und Bestrafung in den Blick, stellt man fest, dass die Maßnahme Bestrafung gleich aus mehreren Gründen nicht optimal ist. Zum einen muss man auch unproduktive Strafen in Kauf nehmen (*counter-* und *anti-social punishment*, vgl. Abbildung 4), was negative Emotionen weckt, zudem braucht man eine starke Institution, um Strafen wirkungsvoll durchführen zu können. Das ist gerade für globale Fragen kein triviales Problem: selbst wenn sich viele Staaten auf globale CO₂-Minderungen einigen könnten, welche Institution könnte neutral und glaubhaft überprüfen, ob sich ein Staat daran hält, und Sanktionen durchsetzen, die akzeptiert würden?

Zum anderen kommt eine Reihe von Studien zu dem Schluss, dass im Vergleich zu einem Standard-PGG ohne Strafen die Einzahlungen in das öffentliche Gut zwar bei Einführung von Strafen stark steigen, die tatsächliche soziale Effizienz (also, was Spieler als Auszahlung erhalten) aber niedriger ist (z. B. Fehr & Gächter 2002). Im Vergleich dazu scheinen Maßnahmen wie Kommunikation, Einsatz von Reputation oder die Trennung von Teilnehmern

in realen Situationen effizienter und leichter zu implementieren zu sein.

4.1 Sanktionen als Mittel, um Kooperation zu erhöhen

In Hinblick auf eine rationale Nutzenmaximierung ist eine theoretische Erklärung für Strafen nicht sofort einleuchtend. Zumindest für nicht-wiederholte Interaktionen sollte es weder Kooperation noch Bestrafung geben, da in beiden Entscheidungssituationen ein Gefangenendilemma entsteht, bei dem die dominante Strategie Defektion ist. Unter bestimmten Randbedingungen (*folk-theorem*), unter anderem Wiederholung, können Kooperation und Bestrafung jedoch nutzenmaximierend sein (Binmore 2006). Die Frage ist also, warum in PGG und ähnlichen sozialen Dilemmasituationen bestraft wird, wenn es doch sowohl kostspielig für den Strafenden als auch den Bestraften ist und damit die Effizienz der Gruppe erniedrigt wird. Mögliche evolutionäre Strategien, um mit Trittbrettfahrern umzugehen, sind neben Strafen (Clutton-Brock & Parker 1995) auch die Beendigung der Interaktion (Gurven 2004).

Bestrafung wird mit verschiedenen Bedeutungen in der Literatur verwendet (vgl. Frey & Rusch 2012). Diese Arbeit verwendet es im Sinne von „kostspieliger Bestrafung“ (*costly punishment*), d.h. in dieser Verwendung sind Nutzensvorteile für den Bestrafenden impliziert, selbst wenn man die Kosten einrechnet; in diesem Sinne entspricht es einer egoistischen Bestrafung (*selfish punishment*). Damit ist der Alternativbegriff, nämlich altruistisches Strafen, ausgeschlossen. Es wird also angenommen, dass kostspielige Bestrafung einen ultimativen Nutzen für den Strafenden hat, nämlich erhöhte Gesamtfitness. Sanktionen sind in diesem Sinne ein Zwangsmechanismus, um Trittbrettfahrer davon abzuhalten, die Belohnung für ihre Strategie einzufahren (vgl. Frey & Rusch 2012).

4.2 Effizienz von Sanktionen

Die Fragestellung, ob Sanktionen in der Gesamtschau effizienter als Institutionen ohne Sanktionen sind, ist noch ungeklärt (vgl. Fehr & Gächter 2002). Frey & Rusch (2012) weisen darauf hin, dass viele PGG mit 10 Runden nur eine sehr kurze Zeitspanne in den Blick bekommen, was die Ergebnisse wohl verzerrt. Zudem sind die Ausgaben für Strafen anfangs extrem hoch, auch wenn sie dann dramatisch abfallen. In dieser Studie zeigt ein Vergleich bisheriger Analysen zur Effizienz von Strafen, dass die Mehrheit der Studien Bestrafung als ein notwendiges Übel sehen, um Einzahlungen stabil zu halten. Da die Wohlfahrt, d.h. die Effizienz der Gruppe zerstört wird, werden Versuchsbedingungen mit Sanktionen (P-Treatments) im Vergleich zu Bedingungen ohne (NP-Treatments) generell als unterlegen angesehen. Betrachtet man allerdings den späteren Verlauf, so ändert sich das Bild und Treatments mit Bestrafung werden effizienter als nicht-bestrafende Treatments (vgl. Frey & Rusch 2012).

4.3 Sanktionen in sozial-ökologischen Systemen

Eine oft geäußerte Kritik an Laborexperimenten ist die mangelnde ökologische Validität (siehe auch Abschnitt A 3). Ein weiterer Kritikpunkt betrifft die Auswahl der Versuchspersonen. Mehr als 96 % aller Versuchspersonen in führenden Psychologie-Fachzeitschriften kommen aus westlichen, industrialisierten Ländern und sind Studienanfänger aus den USA. Gerade diese Gruppe könnte in ihrem kooperativen Verhalten ein Sonderfall sein (Henrich et al. 2010). Dadurch gewinnen Experimente oder Studien mit realweltlichem Bezug an Attraktivität.

So kann an Hand von Allmendegütern (*common pool resources*), wie etwa Bewässerungssystemen, Waldbewirtschaftung oder kleinen Fischereien aufgezeigt werden, dass Überwachung und Sanktionen Kernelemente des sozialen und ökologischen Erfolgs dieser kooperativen Gruppen ausmachen (Ostrom 1990; Chhatre & Agrawal 2008). Die Existenz von Strafen scheint ein notwendiges Kriterium für das erfolgreiche Management von Allmendegütern zu sein (Gibson et al. 2005). Dies geschieht bei erfolgreichen Institutionen mit minimalen Ausgaben für Sanktionen – diese werden den örtlichen Gegebenheiten angepasst (Ostrom 1990). Solide angepasste Regeln entwickeln sich allerdings weder schnell noch automatisch oder einfach. In den meisten Systemen dauert es Jahre oder sogar Jahrzehnte, um die Regelsysteme so zu verfeinern, dass sowohl die Kosten minimal sind als auch die Fairness gewahrt bleibt. Hinreichend häufig gelingt es Gemeinschaften nicht, angepasste Regelsysteme mit geringen Transaktionskosten zu entwickeln. Diese Befunde stützen die Behauptung, dass Sanktionen ein Kernelement für menschliche Kooperation und soziale Koordination sind (vgl.

Frey & Rusch 2012). Wie oben diskutiert (Abschnitt A 3) ist es wichtig, diese Befunde mit unserer evolutionären Vergangenheit als Jäger und Sammler in Einklang zu bringen.

4.4 Sanktionen bei Jägern und Sammlern

Menschen lebten Millionen von Jahren als Jäger und Sammler in kleinen Gruppen, in denen jeder jeden kannte (Hill 2001). Frey & Rusch (2012) argumentieren, dass eine solche Umgebung zwei strukturelle Eigenschaften besitzt: Erstens waren soziale Interaktionen häufig, sie wiederholten sich und sie waren langfristig angelegt und dauerten oft Jahre oder Jahrzehnte. Sowohl Verwandte als auch Nicht-Verwandte waren daran beteiligt. Zweitens favorisierte die Auszahlungsstruktur der meisten Interaktionen Kooperation, da langfristige Kooperationen im Vergleich zu kurzfristiger Defektion weitaus höheren Nutzen erzielen. Ein einfaches Beispiel ist die klassische Arbeitsteilung, bei der ein Individuum Kinder beaufsichtigt, während ein anderes Nahrung sucht. In diesem Beispiel bringt das Verlassen des Kindes zwar einen kurzfristigen kleinen Gewinn für ein Individuum (Zeit, in der etwas anderes gemacht werden kann), beendet aber gleichzeitig eine weitaus profitablere, weil potenziell langfristige Partnerschaft zu gegenseitigem Nutzen.

Aus diesen beiden strukturellen Eigenschaften schließen Frey & Rusch (2012), dass langfristige Allianzen und Partnerschaften substanzielle Vorteile für kooperierende Personen hatten, weil Bestrafung selten und billig war. Obwohl das Problem von Trittbrettfahrern existierte, konnten diese in Kontexten mit diesen Eigenschaften nur selten kleine Vorteile heraus schlagen. Sobald sie dabei entdeckt wurden, wurden sie öffentlich als Trittbrettfahrer angeprangert (Guala 2012). Bei Beibehaltung dieser Strategie verloren sie schrittweise den Zugang zu sozialen Netzwerken, in denen Nahrung und anderes geteilt wurde, bis sie bei Fortführung ihres Verhaltens schließlich in extremen Fällen aus der Gemeinschaft ausgeschlossen wurden. Anfängliche, kleine Vorteile durch Defektion sind mit einer hohen Wahrscheinlichkeit behaftet, zu großen Verlusten zu werden, vor allem im Vergleich mit langfristiger Kooperationen, die sichere und gegenseitige Gewinne verspricht (vgl. Frey & Rusch 2012).

Dafür spricht, dass der Zeithorizont sich über Jahre bzw. Jahrzehnte erstreckt – was etwa das Erlernen von Jagdfähigkeiten betrifft (Hill 2001). Darüber hinaus ist etwa das Feld der Nahrungsteilung von Überlebensrelevanz, da selbst erfahrene Jäger nur an 10-50 % aller Tage im Mittel erfolgreich sind (Hill & Hurtado 2009). Soziale Netzwerke lösen das Problem des nicht-vorhersagbaren Jagderfolges, beruhen jedoch auf Reziprozität, was zwar im Prinzip ausbeutbar ist, bei kleinen bekannten Gruppen mit wiederholten Interaktionen jedoch kaum

möglich ist. Viel effizienter scheint es zu sein, mit Reputation und Strafen zu arbeiten und unprofitable Partnerschaften schnell wieder zu lösen (Semmann et al. 2003; Aktipis 2011). Der Ausschluss von Defektierenden ist tatsächlich ein probates Mittel bei Jägern und Sammlern (Gurven 2004; Guala 2012).

Das eben geschilderte Argument, dass den Nutzen langfristiger Kooperation um ein Vielfaches höher als den Nutzen aus kurzfristiger Defektion einschätzt, stellt die Basis für die folgenden spezifischen Arbeitshypothesen (Abschnitt B 5) dar. In diesen Hypothesen geht es um die Zusammenhänge von Bestrafung und Spielertypen sowohl in realen Allmendesystemen (Ostrom 1990) als auch Daten zu PGG im Labor (Janssen et al. 2010), bei denen Bestrafung vor allem zu Beginn beobachtet wird und dann als bloße, kaum Kosten verursachende Drohung ausreicht, um Trittbrettfahrer zur Kooperation zu bewegen – auch wenn sie von Zeit zu Zeit in die Tat umgesetzt werden muss. Die Vermutung ist zudem (Frey & Rusch 2012), dass 10 Runden in PGG nicht ausreichen, um die Effizienz dieser Institution in den Blick zu bekommen. Zudem wird in dieser Studie darauf hingewiesen, dass es zunächst gegen jegliche Intuition zu sein scheint, dass Gruppen mit Strafen sich als effizienter als NP erweisen sollten, da erstere stets durch ihre Kosten belastet sind. Der Hinweis auf die Belastung ist zwar richtig, doch gibt es genau eine Ausnahme: *wenn Strafen bei geringen Kosten zu hohen und stabilen Beiträgen führen*, können Treatments mit Strafen zu einer höheren Effizienz führen. Beide Voraussetzungen sind in der Realität relativ häufig gegeben.

5 Arbeitshypothesen zur Studie 2 (Effiziente Bestrafung)

Die folgenden Arbeitshypothesen bauen auf den eben genannten Vorarbeiten (Frey & Rusch 2012) und der reichhaltigen Literatur zur Bestrafung in PGG auf (z. B. Herrmann et al. 2008; Nikiforakis 2008; Fehr & Gächter 2002). *Allerdings differenziert keine mir bekannte Studie ihre Erkenntnisse nach Spielertypen*. Die folgenden Arbeitshypothesen nehmen deshalb Befunde aus der Literatur zu Sanktionen auf und differenzieren diese allgemeinen Hypothesen nach Spielertypen. So ist zwar klar und gut bestätigt, dass generell dann stärker bestraft wird, wenn weniger in das öffentliche Gut investiert wird (Herrmann et al. 2008), es fehlen jedoch detaillierte Analysen, wer diese Strafen wann vornimmt und welche Strategie – in Hinblick auf Höhe der Investitionen und Höhe der Strafen – am erfolgreichsten ist. Aus diesen Befunden (Herrmann et al. 2008) werden die folgenden ersten beiden Arbeitshypothesen abgeleitet:

Hypothese 1a: Kooperativere Spielertypen strafen häufiger und mehr als egoistische.

Hypothese 1b: Kooperativere Spielertypen werden seltener und weniger als egoistische bestraft.

Für die Hypothese 1a spricht die Vermutung, dass kooperationswillige Individuen über Strafen versuchen, das Niveau der Investitionen in das öffentliche Gut in der Gruppe hoch zu halten. Analog dazu ist es wahrscheinlich, dass Trittbrettfahrer, da sie nicht bereit sind, Beiträge für das öffentliche Gut zu leisten, auch keine Strafkosten für die Gruppe übernehmen.

Für die Hypothese 1b gibt es in der Literatur robuste Befunde, dass geringe Investitionen eines Individuums ein Motiv für Bestrafung sind. Daher liegt es nahe, dass kooperative Spielertypen weniger bestraft werden als egoistische. Möglicherweise kommt es jedoch zu anti-sozialer Bestrafung und Gegenbestrafung (Nikiforakis 2008; Herrmann et al. 2008). Diese wird sich gemäß Definition vor allem gegen kooperativere Spielertypen wenden, die ja voraussichtlich mehr bestrafen als Trittbrettfahrer. Diese Differenzierung kann nun über die Klassifizierung der Spielertypen vorgenommen werden – was bislang nicht möglich war.

So wie das Kooperationsniveau eine Dimension einer Strategie ist, so ist das Strafverhalten eine weitere. Es ist nun von Interesse, die Wechselwirkungen dieser beiden Dimensionen zu analysieren. Daraus ergibt sich eine weitere Hypothese von Spielertypen und Straftypen:

Hypothese 2: Die beiden Strategie-Dimensionen Spielertyp und Straftyp treten bevorzugt in bestimmten Kombinationen auf.

Man kann vermuten, dass ein bestimmtes Kooperationsniveau mit einem dazu angepassten Strafniveau einhergeht. In diesem Fall müssten sich deutliche Häufigkeitsunterschiede zwischen bestimmten Kombinationen von Spielertypen und Straftypen zeigen. Dagegen spricht, dass Bestrafung sehr situationsabhängig eingesetzt werden muss, und von daher je nach Kontext stark variiert. Auf diese vermutete Kontextabhängigkeit zielt Hypothese 3:

Hypothese 3: Je nach Gruppenzusammensetzung ändert sich das Strafverhalten.

Bei dieser Hypothese kommt es nicht nur auf die Höhe der Strafen an, die auf dem Gruppenkontinuum von egoistisch bis hoch kooperativ stetig abnehmen sollte, sondern auf die *Rolle* der Spielertypen: Wer übernimmt in welchen Gruppen die Rolle der Bestrafenden? So ist denkbar, dass in kooperativeren Gruppen mit wenig Sanktionen nur noch diejenigen Spieler bestrafen, welche auch am stärksten in das Gemeingut investieren, da nur noch wenige Strafen nötig sind.

Die nächste Hypothese baut auf den Resultaten von Frey & Rusch (2012) auf, differenziert die Befunde jedoch nach Spielertyp aus:

Hypothese 4: Unabhängig vom Spielertyp wird anfangs im Vergleich zum späteren Verlauf überproportional viel bestraft.

Das beobachtete Muster in PGG besteht in häufigen und hohen Strafen zu Beginn des Spiels, die dann relativ schnell seltener werden und abnehmen. Im Laufe des Spiels sind dann fast keine Strafen mehr nötig; die bloße Drohung einer Strafmöglichkeit scheint auszureichen, um ein hohes Kooperationsniveau zu gewährleisten. Typischerweise beobachtet man in der letzten Runde ein Ansteigen der Defektion und gleichermaßen einen Anstieg der Bestrafung. Möglicherweise differenziert sich dieses allgemein beobachtete Muster jedoch je nach Spielertyp aus. Für diese Vermutung gibt es allerdings noch keinerlei Bestätigung oder Widerlegung. Aus diesem Grund wird eine detaillierte Analyse der Bestrafung nach Runde und Spielertyp vorgenommen.

Als letzten Punkt untersuche ich den Zusammenhang dreier entscheidender Kenngrößen in PGG: den Einzahlungen in das öffentliche Gut, dem Verdienst vor der Bestrafung und nach der Bestrafung. Dies lässt sich wie folgt formulieren:

Hypothese 5: Hohe Einzahlungen führen zusammen mit Bestrafung der Trittbrettfahrer zu einem Wettbewerbsvorteil für kooperative Individuen.

Das Verhältnis von Einzahlungen zum letztlichem Erfolg (Verdienst) bei sozialen Dilemmata ist logischerweise invers: je mehr ein Individuum einzahlt, desto weniger bleibt ihm und desto mehr profitieren im Vergleich die anderen. Trotzdem führen Institutionen mit Sanktionsmöglichkeit dazu, dass ein Gleichgewicht gefunden werden kann, das näher am sozialen Optimum liegt als ohne Sanktionen. In diesem Fall sollten Gruppen, die aus kooperationsbereiten Individuen bestehen, egoistischen Gruppen überlegen sein, da sie durch die hohen Einzahlungen effizienter sind und möglicherweise nur geringe Strafkosten verursachen.

6 Stand der Forschung zur Studie 3 (Erfolg Kooperierender) – Die evolutionäre Funktionslogik sozialer Dilemmata

Vergleichsweise wenig Literatur gibt es für den direkten Vergleich des (evolutionären) Erfolgs zwischen altruistischen Strategien und Trittbrettfahrern. Für diese Forschungslücke gibt es mehrere Gründe. Aus theoretischer Sicht lässt sich für solche Vergleiche wenig Überraschendes erwarten, da Trittbrettfahrer Kooperierenden überlegen sein müssten, weil solche

Strategien weniger Kosten für den Trittbrettfahrer verursachen; dies gilt zumindest in Situationen, in denen Verwandtschaft keine Rolle spielt. Aus praktischer Sicht ist diese Forschungslücke ebenfalls verständlich, da es extrem schwierig ist, bei verschiedenen Verhaltensweisen wie eben Trittbrettfahren Fitness zu messen und deren Fitnesskonsequenzen zu identifizieren (Bshary & Bergmüller 2008).

Ich meine allerdings, dass es eine einfachere Erklärung gibt, die oft nicht die Beachtung findet, die sie verdient. Unter einigen Umständen könnten Kooperierende erfolgreicher sein als Trittbrettfahrer (d.h. eine höhere Fitness haben). Man kann zudem vermuten, dass diese Umstände bzw. Umweltparameter während der Zeit als Jäger und Sammler gegeben waren.

Das *Explanandum* ist demnach die kulturübergreifende Bereitschaft von Menschen, in Gemeingüter einzuzahlen. Die Belege reichen von natürlichem Verhalten wie Fleisch- und Arbeitsteilung bei Jägern und Sammlern (Hill 2001) über Experimente in ebendiesen Gesellschaften (Henrich et al. 2001) bis hin zu kulturübergreifenden Befunden bei modernen Gesellschaften in Experimenten (Gächter et al. 2010), wie bereits oben diskutiert (Abschnitt A 2). Dies findet sich auch in vielen Phänomenen des modernen täglichen Lebens, wie etwa die fast durchgängige Bereitschaft, Steuern zu zahlen.

Die vielleicht wichtigste Frage bei den untersuchten sozialen Dilemmata ist, ob es sich überhaupt um altruistisches Verhalten handelt. Erleidet der Einzahlende tatsächlich Einkommens-, Ressourcen- bzw. Finesseinbußen? In der herkömmlichen Interpretation ist dies der Fall, da ein Einzahlender Kosten in Kauf nimmt, von denen andere profitieren können. Trotzdem legt die ubiquitäre Natur dieses Verhaltens nahe, dass es – da auch funktional in Abwesenheit von Verwandten – einen fundamentalen anderen Grund und Nutzen dieses Verhaltens geben muss. Diese Arbeit vermutet diesen Grund in einer Kombination von Situationseigenschaften, die ein solches Verhalten nicht mit *Finesseinbußen* assoziieren, sondern im Gegenteil mit proximatem *Fitnessvorteilen*.

Sobald es Individuen möglich ist, durch räumliche Trennung oder andere Abwehrmaßnahmen wie Bestrafung Trittbrettfahrer von den Profiten eines Gemeingutes abzuhalten, verfügen sie nämlich über die dann *exklusiven* Vorteile (etwa einer sonst nicht vorhandenen Infrastruktur, einer effizienten Verteilung von Gütern, usw.) des Gemeingutes. Diese Vorteile sollten – auf lange Sicht – zu deutlichen Vorteilen gegenüber Defektionsstrategien führen. In diesem Sinne dürften wir auch nicht von einem proximatem altruistischem Verhalten sprechen, sondern von einem sowohl proximat als auch ultimat (gen-)-egoistischen.

Deshalb werden erwartet:

- konditionale Kooperationsstrategien, d.h. bei Erkennen von Trittbrettfahrern erfolgt die schnellstmögliche Beendigung der Interaktion; im Umkehrschluss: Beibehaltung von profitablen Kooperationsbeziehungen; daraus folgt eine relativ hohe Variabilität von Strategien, je nach Strategien und Häufigkeit anderer Spielertypen,
- Versuche, soziale Netzwerke mit anderen Investoren aufzubauen,
- Versuche, sich räumlich von Trittbrettfahrern abzugrenzen bzw. zu trennen,
- Etablierung von Abwehrmaßnahmen gegenüber Trittbrettfahrern wie Bestrafung, der Reputationsaufbau bei guten Investoren, verbale oder sonstige „Kennzeichnung“ von Trittbrettfahrern.

Diese Liste erhebt keinen Anspruch auf Vollständigkeit, sondern soll einen Überblick über mögliche Folgen geben, falls die obige Hypothese stimmen sollte. Ich komme in der Diskussion (Abschnitt D) darauf zurück, inwiefern sich diese Erwartungen in der Datenanalyse bestätigen.

Natürlich ist es Kooperierenden und Reziprokatoren (konditional Kooperierenden) nicht in allen Situationen möglich, Trittbrettfahrer abzuwehren. Beispiele für solche Situationen sind Sozialgemeinschaften, in denen es schwierig ist, unkooperatives Verhalten zu beobachten, oder Situationen, in denen Interaktionspartner nicht gewählt werden können, sondern zugewiesen werden. Allerdings ist es selbst in solchen oder ähnlichen Situationen oft möglich, Maßnahmen wie etwa Bestrafung zu ergreifen, um unkooperatives Verhalten zu verringern (Güerck et al. 2006; Ostrom et al. 1994).

Unterstützung für diese Argumentation kommt auch aus der Struktur der biologischen Nische der Menschen. Die Entwicklung eines solchen Verhaltens könnte durch Verwandtschaft in den Kleingruppen typischer Jäger und Sammler-Gesellschaften (Hill et al. 2011) extrem begünstigt gewesen sein. Weiterhin benötigt die Amortisation von Vorteilen durch Gemeingüter eine lange biologische Lebensdauer, was ebenfalls gegeben ist (so konnte z. B. für Vögel gezeigt werden, dass hier geringe Erwachsenensterblichkeit gekoppelt mit geringer Zerstreung (*dispersal*) zu höheren Raten von kooperativem Brüten führen, Hatchwell & Komdeur 2000). Ein weiterer Punkt, der hier in das Bild passt, sind unsere hoch entwickelten sozialen Fähigkeiten (Dunbar 2003). Wir sind fähig, eine andere Perspektive einzunehmen (*theory of mind*), können Betrüger gut entdecken (Cosmides & Tooby 1992), und verstehen die Intentionen

anderer (Volland 2007). Diese Fähigkeiten sind die Basis komplexer sozialer Interaktionen, wie sie Investitionen in Gemeingütern zugrunde liegen.

Nur eine Auftrennung in Gruppen kann das Problem überhaupt in den Blick bekommen, *da nicht angezweifelt wird, dass im Allgemeinen Trittbrettfahrer besser fahren als Individuen, die viel in ein Gemeingut einzahlen*. Ergeben sich aber durch stochastische Prozesse oder sogar durch bewusste Steuerung Gruppen, die nur aus hohen Investoren bestehen, greift die Überlegung, dass hier Wettbewerbsvorteile entstehen. Solche Gruppen aus Investoren sind auf Individualebene überlegen, da sie hohe Effizienzgewinne erzielen. Da eine Einwanderung von Trittbrettfahrern, sei es durch Mutation oder durch Migration nie verhindert werden kann, ist diese Strategie allerdings nicht evolutionär stabil. Sie kann sich aber immer wieder neu bilden, wenn geeignete Bedingungen herrschen (für Bakterien siehe MacLean & Gudelj 2006; Kerr et al. 2006). Vor allem der Migrationsdruck sollte entscheidend für die Häufigkeit sein, mit der Strategien auftreten, die viel in Gemeingüter investieren. Wie oben erwähnt, kann man deshalb einen hohen Selektionsdruck auf die Fähigkeit erwarten, zwischen kooperativeren und nicht kooperativen Individuen zu unterscheiden, was durch Experimente für Menschen auch bestätigt wird (Barclay 2008; Oda 1997). Die Erinnerung an Strategien von Interaktionspartnern wird dabei möglicherweise durch deren Häufigkeiten modifiziert (Bell et al. 2010).

Eine weitere Frage, die erst später (D 2 und 3) beantwortet werden kann, ist, wie stabil die Strategie gegen Defektoren ist und über welche Mechanismen die Abwehr erfolgt. Diese Arbeit vermutet auf Basis vorhandener Belege, dass die Abwehr vor allem über konditionale Kooperation (Gächter 2007) in Verbindung mit Gruppierungsmechanismen zusammen mit anderen Kooperierenden (*assortment*) erfolgt (Page et al. 2005). Zusätzlich scheint die Defektion gegen Defektoren wichtiger zu sein als Bestrafung gegen Dritte (Marlowe et al. 2011).

7 Arbeitshypothesen zur Studie 3 (Erfolg Kooperierender)

Auf Basis dieser Überlegungen kann bereits vermutet werden, wie häufig welche Strategien auftreten, da die Häufigkeit kooperativer Verhalten stark abhängig ist vom Auftreten anderer Strategien. Als vorläufige Vermutung kann ein hoher Anteil an konditional Kooperierenden (siehe vorheriger Absatz), ein kleiner Anteil an hohen Investoren (da sie eine Nischenstrategie sind), und ein relativ konstanter Bereich der dominanten Strategie (Defektion) erwartet werden, da letztere sowohl konditional Kooperierende als auch Investoren ausnutzen können.

Weiterhin leite ich aus den oben angestellten theoretischen Überlegungen ab (Abschnitte A 2 sowie B 1), dass der Erfolg der gewählten Strategie bei der Investition in Gemeingüter vor

allem davon abhängt, wie viele andere Individuen bereit sind, ebenfalls zu investieren und inwiefern es gelingt, Trittbrettfahrer davon abzuhalten, vom Ertrag des Gemeingutes zu profitieren. Daraus ergibt sich ein Kontinuum, bei dem im einen Extrem ein einzelnes Individuum ein Gemeingut erzeugt, von dem viele (es selbst auch) profitieren. Hier ist der Versuch von „Kooperation“ gescheitert und die Kosten extrem hoch. Wenn sich weitere Individuen beteiligen, sinken zwar die Kosten, Trittbrettfahrer profitieren aber im Vergleich noch mehr, da sie überhaupt keine Kosten tragen müssen. Gelingt es aber außerdem noch, innerhalb eines sozialen Netzwerkes von Kooperierenden zu bleiben, sind also Trittbrettfahrer durch räumliche Trennung oder Abwehrmaßnahmen wie Sanktionen (die zu den Kosten hinzugerechnet werden müssen, aber im Vergleich zur Erzeugung vergleichsweise niedrig sind) nicht in der Lage, vom Gemeingut zu profitieren, dann amortisieren sich – abhängig von Zahl der Beteiligten, Kosten der Erzeugung und Nutzen des Gemeingutes – nach einiger Zeit die Kosten. In solchen Situationen sind Kooperierende Individuen Egoisten überlegen, wobei die Schere mit fortschreitender Zeit immer weiter zugunsten der Kooperierenden aufgehen sollte.

Diese grundsätzlichen Überlegungen führen zu spezifischen und testbaren Hypothesen. Diese Hypothesen beziehen sich alle auf proximate Mechanismen.

Hypothese 1: Egoistischere Strategien sind – ohne zusätzliche Randbedingungen – kooperierenden Strategien überlegen.

Je weniger ein Individuum in das öffentliche Gut investiert, desto geringer seine Kosten, während es gleichzeitig von dessen Erzeugung durch andere profitiert. Diese Folgerung leitet sich direkt aus der Struktur von sozialen Dilemmata ab und ist in linearen PGG eine dominante Strategie.

Hypothese 2: Sind Kooperierende unter sich, sind sie Trittbrettfahrergruppen überlegen.

Je mehr Kooperierende in ein Gemeingut investieren, desto näher operieren sie am sozialen Optimum. Dadurch erreichen sie eine hohe Effizienz und können die damit verbundenen Vorteile abschöpfen, wie es egoistische Strategien nicht können.

Hypothese 3: Der Erfolg von Strategien ist unmittelbar von der Gruppenzusammensetzung abhängig, d.h. von denjenigen Strategien, mit denen interagiert wird.

Je nach Gruppenzusammensetzung sollte sich der Erfolg wie folgt staffeln: Nur Kooperierende (K), Kooperierende und Reziprokatoren (KR), nur Reziprokatoren (R), Kooperierende, Reziprokatoren und Trittbrettfahrer (KRT), Kooperierende und Trittbrettfahrer (KT), Rezipro-

katoren und Trittbrettfahrer (RT), und Trittbrettfahrer (T).

Kooperierende sind dann im Vorteil, wenn Trittbrettfahrer ausgeschlossen sind (siehe Hypothese 2). In solchen Fällen können Sie die Vorteile in Erfolg ummünzen. Je stärker Egoisten Gruppenzusammensetzungen prägen, desto schlechter sollten diese Gruppen im Vergleich abschneiden.

Hypothese 4: In der direkten Auseinandersetzung zwischen den beiden Extremstrategien Kooperation (K) und Trittbrettfahren (T) sollte sich der größte Erfolgsunterschied zu Gunsten von Trittbrettfahrern zeigen.

Trifft kooperierendes Verhalten, das nicht konditional auf die Investitionen anderer Spieler reagiert, auf reine Trittbrettfahrer, so wird es ausgebeutet und bricht bezüglich des Erfolges ein. Dies lässt sich beispielsweise mit den Auszahlungen des Gefangenendilemmas und den Strategien Falke (*always defect, All-D*) und Taube (*always cooperate, All-C*) gut demonstrieren (Axelrod 1984/2000).

Hypothese 5: Alle Spielertypen sollten eine längere Interaktionsdauer mit Kooperierenden als mit Egoisten aufweisen.

Konditional Kooperierende machen ihre Kooperationsbereitschaft von den Einsätzen anderer abhängig. Wenn Sie auf Trittbrettfahrer treffen, sollten sie ihre Einsätze verringern, bzw. die Interaktion so schnell wie möglich beenden. Da Trittbrettfahrer von kooperativem Verhalten profitieren, gilt die umgekehrte Logik für ihre Interaktion mit Kooperierenden. Auch Trittbrettfahrer sollten ihre Interaktionen mit anderen Trittbrettfahrern so schnell wie möglich beenden, da sie daran interessiert sein müssen, kooperativere Individuen auszubeuten. Dafür ist es nötig, bei Bedarf zu Interaktionspartnern zu wechseln, die leichter ausgebeutet werden können. Dieselbe Funktionslogik gilt für Kooperierende selbst – jede Strategie sollte von Trittbrettfahrern weg und hin zu Kooperierenden wechseln.

Hypothese 6: Je länger die Interaktionsdauer, desto größer die Differenzen im Erfolg zwischen kooperierenden und egoistischen Strategien.

Aus mehreren Gründen profitieren kooperative Interaktionen von längeren Zeiträumen. Zum einen nimmt das Vertrauen in den Partner zu (Axelrod 1984/2000), zum anderen verändern sich die Erwartungen im positiven Sinne (Gächter et al. 2008) und auch die Effizienz von flankierenden Maßnahmen wie Bestrafung nimmt mit der Zeit zu (Abschnitte B 4 und D 2). In unserer evolutionären Vergangenheit waren wiederholte Interaktionen die Regel (Hill 2001).

Ein weiterer Grund besteht darin, dass öffentliche Güter (egal welcher Natur), die von Gruppen aus Egoisten nicht erzeugt werden können, einmal vorhanden, ihre Effizienzvorteile über längere Zeiträume ausbauen können. Zudem benötigen sie meist kaum weitere Investitionen, um aufrechterhalten werden zu können (z. B. Gemeinschaftshäuser o. ä.). Andere öffentliche Güter wie beispielsweise Bewässerungssysteme benötigen zwar stetiges kollektives Handeln, was aber im Vergleich zur erstmaligen Bereitstellung nicht zu stark ins Gewicht fällt.

Hypothese 7: Strategien, die in der Lage sind, einen Trade-off zwischen der Erzeugung öffentlicher Güter und minimaler persönlicher Investition zu finden, sollten hohen Erfolg zeigen.

Gerade so viel zu investieren, dass konditionale Kooperierende ebenfalls investieren, dabei aber eigene Verluste in Form von Kosten so gering wie möglich zu halten, sollte das Beste beider Extremstrategien miteinander verbinden und daher sehr erfolgreich sein. Der Hintergrund für diese Überlegungen kommt aus der experimentellen Ökonomik, wo vermutet wird, dass das Absinken des Kooperationsniveaus, das typischerweise in PGG beobachtet wird (Chaudhuri 2011), darauf zurückzuführen ist, dass es verschiedene Spielertypen gibt (Fischbacher et al. 2001). Da die größte Gruppe an Spielertypen konditionale Kooperierende (Reziprokatoren) sind, die ihre Investitionen von denen der anderen abhängig machen, und es gleichzeitig etwa 20-25 % an Trittbrettfahrern gibt (Rustagi et al. 2010; Kocher et al. 2008), geht man davon aus, dass erstere Gruppe frustriert von den fehlenden Kooperation der Letzteren ihre Bemühungen mit der Zeit einstellt. Erfolgversprechender als nichts zu geben, könnte deshalb eine Strategie sein, die mit minimalem Einsatz die Bereitschaft zur Investition konditional kooperierender Spieler aufrechterhält.

Um solche Fragen überhaupt beantworten zu können, ist allerdings geeignetes Datenmaterial nötig. Spielertypen müssen a) identifiziert, b) in gemischte (K und T) und homogene (nur T) Gruppen eingeteilt werden können und c) ein klar messbares Kriterium für den Erfolg muss vorhanden sein. Aus diesen Gründen werden zwei kontrollierte Umgebungen (fünf PGG und das Online-Spiel Ikariam, siehe C 1 und C 2) verwendet. Bei beiden ist die Problemstruktur ein soziales Dilemma in Form eines Gemeingutproblems (PGG); die Daten des ersten stammen aus fünf veröffentlichten Labor-PGG. Der nächste Teil der Arbeit (Teil C) beschreibt die verwendeten Datensätze und Methoden.

C Daten und Methoden

Um die Arbeitshypothesen zu belegen oder zu falsifizieren, wählt diese Arbeit zwei in ihren Parametern sehr unterschiedliche Datensätze aus, deren Kern jedoch jeweils ein soziales Dilemma in Form eines Gemeinguts ist.

1 *Beschreibung der Daten aus PGG-Laborexperimenten*

Ich analysiere Daten aus fünf im *peer-review*-Verfahren publizierten Studien aus der experimentellen Ökonomik, die ihre Daten öffentlich zur Verfügung gestellt haben. Diese Arbeiten sind: Fehr & Gächter 2000; Fehr & Gächter 2002; Gächter et al. 2008; Nikiforakis & Normann 2008 und Nikiforakis 2008. Diese unabhängigen Daten bieten einige Vorteile:

- die Stichprobe ist größer,
- mögliche Verzerrungen in den Daten werden dadurch reduziert (z.B. Erwartungseffekte, siehe Frey 2007),
- Effekte können als robuster gelten, wenn sie in allen Teilstichproben gefunden werden, und
- die Datenqualität ist hoch, da alle Studien *peer-reviewed* wurden.

Für Details zu den einzelnen Studien siehe Tabelle 2.

Es gibt mehrere Gründe für die Auswahl: jede der Studien hat sowohl ein Referenz-Treatment (*baseline*), das ohne Manipulation von Parametern stattfindet und einen Vergleich zur jeweiligen experimentellen Beeinflussung ermöglicht, sowie ein Treatment mit Bestrafung. Treatments mit Bestrafung werden im Folgenden mit P (für *punishment*), solche ohne mit NP (*no punishment*) abgekürzt. Des Weiteren sind Gruppengrößen, *marginal payoff*, Ausstattung (*endowment*) und Rundenzahl weitgehend vergleichbar (siehe Tabelle 2). Ein weiteres Kriterium ist die Datenverfügbarkeit, da andere Studien ihre Daten in Rohform oft nicht zur Verfügung stellen. Diesbezüglich ist als einzige Einschränkung in diesem Datensatz zu nennen: Nikiforakis & Normann (2008) stellen die Daten *nach* Bestrafung nicht zur Verfügung, so dass dieser Teil nicht vollständig ist. Alle weiteren Treatments neben dem Referenz- und dem Bestrafungs-Treatment wurden ausgeschlossen (z. B. Counter-Punishment in Nikiforakis & Normann 2008). In Nikiforakis (2008) wird die Effektivität von Bestrafung von 1:1 bis auf 1:4 variiert (d.h. je eingesetzter Geldeinheit für die Bestrafung bekommt der Bestrafte 1, 2, 3

und 4 Geldeinheiten abgezogen). Diese Daten werden als P-Treatment aggregiert, ohne weiter zwischen ihnen zu differenzieren.

Diese Arbeit konzentriert sich in Studie 2 auf drei Studien: Nikiforakis & Normann (2008), Fehr & Gächter (2000) und Fehr & (2002). In Studie 3 werden dagegen alle fünf Studien verwendet. Die Beschränkung auf drei Studien für Fragen der Bestrafung ist durch die fehlenden Daten für eine Einzelspielerzuordnung der Bestrafung in den Studien Gächter et al. (2008) und Nikiforakis (2008) bedingt.

Die untenstehende Tabelle 2 gibt zunächst einen Überblick über die wichtigsten Parameter der verwendeten PGG-Studien.

Studie	Gruppen- größe	Partner oder Stranger	Zahl der Teilnehmer	Runden- zahl	Aus- stattung	Bestra- fungs- ratio	Marginal payoff
Fehr & Gächter 2000	4	P und S	224	10	20	1:1 ansteigen d zu 1:3	0,4
Fehr & Gächter 2002	4	S	240	6	20	1:3	0,4
Gächter et al. 2008	3	P	207	10 (T1) 50 (T2)	20	1:3	0,5
Nikiforakis & Normann 2008	4	P und S	192	10	20	1:1 ansteigen d zu 1:3	0,4
Nikiforakis 2008	4	P	120	10	20	1:1 (T1) 1:2 (T2) 1:3 (T3) 1:4 (T4)	0,4

Tabelle 2. Vergleich der wichtigsten Parameter der fünf verwendeten PGG-Studien

Eine Bemerkung zur Wortwahl: aus Mangel an geeigneten bzw. gebräuchlichen griffigen deutschen Alternativen bleiben die Begriffe *Treatment* (spezifische Versuchsanordnung), *Partner* (gleichbleibende Gruppenzusammensetzung während des PGG), *Stranger* (Gruppenzusammensetzung, die jede Runde wechselt) unübersetzt und sind im Folgenden auch nicht mehr kursiviert.

Ein zentraler Forschungsaspekt der vorliegenden Arbeit ist die Differenzierung verschiedener Spielertypen. Sie werden nach ihren Beiträgen zum Gemeingut identifiziert und wie folgt definiert:

Spielertyp	Im Text verwendete Abkürzung	Definition
Trittbrettfahrer	T	0% der eigenen Ressourcen sind Beiträge für das Gemeingut
Reziprokator	R1	< 10%
Reziprokator	R2	< 20%
Reziprokator	R3	< 30%
Reziprokator	R4	< 40%
Reziprokator	R5	< 50%
Kooperierender	K	>= 50% der eigenen Ressourcen sind Beiträge für das Gemeingut

Tabelle 3. Definition der Spielertypen

Ursprünglich war eine Aufteilung von Spielertypen nur nach K, R und T geplant, da in der Literatur auch nicht zwischen R1 bis R5 unterschieden wird (Fischbacher et al. 2001). Da die strategische Breite des Spielertyps Reziprokator (R) aber sehr hoch ist, ist eine Aufteilung für eine differenziertere Analyse in fünf Untergruppen R1 bis R5 sinnvoll. Die Wortwahl „Reziprokator“ ist dabei ein Kompromiss. Gemeint sind konditionale Strategien, also Strategien, die ihre Kooperationsbereitschaft von der anderer abhängig machen – im Englischen *conditional cooperator*. Durch die sprachliche Nähe zu Kooperierenden (K) ziehe ich deshalb Reziprokatoren vor (was in einigen andere Studien auch so gehandhabt wird, z. B. Kurzban & Houser 2001). Der evolutionäre Mechanismus der direkten oder indirekten Reziprozität kann, muss aber nicht notwendigerweise assoziiert werden.

In der Literatur werden konditionale Kooperierende (hier R1 bis R5) meist über einen positiven Spearman-Koeffizienten ihrer Investitionen in Hinblick auf die Investitionen der anderen Spieler definiert. Diese Arbeit wählt stattdessen ein einheitliches Merkmal (< X % der Investitionen, siehe Tabelle 3 und Abschnitt C 2.1), um ein Kontinuum zwischen den Spielertypen zu gewährleisten. Anderenfalls besteht die Gefahr, dass Spieler sowohl die Kriterien für Kooperierende (=> 50%) und konditional Kooperierende erfüllen und doppelt gezählt würden. Da ich zudem die Häufigkeiten der Spielertypen für verschiedene Messmethoden angebe, ist ein Vergleich zur herkömmlichen Messmethode immer möglich (siehe Tabelle 13). Es zeigt sich überdies, dass die Häufigkeiten der Typen mit den eben diskutierten zwei Methoden vergleichbar sind.

Analog dazu definiere ich die sieben möglichen Gruppenzusammensetzungen in der folgenden

Tabelle 4:

Gruppentyp	Im Text verwendete Abkürzung	Definition
Nur Kooperierende	K	Gruppe, die nur aus dem Spielertyp K besteht (mind. 2 Spieler)
Kooperierende und Reziprokatoren	KR	Gruppe, die ausschließlich aus den Spielertypen K und R besteht (mind. 2 Spieler, mind. 1 Spieler K und 1 Spieler R)
Nur Reziprokatoren	R	Gruppe, die nur aus den Spielertypen R1 bis R5 besteht (mind. 2 Spieler)
Kooperierende, Reziprokatoren und Trittbrettfahrer	KRT	Gruppe, die aus den Spielertypen K, R und T besteht (mind. 3 Spieler, mind. 1 Spieler K, mind. 1 Spieler R, mind. 1 Spieler T)
Kooperierende und Trittbrettfahrer	KT	Gruppe, die ausschließlich aus den Spielertypen K und T besteht (mind. 2 Spieler, mind. 1 Spieler K und 1 Spieler T)
Reziprokatoren und Trittbrettfahrer	RT	Gruppe, die ausschließlich aus den Spielertypen R und T besteht (mind. 2 Spieler, mind. 1 Spieler R und 1 Spieler T)
Nur Trittbrettfahrer	T	Gruppe, die nur aus dem Spielertyp Trittbrettfahrer besteht (mind. 2 Spieler)

Tabelle 4. Definition der Gruppentypen

Allerdings sind selbst bei dieser vergleichsweise großen Anzahl an Versuchspersonen aus der Gesamtheit der PGG (n = 983, siehe Tabelle 2) die Grenzen schnell erreicht, wenn nach speziellen Gruppentypen gefiltert wird.

Auch Strafende und Bestrafte werden – je nach Höhe der Strafe / Bestrafung – in sieben Gruppen eingeteilt (Tabelle 5). Die Einteilung erfolgt nach Strafpunkten, nicht nach tatsächlichen Kosten für Strafer bzw. Bestraften, um einen Vergleich zwischen den Studien zu ermöglichen. Die Grenzen für die Typen 1 bis 7 sind nach Prozent festgelegt und entsprechen 0%, 1%, 3%, 5%, 7%, 9% und mehr als 11%.

Straftyp	Bestrafttyp	Definition (Höhe der gegebenen bzw. erhaltenen Strafpunkte) in %
Straftyp 1	Bestrafttyp 1	0
Straftyp 2	Bestrafttyp 2	1-3
Straftyp 3	Bestrafttyp 3	4-9
Straftyp 4	Bestrafttyp 4	10-15
Straftyp 5	Bestrafttyp 5	16-21
Straftyp 6	Bestrafttyp 6	22-27
Straftyp 7	Bestrafttyp 7	>= 28

Tabelle 5: Definition der Straf- und Bestrafttypen

Als abhängige Variable wird der Gesamtverdienst verwendet. Dieser wird über alle Spielertypen aggregiert. Da der Gesamtprofit zwischen den Studien nicht vergleichbar ist, weil sowohl die Rundenzahl als auch der Auszahlungsmultiplikator zwischen den Studien unterschiedlich sind, wird der Gesamtverdienst *in Prozent des maximal erreichbaren Verdienstes* angegeben, um das Ergebnis zwischen den Studien vergleichbar zu machen. Diese Kennzahl unterscheidet sich vom sozialen Optimum. Das soziale Optimum errechnet sich aus:

$$\text{Spieleranzahl pro Gruppe} * \text{Ausstattung (endowment)} * \text{Multiplikator} * \text{Rundenzahl}$$

Die einzelnen Werte für diese Variablen sind in Tabelle 2 oben angegeben. So beträgt beispielsweise das soziale Optimum für Fehr & Gächter 2000:

$$4 * 20 * 0,4 * 10 = 320$$

Das bedeutet, jeder Spieler kann aus seiner Ausstattung von 200 fiktiven Geldeinheiten (20 Einheiten * 10 Runden) maximal 320 machen. Das geht allerdings nur, wenn jeder Spieler in allen Runden seine gesamte Ausstattung in das Gemeingut investiert. Der minimale Verdienst ist dagegen mit 200 gegeben (20 Einheiten * 10 Runden) – in diesem Fall wird nichts in das Gemeingut investiert.

Der maximal erreichbare Verdienst dagegen errechnet sich aus der maximalen Einzahlung aller anderen Spieler bei gleichzeitiger Enthaltung eines Spielers. Die Beispielrechnung ist wiederum für Fehr & Gächter 2000:

$$3 * 20 * 0,4 * 10 = 240 + 200 = 440$$

Hier werden 240 Einheiten aus dem Gemeingut erwirtschaftet (und zwar ausschließlich durch die anderen Spieler). Zu diesem Verdienst kommen noch 200 Einheiten der eigenen Ausstat-

tung, die erhalten und eben nicht in das Gemeingut eingezahlt wurden. Im Extremfall kann also ein Spieler mehr als das soziale Optimum verdienen, der gesamte Verdienst für die Gruppe fällt dann allerdings im Beispiel vom sozialen Optimum 1280 ($4 \cdot 320$) auf 1160 ($3 \cdot 240 + 440$). Ein stabiles Nash-Gleichgewicht liegt im Beispiel bei 800 ($4 \cdot 200$).

2 Beschreibung der Daten des Online-Browser-Spiels Ikariam

Der zweite Datensatz, den ich für meine Analysen heranziehe, ist – in meinen Augen – einzigartig auf der Welt. Es handelt sich um Daten aus dem Online-Browser-Spiel Ikariam, in dem völlig analog zu Standard-PGG, wie sie im Labor gespielt werden, drei PGG eingebaut sind. Mehrere Besonderheiten zeichnen diese Daten aus: Erstens gibt es kein PGG, das jemals mit mehr als einigen hundert Teilnehmern gespielt wurde – hier sind es über 18 000. Diese Teilnehmer spielen unter identischen Bedingungen und sind damit unmittelbar vergleichbar. Zweitens sind länderübergreifende Vergleiche für PGG selten (für Ausnahmen siehe Herrmann et al. 2008; Henrich et al. 2001; Kocher et al. 2008) – mit dem Ikariam-Datensatz können Spieler aus fünf Ländern (Deutschland, England, Frankreich, Griechenland und der Türkei) verglichen werden. Drittens ist die Spieldauer in PGG meist auf zehn Runden bzw. einige Stunden realer Zeit beschränkt (was nicht unerhebliche negative Folgen hat, siehe Frey & Rusch 2012) – bei Ikariam liegen dagegen Daten für zehn Monate und tausende von „Runden“ vor. Viertens, die Daten sind sehr viel reichhaltiger als die Daten herkömmlicher PGG, erlauben also Analysen, die anders nicht möglich wären. Fünftens halte ich den Kontext und die Motivation der Spieler für sehr viel realistischer als in Laborexperimenten, weil das Spiel aus intrinsischen, nicht extrinsischen Gründen gespielt wird und weil diese Motivation bei vielen Spielern über Monate und Jahre aufrecht erhalten wird. Auch der Kontext ist realistischer, da Ikariam zum einen vollständig grafisch basiert ist und zum anderen bekannte Konzepte wie den Aufbau von Städten, Bevölkerungswachstum, Handel und Allianzen usw. umsetzt und nicht abstrakt bleibt. Dies steht bei einer virtuellen Welt natürlich zur Diskussion.

Ich beginne mit einer allgemeinen Beschreibung des Spieles selbst und der Spielwelt. Ikariam ist ein frei zugängliches Online-Browser-Computerspiel (<http://us.ikariam.gameforge.com/>). Ikariam ist in 60 Ländern zugänglich und hat mehr als 51 Millionen registrierte Mitspieler. Weitere Informationen über Ikariam sind unter http://ikariam.wikia.com/wiki/Main_Page zu finden. Es wird von Spielern finanziert, die mit echtem Geld für Vorteile innerhalb der Spielwelt bezahlen.

Alle Teilnehmer interagieren mit anderen menschlichen Spielern in Echtzeit – das heißt, das

Spiel geht weiter, auch wenn man ausgeloggt ist. Die Spieler loggen sich – unabhängig voneinander – ein und aus. Das Spielziel ist es, bis zu zwölf unterschiedliche Städte auf beliebigen Inseln in einer „Welt“ weiterzuentwickeln. In jeder Welt gibt es 5351 Inseln und eine variable Anzahl von Spielern. In dem uns zur Verfügung gestellten Datensatz, der zehn Monate umfasst, nahmen – abhängig vom untersuchten Land der echten Welt – durchschnittlich zwischen 2257 (England) und 6112 (Türkei) Spieler pro Welt am Spiel teil.

Erfolgreiches Handeln führt zu einem besseren Spielergebnis und höherem *Highscore*. Auf jeder Insel gibt es 17 Orte, an denen Städte gebaut werden können, und drei Rohstoffquellen, die für alle Spieler zugänglich sind und die benötigt werden, um Rohstoffe zu erwerben und Gebäude zu verbessern. Die Entwicklung der Städte hängt hauptsächlich von zwei Größen ab: der Zeit, die der Spieler Ikariam spielt, und – entscheidender – von der Menge der abgebauten Rohstoffe. Letzteres wiederum hängt von der *Ausbaustufe* der Rohstoffquelle der Insel ab. Ist die Ausbaustufe gering, so lässt das Programm nur eine stark eingeschränkte Menge an Rohstoffabbau pro Stunde zu – auch wenn dieser sich nie erschöpfen kann. Die Effizienz des Abbaus wird demnach durch die Ausbaustufe bestimmt. Diese Stufe ist ein Gemeingut, weil die Stufe (also die Effizienz) von den Investitionen der Spieler abhängt. Da die Investitionen aus den privaten Ressourcen der Spieler stammen, die höhere Ausbaustufe aber allen Teilnehmern (auf dieser Insel) zu Gute kommt, entspricht dieses Setup exakt einem sozialen Dilemma, wie es auch in PGG in der experimentellen Ökonomik gespielt wird. Eine von Ikariam auf ihrer deutschen Website unternommene Online-Befragung (14.3.2014 – 30.3.2014), an der 429 Spieler aus Deutschland teilnahmen, zeigt, dass die Spieler sich dieses Dilemmas ebenso bewusst sind wie der damit einhergehenden Strategien wie Trittbrettfahren oder kooperativem Verhalten (vgl. Anhang G für weitere Informationen). Der folgende Screenshot zeigt die Bestandteile des Gemeinguts in Ikariam und macht deutlich, dass das soziale Dilemma einem PGG analog strukturiert ist.

Teil 1 der Abbildung zeigt die Arbeiter, die dem Privateinkommen der Rohstoffproduktion eines Spielers zugeteilt sind (30 von 40 Arbeiter sind hier zugeteilt). Teil 2 zeigt die Ausbaustufe (Level 26), den Schwellenwert der Kosten, die notwendig sind, um diese Stufe zu erreichen (767 723 Einheiten Holz) sowie die Rohstoffe, die für das Erreichen dieser Stufe bereits in das Gemeingut investiert wurden (200 Einheiten Holz). Teil 3 zeigt die Spielerpseudonyme, die Städtenamen, das Level der Städte und die Anzahl der Arbeiter, die von den einzelnen Spielern dieser Rohstoffressource pro Stadt zugeteilt sind sowie den Gesamtwert der Beiträge pro Spieler, die bereits für das Gemeingut geleistet wurden (der Spieler „mortis“ spendete

beispielsweise 30 000 Einheiten, wohingegen „pseudonym“ 0 spendete).

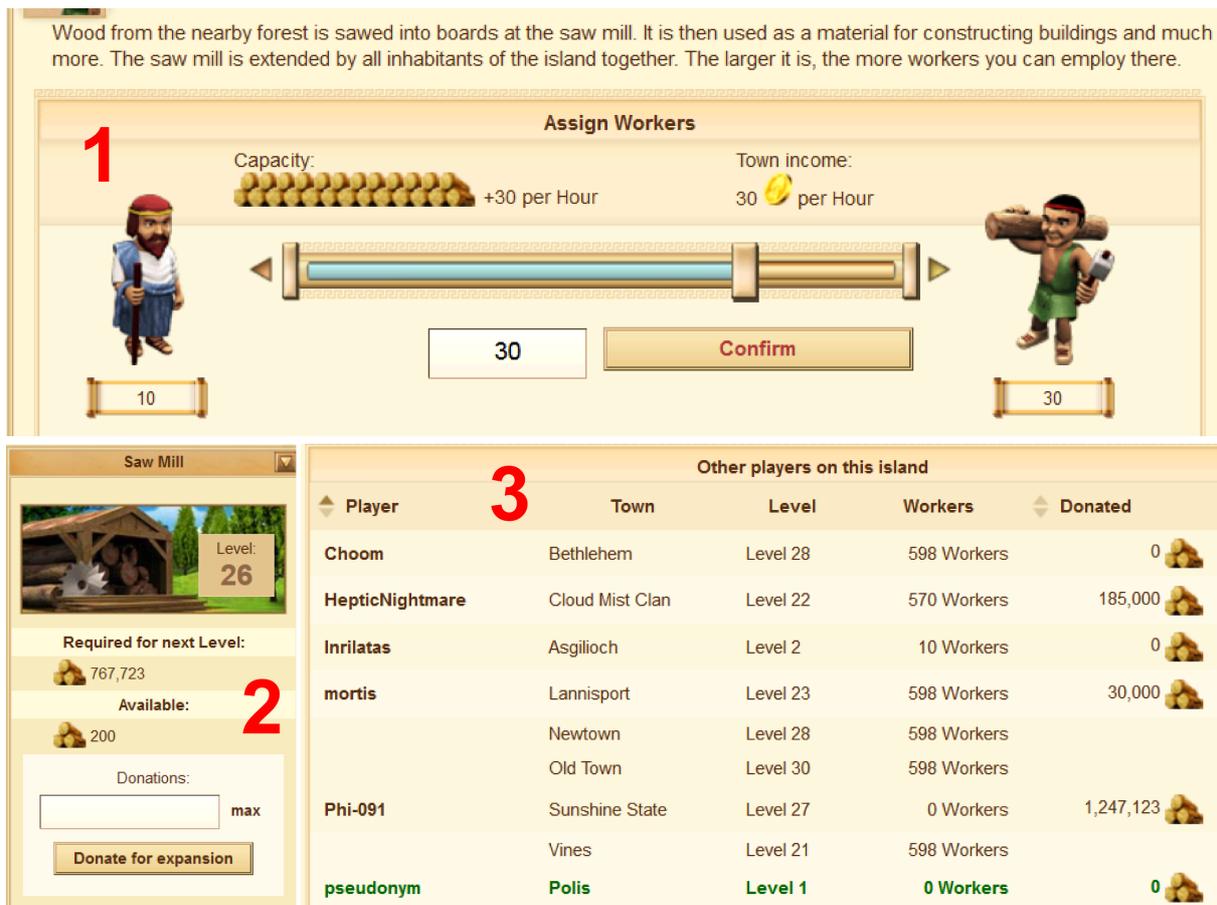


Abbildung 5. Screenshot der Hauptbestandteile der Gemeingutsituation in Ikariam

Das in dem Spiel vorkommende soziale Dilemma kann wie folgt genauer beschrieben werden. Erstens: alle drei Gemeingüter auf einer Insel (Sägewerk, Luxusgüter, Wunder) weisen die zwei grundlegenden Merkmale eines Gemeinguts auf: sie sind nicht-ausschließbar (*non-excludable*, d.h. jeder Spieler kann Arbeiter entsenden, um davon zu profitieren) und nicht subtrahierbar (*non-rivalrous*, d.h. nutzt ein Spieler den Rohstoff, so wird dessen Verfügbarkeit für andere Spieler dadurch nicht reduziert).

Zweitens: es gibt einen Schwellenwert (*threshold* auch *provision point* bzw. *step-level* genannt), da eine bestimmte Menge privater Ressourcen in das Gemeingut investiert werden muss, bevor dieses bereitgestellt werden kann. Jedes Gemeingut beginnt bei Stufe 1 und kann bis Stufe 50 gesteigert werden (Wunder lediglich bis Stufe 5). Höhere Stufen haben auch höhere Schwellenwerte, so dass auch mehr private Ressourcen notwendig sind, um diese Stufe zu erreichen (vgl. Anhang, Tabellen 33, 34, und 35 für die Kosten aller Stufen). Wird die nächste Stufe nicht erreicht, so wird das Gemeingut – höhere Effizienz – nicht zur Verfügung gestellt und die investierten Rohstoffe bleiben ohne jede Auswirkung. Es gibt keine Rückzah-

lung – die Zuwendungen verbleiben im Gemeingut, bis der Schwellenwert erreicht wird. Es ist nicht möglich, eine größere Summe als den Schwellenwert zu investieren (d.h., es gibt keine Erstattung, auch *rebate* genannt), da jede zusätzliche Investition automatisch auf den exakten Schwellenwert verringert wird.

Die Entscheidung, private Rohstoffe zu investieren, wird nacheinander getroffen, wobei jeder weiß, wie viel jeder Spieler bisher beigetragen hat (vgl. Abbildung 5). Die Informationen über die Investitionen, das Wissen über die erreichte Ausbaustufe des Gemeinguts sowie die bisher getätigten Investitionen werden nur den Spielern der jeweiligen Insel zugänglich gemacht, so dass eine Insel einer Gruppe in einem PGG-Laborexperiment entspricht. Spieler können über die Spitznamen, die sie sich selbst ausgesucht haben, identifiziert werden, bleiben aber grundsätzlich anonym, wenn sie dies wünschen – was ebenfalls analog zu herkömmlichen Experimenten ist. Obwohl Kommunikation zwischen den Spielern möglich ist, wird dies so gut wie nie genutzt. Die Spieler bleiben überwiegend auf einer Insel, so dass dieser Versuchsaufbau eher mit einem Partner-Treatment als einem Stranger-Treatment vergleichbar ist. Da 17 Orte auf einer Insel verfügbar sind, variiert die Gruppengröße zwischen $1 \leq n \leq 17$.

Jedes Individuum ist mit der Entscheidung konfrontiert, ob es seine privaten Ressourcen in ein Gemeingut investieren oder für sich behalten und etwa für die Verbesserung der eigenen Gebäude verwenden möchte. Die jedem Spieler zur Verfügung stehende Menge an Ressourcen hängt hauptsächlich von der Spieldauer, der Anzahl seiner Städte, der Anzahl der im Gemeingut eingesetzten Arbeiter sowie einiger besonderer Gebäude, welche die Ernteeffizienz erhöhen, ab. Das bedeutet, dass die zur Verfügung stehenden Ressourcen für das Investment zwischen den Spielern höchst unterschiedlich sind (vgl. Anhang G und Tabelle 28 für detaillierte Berechnungen der Auszahlungen).

Es gibt offenbar klare Anreize zu spielen, da mehr als 18 000 Spieler in diesem Datensatz länger als zwei Monate und zwischen 4 und 9 % der Spieler (je nach Land) länger als acht Monate (von möglichen zehn) teilgenommen haben. Zudem sind Spieler mit durchschnittlich etwa vier täglichen Logins ausgesprochen aktiv (Quelle: Ikariam-Umfrage 2014, siehe dazu Anhang G). Diese Zahlen zeigen, dass beträchtliche Zeit investiert wird, um an diesem Spiel teilzunehmen. Die Anreize sind offenbar weniger monetärer Art als die Zufriedenheit mit den Erfolgen innerhalb des Spiels, dem Spaß am Spiel sowie gemeinsame Zeit mit Freunden.

Das verwendete Online-Browser-Spiel stellt nahezu perfekte Daten über viele für die Arbeits-hypothesen (Abschnitte B 3, 5 und 7) notwendigen Parameter bereit. Jeder Datensatz besteht

aus 69 Variablen und stellt detaillierte Informationen über Spieler, Städte, Inseln sowie die Investitionen in das Gemeingut bereit (vgl. Tabellen 29, 30, 31 und 32 im Anhang für eine komplette Auflistung aller Originalvariablen). Der Rahmen des PGG (z. B. Ressourcen ernten) ähnelt Rahmen der realen Welt. Gleichwohl mag die *Validität* bzw. *Übertragbarkeit* der Ergebnisse dieses PGG in die reale Welt ein Einwand sein, obwohl zahlreiche Laborexperimente in Computersettings erfolgreich wiederholt werden konnten (Horton et al. 2011). Dies wurde auch für eine virtuelle Welt wie die hier vorgestellte gezeigt (Spiel eines PGG in einer virtuellen Welt, siehe Chesney et al. 2009). Die letztgenannte Studie bestätigt zudem Ergebnisse von Offline-Forschung zu reziprokem Verhalten. Darüber hinaus hat sich gezeigt, dass die Kommunikationsmuster, die Gefühle und die sozialen Interaktionen (Ducheneaut et al. 2006), ja sogar die Motivationen (Schultheiss 2007) der Computerspielteilnehmer sich nicht von den realweltlichen unterscheiden. Dies trifft auch auf das hier analysierte Spiel Ikariam zu: die Teilnehmer investieren eine beträchtliche Menge an Zeit (manchmal Jahre ihres Lebens), einige wenige auch echtes Geld, um bestimmte Ziele innerhalb des Spiels zu erreichen.

Im Folgenden werden einige grundlegende Eckdaten dieses Datensatzes vorgestellt, da er relativ komplex aufgebaut ist. Zunächst zu den Spielerzahlen. Zu jedem Messzeitpunkt ist die Zahl der Spieler exakt bekannt. Die untenstehende Tabelle 6 führt zu diesen Häufigkeiten Maximum, Minimum und Durchschnitt der Spielerzahlen für alle 28 Zeitpunkte auf, zu denen Daten zur Verfügung stehen (siehe auch Abbildung 6). Jeder Zeitpunkt wird im Folgenden als *Snapshot* bezeichnet. In viele Berechnungen bei Ikariam fließt der erste Snapshot jeweils nicht ein, da die Anfangsphase in vielerlei Hinsicht nicht repräsentativ für den Spielverlauf ist (siehe auch Abbildungen 63, 64 und 65). Ein Beispiel: der Anteil der Trittbrettfahrer für die deutschen Daten bewegt sich über 10 Monate hinweg zwischen 17% und 29% (Standardabweichung: 3,4) mit flach abfallendem Trend, beginnt aber mit einem Sprung von 45% für den ersten Snapshot auf 29%. Dies ist für alle anderen Länder sehr ähnlich. Es wird in jeder Abbildung gesondert darauf hingewiesen.

Die Türkei ist dabei mit Abstand der größte Datensatz, England der kleinste.

	Durchschnitt	Maximum	Minimum	Standardabw.
DE	2918	4292	786	976,10
EN	2257	3094	762	675,16
FR	3177	4422	921	992,11
GR	3740	5013	998	1164,18
TR	6112	9103	1770	2097,98

Tabelle 6. Deskriptive Statistik der Ikariam-Spielerzahlen im Ländervergleich

Die Spielerzahlen verändern sich über die Zeit in einem Muster, das für alle fünf Länder vergleichbar ist, wobei die Türkei die höchsten Teilnehmerzahlen verzeichnet (Abbildung 6). In allen Ländern steigt die Beteiligung bis zur Mitte des Beobachtungszeitraumes an, um danach wieder abzufallen. Dies ist mit dem Wechsel von Spielern auf andere, neuere Server zu erklären, da auf alten Welten die Aktivität generell und damit oft der Spielspaß abnimmt.

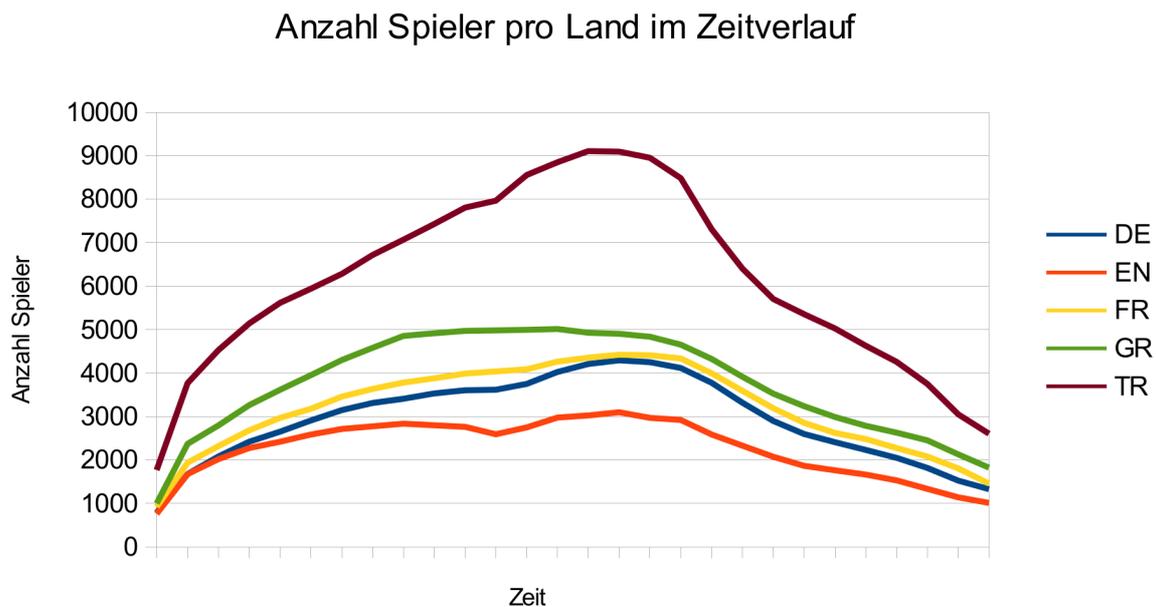


Abbildung 6: Entwicklung der Spielerzahlen im Zeitverlauf im Ländervergleich

Es ist wichtig zu wissen, wer sich in demographischer Hinsicht hinter den Spielern verbirgt. Der Datensatz ist zu männlichen Teilnehmern hin verschoben (83%), wobei das für Computerspiele keine ungewöhnliche Verschiebung darstellt (Schultheiss 2007). Die Altersverteilung ist etwas älter als in vergleichbaren Online-Computerspielen (Gomes & Carvalho 2011):

Altersgruppe	Prozentanteil
< 16 Jahre	9,7
16 bis 21 Jahre	28,3
22 bis 27 Jahre	7,5
28 bis 35 Jahre	16,4
> 35 Jahre	38,5

Tabelle 7. Altersverteilung bei deutschen Ikariam-Spielern (Daten aus eigener Online-Umfrage 2014, siehe Anhang G)

Der Mittelwert liegt bei 31 Jahren, der jüngste der 235 gültigen von insgesamt 428 Teilnehmern der Umfrage ist 12, der älteste 65 Jahre alt.

Da im Zentrum der folgenden Analysen die im Spiel vorhandenen öffentlichen Güter (Sägewerk, Luxusmine, Wunder) stehen, gebe ich die wichtigsten statistischen Kennwerte an (Tabelle 8). Da sich die späteren Analysen vor allem mit verschiedenen Strategien beschäftigen, die von egoistisch zu kooperativ reichen, wäre es denkbar, dass sich Gruppen von Kooperierenden einen kleinen Anteil an fast optimal ausgebauten Inseln erarbeitet haben bzw. dass man bei reinen Trittbrettfahrer-Inseln extrem niedrige Ausbaustufen beobachten kann. Die untenstehende Tabelle 8 zeigt die statistischen Kennwerte der Gemeingüter und die Abbildungen 7, 8 und 9 den Verlauf über die Zeit.

Land	Stufen des Gemeinguts Sägewerk			Stufen des Gemeinguts Luxusminen			Stufen des Gemeinguts Wunder		
	Mittelwert	Max	Min	Mittelwert	Max	Min	Mittelwert	Max	Min
DE	18,8	33	1	15,3	25	1	2,7	5	1
EN	17,4	36	1	14,0	25	1	2,7	5	1
FR	18,8	34	1	15,2	25	1	2,8	5	1
GR	18,0	30	1	14,8	23	1	3,1	5	1
TR	16,6	34	1	13,8	27	1	2,3	5	1

Tabelle 8. Deskriptive Statistik für den Ausbau der öffentlichen Güter

Aus dem Verlauf wird unmittelbar klar, dass mittlere Ausbaustufen dominieren; ausgeprägte

Extrema werden nicht sichtbar: zum einen müsste der Anstieg konvex und nicht konkav verlaufen und zum anderen die theoretisch möglichen Ausbaustufen von 50 erreichen. Dies ist nicht der Fall. Die höchste Ausbaustufe bei Wundern (5) wird zwar erreicht, doch lässt sich dies durch die vergleichsweise leichte Erreichbarkeit erklären. Der etwas andere Verlauf bei Wundern erklärt sich auch dadurch, dass Wunder bei weitem nicht so spielentscheidend sind wie Sägewerke und Luxusminen.

Verteilung der Ausbaustufen der Sägewerke

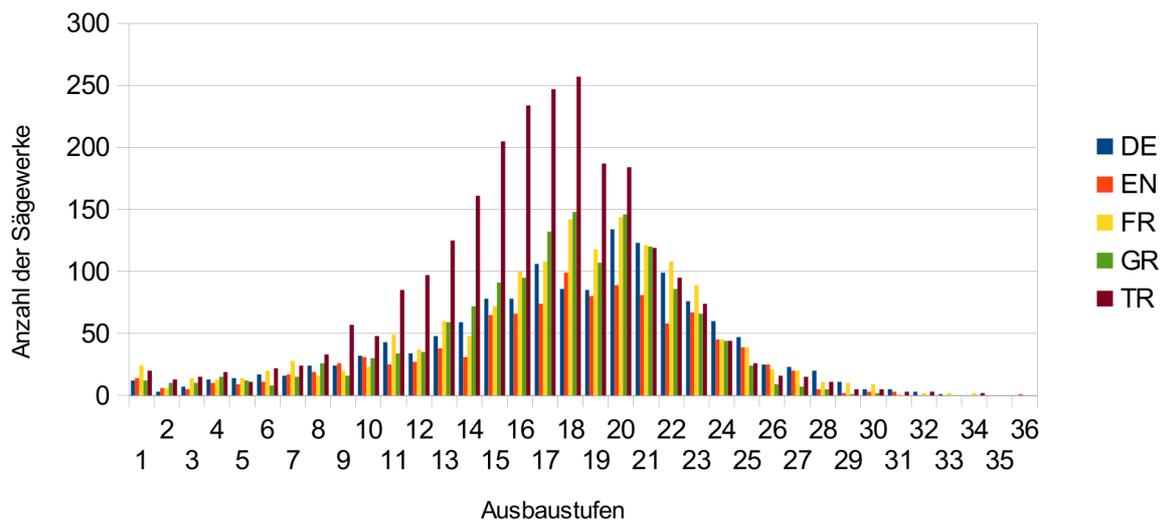


Abbildung 7. Verteilung der Ausbaustufen der Sägewerke im Ländervergleich am 13.2.2014 (28. Snapshot)

Ähnlich den Sägewerken ist auch die Verteilung der Luxusminen:

Verteilung der Ausbaustufen der Luxusminen

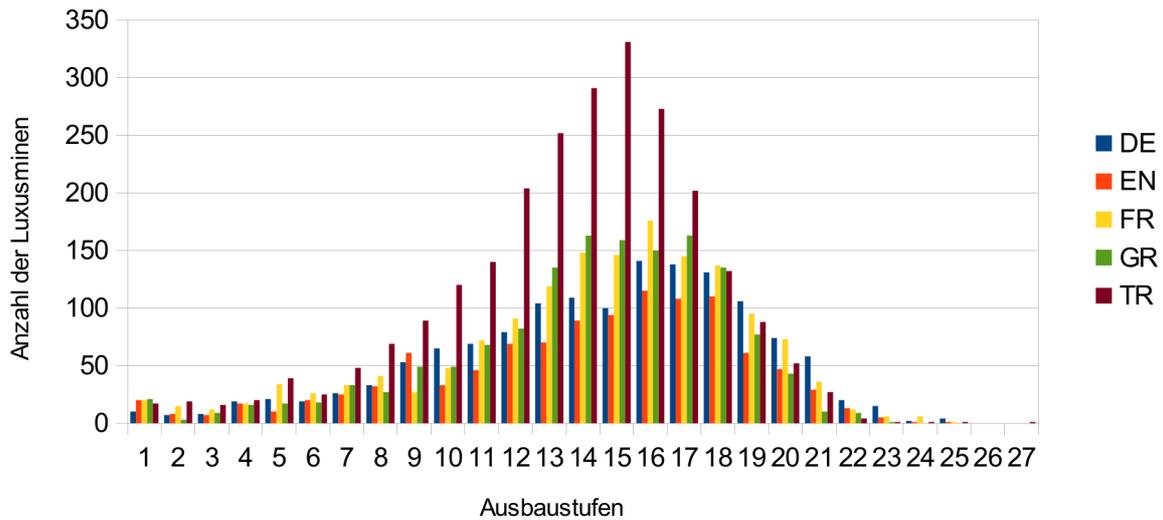


Abbildung 8. Verteilung der Ausbaustufen der Luxusminen im Ländervergleich am 13.2.2014 (28. Snapshot)

Es folgt der Ausbau der Wunder mit nur fünf Stufen:

Verteilung der Ausbaustufen der Wunder

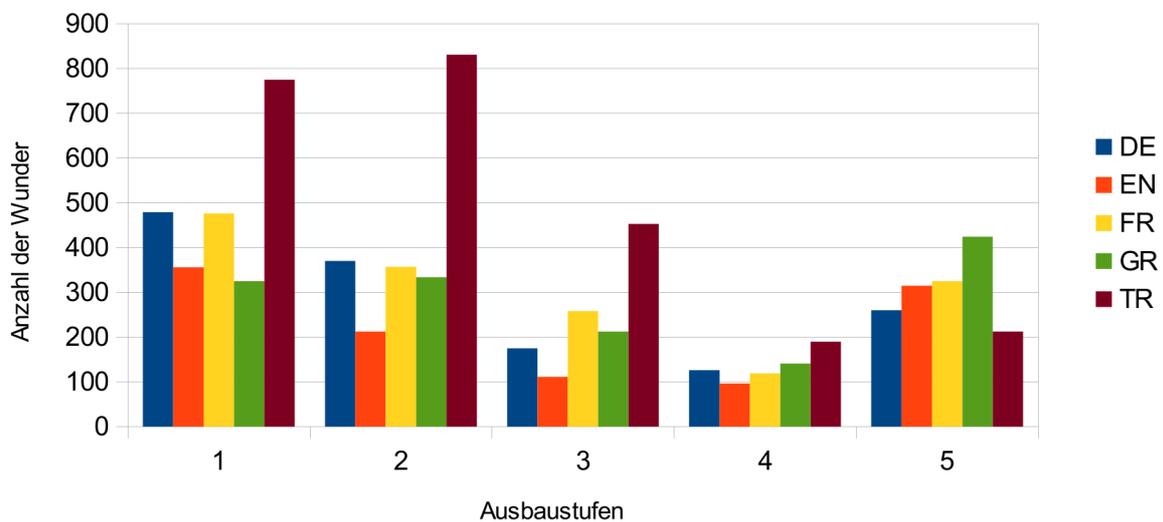


Abbildung 9. Verteilung der Ausbaustufen der Wunder im Ländervergleich am 13.2.2014 (28. Snapshot)

Der steilere Anstieg am Anfang der Ausbaustufen (0-10) spiegelt wohl die niedrigeren Kosten und den *step return* über 1 (siehe Anhang G) wider. Die höchsten Ausbaustufen von 35 (Sägewerk) und 25 (Luxusminen) von 50 maximal möglichen ergeben sich zum einen aus dem

Fehlen extrem kooperativer Strategien, zum anderen aus der fehlenden Spielzeit (da die Welten mit zehn Monaten noch nicht allzu alt sind) und fehlenden Notwendigkeit, über diese Ausbaustufen hinaus auszubauen. Je weiter die Verteilung nach rechts verschoben ist, desto mehr haben die Spieler in den Ausbau der Gemeingüter investiert. Aus der im Vergleich rechtsverschobenen Verteilung von England gegenüber der Türkei lässt sich folgern, dass England im Mittel kooperativere Teilnehmer hat. Die in der Anzahl abweichenden Zahlen für die Türkei ergeben sich einfach aus der hohen Zahl der türkischen Spieler, was in einer größeren Zahl besiedelter Inseln resultiert.

Auch bei dieser Fragestellung ist wiederum der Blick auf den zeitlichen Verlauf wichtig, da ungewöhnliche Verlaufsmuster Einfluss auf Hypothesen und Analysen hätten. Man beobachtet aber das Gegenteil (siehe Abbildung 10), nämlich einen Verlauf ohne Brüche. Der Ausbau der Ressource ist demnach gleichmäßig über die Zeit, mit einem stetigen Anstieg des Ausbaus von einem Stützzeitpunkt zum nächsten.

Entwicklung der Ausbaustufen der Sägewerke über die Zeit im Ländervergleich

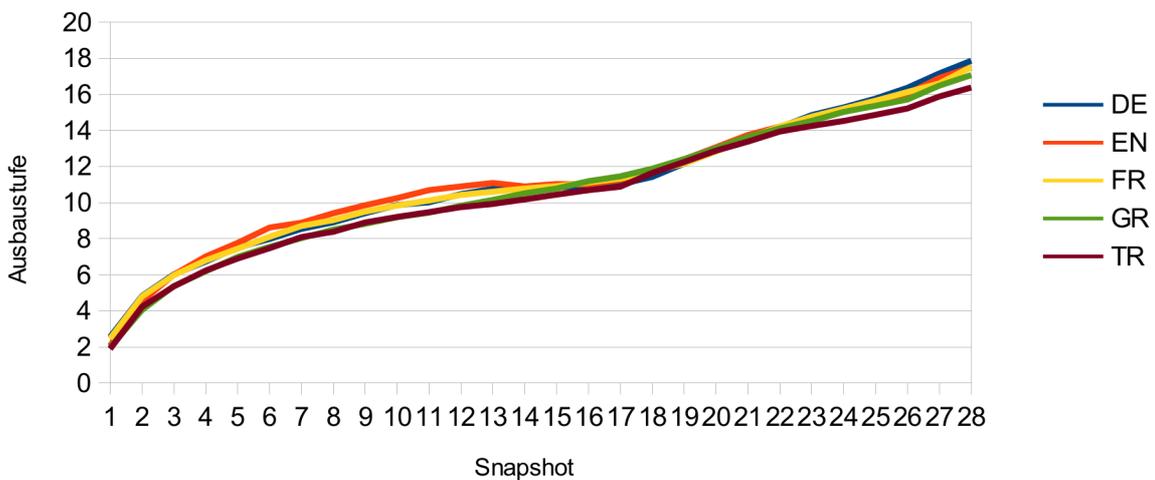


Abbildung 10. Entwicklung der Ausbaustufen der Sägewerke im Ländervergleich über die Zeit

Es zeigt sich, dass der Ausbau des öffentlichen Guts (die beiden anderen öffentlichen Güter zeigen einen fast identischen Verlauf) im Ländervergleich extrem ähnlich verläuft und auch über die Zeit keine Besonderheiten aufweist. Nach einem etwas schnelleren Ausbau zu Beginn, da die Stufen leicht zu erreichen sind, setzt nach einem kleinen Knick der stete Anstieg der ersten Monate wieder ein.

Zudem besteht ein enger Zusammenhang zwischen dem Ausbau eines Gemeinguts und dem

der anderen. Ich finde sehr hohe Korrelationen ($> 0,9$) zwischen dem Ausbau des Sägewerks auf einer Insel, dem wichtigsten Gemeingut, und dem der Luxusminen nach zehn Monaten:

Land	Pearson-Korrelationskoeffizient Sägewerk / Luxusmine ($p < 0,01$)	Pearson-Korrelationskoeffizient Sägewerk / Wunder ($p < 0,01$)	Pearson-Korrelationskoeffizient Luxusmine / Wunder ($p < 0,01$)
DE	0,91	0,55	0,55
EN	0,90	0,57	0,57
FR	0,93	0,57	0,56
GR	0,90	0,57	0,57
TR	0,90	0,50	0,50

Tabelle 9: Korrelationen zwischen den Ausbaustufen der Gemeingüter nach zehn Monaten Spielzeit (Daten des letzten Snapshots, am 13.2.2014)

Die Korrelationskoeffizienten zwischen Sägewerk bzw. Luxusminen mit jeweils dem Wunderausbau sind ebenfalls stark und sehr ähnlich zwischen den Ländern; sie schwanken zwischen 0,50 und 0,57 schwanken.

2.1 Definition der Spielertypen

Frühere Studien haben Spielertypen in Trittbrettfahrer, konditional Kooperierende und Dreiecks-Kooperierende (*triangle-shaped cooperators*, z. B. Fischbacher et al. 2001) eingeteilt. Einige neuere Artikel fügen Altruisten hinzu (Rustagi et al. 2010). Zahlreiche Studien verwenden die Klassifikation von Fischbacher et al. (2001) und definieren Trittbrettfahrer als Spieler, die zu keiner Zeit einen Beitrag leisten. Konditional Kooperierende haben nach diesem Schema entweder einen signifikanten positiven Spearman-Korrelationskoeffizienten zwischen den eigenen Einzahlungen und dem der anderen *oder* einen ansteigenden, schwach monotonen Trend zwischen eigenen und steigenden Einzahlungen anderer (Kocher et al. 2008). Dreiecks-Kooperierende erhöhen ihre Einzahlungen bis zu einem gewissen Punkt und verringern sie dann; dieser Punkt ist nicht genauer spezifiziert.

Die Methoden zur Bestimmung der Spielertypen variieren: Es gibt die Strategie-Methode (für Details siehe Fischbacher et al. 2001), die direkte-Antwort-Methode (d.h. das Verhalten in einem PGG, siehe Chaudhuri 2011) oder andere Algorithmen (z.B. Kurzban & Houser 2005). Da mir die Strategie-Methode nicht zur Verfügung steht, bestimme ich Spielertypen über ihr Verhalten im Spiel, vergleichbar der direkten-Antwort-Methode. Es wurde gezeigt, dass beide Methoden sehr ähnliche Ergebnisse erbringen und daher austauschbar sind (Fischbacher et al. 2012).

Da frühere Studien abweichende Definitionen von Spielertypen verwenden, können deren Ergebnisse untereinander nicht oder nur teilweise verglichen werden. Um meine Ergebnisse vergleichbar zu machen, gibt diese Arbeit daher für Spielertypen nicht nur passende eigene Definitionen an, sondern klassifiziert auch gemäß der bisher verwendeten Schemata. Für konditional Kooperierende bzw. Reziprokatoren werden folgende sechs alternative Definitionen verwendet:

- Spieler, deren Einzahlungen signifikant positiv (jeweils auf dem 5%, 1% oder 0,1 %-Niveau, was drei verschiedene Definitionen darstellt) mit denen der anderen korrelieren (diese Definition folgt Fischbacher et al. 2001)
- Spieler, die ihre Einzahlungen immer in Richtung (positiv oder negativ) der Gruppe modifizieren (über alle 28 snapshots, d.h. etwa 10 Monate) (diese Definition soll die grundlegende Idee der Reziprozität widerspiegeln)
- Spieler, die nicht mehr als 10 % (positiv oder negativ) pro Snapshot in ihren Einzahlungen vom Gruppendurchschnitt abweichen (diese Definition soll ebenfalls die grundlegende Idee der Reziprozität widerspiegeln)
- Spieler, die mehr als Trittbrettfahrer ($> 0\%$) und weniger als Altruisten ($\geq 50\%$) einzahlen (diese Definition soll die Vergleichbarkeit zwischen allen Spielertypen erhöhen, indem ein Kontinuum erzeugt wird)

Diese Definitionen stellen jeweils Alternativen zur in der Arbeit verwendeten letzten Definition dar. Damit werden Vergleiche der Häufigkeiten mit anderen Studien möglich. Ein zusätzlicher Nutzen der Verwendung verschiedener Messmethoden ist ein Zugewinn an Robustheit der Ergebnisse.

Für Trittbrettfahrer werden die folgenden alternativen vier Definitionen verwendet:

- Spieler, die während des gesamten Spiels überhaupt nichts in das Gemeingut einzahlen (diese Definition folgt Fischbacher et al. 2001)
- Spieler, die in allen Snapshots durchschnittlich weniger als 3% ihrer Gelder einzahlen
- Spieler, die in allen Snapshots durchschnittlich weniger als 5% ihrer Gelder einzahlen
- Spieler, die in mehr als 50% der Snapshots nichts einzahlen (diese Definition folgt Rustagi et al. 2010).

Die verschiedenen Obergrenzen (0, 3 und 5%) sind dabei auch als Test für Robustheit angelegt – durch den Vergleich dieser drei Häufigkeiten sind bereits gute Abschätzungen der Trittbrettfahrerzahl möglich.

Analog zu den Trittbrettfahrern werden Kooperierende als Spieler definiert, die mehr als einen bestimmten Betrag in das Gemeingut einzahlen:

- Spieler, die in allen 28 Snapshots durchschnittlich mehr als 40% der ihnen zur Verfügung stehenden Gelder zum Gemeingut beitragen
- Spieler, die in allen 28 Snapshots durchschnittlich mehr als 50% der ihnen zur Verfügung stehenden Gelder zum Gemeingut beitragen
- Spieler, die in allen 28 Snapshots durchschnittlich mehr als 60% der ihnen zur Verfügung stehenden Gelder zum Gemeingut beitragen
- Spieler, die mehr als 50% ihrer Ressourcen in mehr als 50% der Snapshots beitragen (diese Definition folgt Rustagi et al. 2010, analog zu Trittbrettfahrern).

Aufgrund der vorliegenden Datenstruktur ist es nicht möglich, ein monotonen Ansteigen oder Abfallen für konditional Kooperierende und Dreieckskooperierende zu berechnen.

Diese Definitionen stellen unterschiedliche Methoden dar, die Häufigkeit von Spielertypen zu messen. Damit sind komplementär zu einander und unabhängig untereinander. Diese Arbeit übernimmt die gebräuchliche Definition für Trittbrettfahrer (0% Einzahlungen). Für Altruisten gibt es nur wenige Studien und zwei Definitionen (100% bzw. mehr als 50% in 50% der Fälle). Die erste Definition ist meines Erachtens viel zu strikt für realistischere Umgebungen. Es wird daher die niedrigere Grenze von 50% der eigenen Ressourcen an Investitionen in das Gemeingut verwendet. Mit diesen beiden Entscheidungen liegt es nahe, für konditional Kooperierende eine ähnliche Definition zu wählen, um die Vergleichbarkeit zu gewährleisten. Dies wäre bei einer Mischung der Methode positiver Korrelationen für konditional Kooperierende und Ober- bzw. Untergrenzen für Trittbrettfahrer bzw. Altruisten nicht gegeben. Daher gibt diese Arbeit für Ikariam die Häufigkeiten für alle Definitionen an, verwendet aber in den eigenen Analysen für konditional Kooperierende die Definition einer Investition zwischen 1% und 49% ihrer Ressourcen – damit liegen diese zwischen den beiden Extremtypen in einem Kontinuum.

2.2 Definition der abhängigen Variablen (Erfolgsmaß und Investitionsbereitschaft)

Das Erfolgsmaß in *public goods games* ist der Gesamtprofit, der am Ende von einem Spieler erzielt wird. In Laborexperimenten geht es für die Spieler demnach darum, den eigenen Verdienst zu maximieren.

Im Gegensatz dazu kann man in Ikariam kein Geld verdienen. Man kann allerdings, was nur ca. 4% der Spieler tun, eigenes Geld in den Erfolg im Spiel investieren – dadurch genießt man bestimmte Vorteile (schnellere Bauzeit, usw.). In Ikariam wird der Erfolg durch den *Erfolg im Aufbau*, also im Vorankommen im Spiel gemessen. Dafür ideal geeignet wäre der *Highscore*, der sich aus verschiedenen Indices zusammensetzt. Diese Daten liegen aus Datenschutzgründen nicht vor. Aus diesem Grund verwendet diese Arbeit den Rathauslevel pro Spieldauer (abgekürzt RHL/SD). In jeder Stadt ist das Rathaus das wichtigste Gebäude. Es ist das Maß für die Größe jeder Stadt und ein großer Teil der Bautätigkeit läuft auf Zulieferung dafür hinaus. Es ist ein sehr gut geeignetes Maß für den Erfolg, muss allerdings um die für den Aufbau benötigte Spielzeit dividiert werden. Die untenstehende Abbildung (Abbildung 11) zeigt die Verteilung der abhängigen Variable (des Erfolgsmaßes) beispielhaft für den letzten Snapshot (13.2.2014) über alle Länder, also nach zehn Monaten Spielzeit. Die enge Verteilung zeigt, dass der Großteil der Spieler ähnlich gut spielt, der kaum sichtbare, lange rechtslastige Schwanz, dass es einigen sehr wenigen Spieler durch Geschick (und/oder Einsatz von Geld) gelingt, viel besser als die Masse zu spielen.

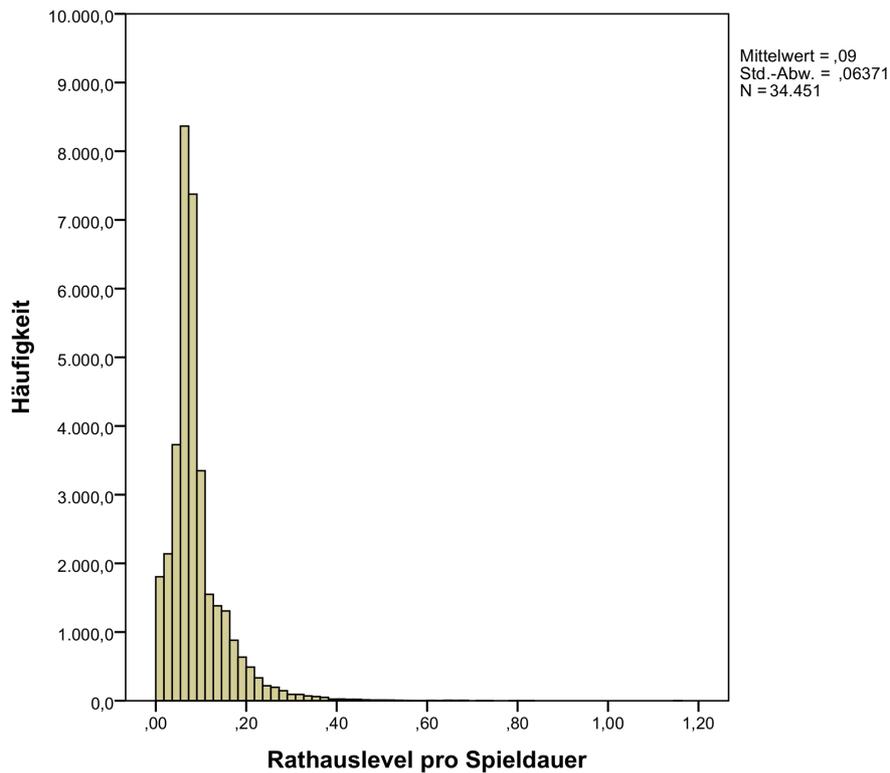


Abbildung 11. Häufigkeitsverteilung Spieler am 13.02.2014 bezüglich der abhängigen Variable Rathauslevel / Spieldauer in Ikariam über alle Länder

Die zugehörigen Kennwerte lauten: Median = 0,799; Standardabweichung = 0,662; Varianz = 0,04; Schiefe = 2,568; Kurtosis = 11,087; Minimum = 0; Maximum = 0,722.

Eine zweite abhängige Variable – die Einzahlungen in das Gemeingut – setzt sich aus der Summe aller Investitionen in alle drei Gemeingüter zusammen und wird als Anteil der dem Spieler insgesamt zur Verfügung stehenden Ressourcen in Prozent ausgedrückt:

$$\text{Investitionen in alle Gemeingüter (absolut)} / \text{Gesamtzahl vorhandener Ressourcen (absolut)}$$

Der Ausdruck in prozentualen Anteilen ist zu Vergleichszwecken auch deshalb nötig, weil die Zahl der Spieler auf den verschiedenen Welten unterschiedlich ist.

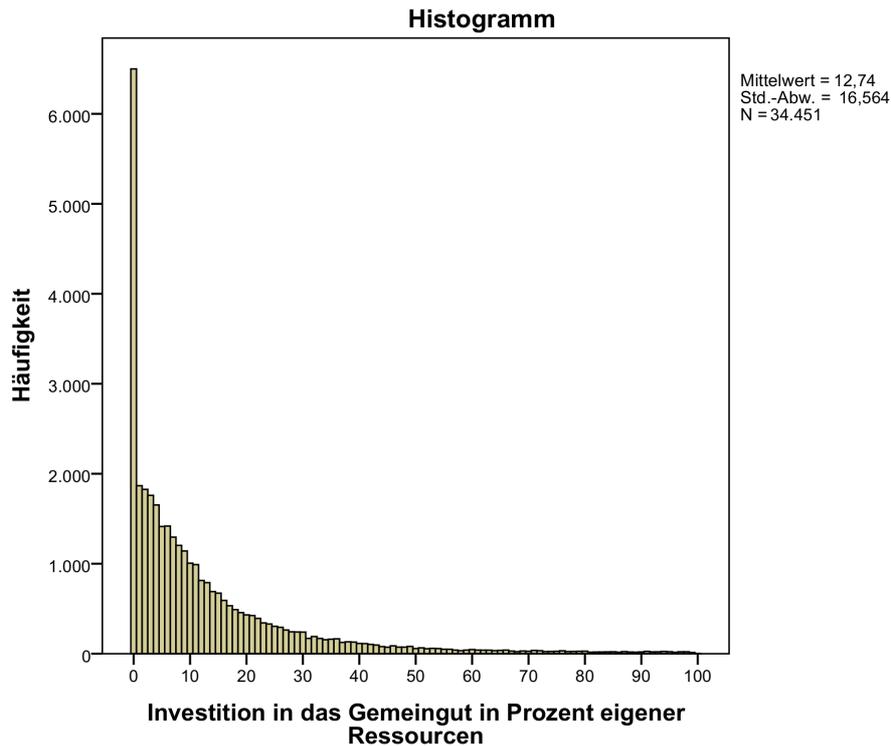


Abbildung 12: Häufigkeitsverteilung der Investitionen in die Gemeingüter: Prozentualer Anteil der investierten Ressourcen am 13.2.2014 (28. Snapshot) über alle Länder

Die zugehörigen Kennwerte lauten: Median = 7,00; Standardabweichung = 15,879; Varianz = 252,151; Schiefe = 2,233; Kurtosis = 6,174; Minimum = 0; Maximum = 100.

3 Beschreibung der Arbeitsschritte zur Datentransformation

3.1 PGG-Laborexperimente

Die Daten aus den analysierten PGG sind teilweise im Internet zugänglich. Um die Kalkulation der Ergebnisse nachvollziehen zu können, sind Links zu diesen Rohdaten im Anhang G zu finden. Die Arbeitsschritte der Umkodierung sind die folgenden:

1. Die Rohdaten wurden in MS Excel 2013 für den Export vorbereitet (Dezimalzeichen, Anordnung, Trennung der benötigten und nicht benötigten Spalten).
2. Duplikate, die durch die mehrfache Berücksichtigung von Strafen für andere Analyse-zwecke angegeben waren, wurden entfernt.
3. Die Daten wurden als CSV gespeichert und exportiert.
4. Die Daten wurden mit folgender Struktur in MySQL importiert:

- Treatment
 - Gruppe
 - Runde
 - Versuchsperson
 - Einzahlung pro Runde
 - Verdienst vor Strafen
 - Verdienst nach Strafen
5. Die folgenden Spalten wurden hinzugefügt: ID, Partner oder Stranger-Treatment (Treat1), Strafen oder keine Strafen (Treat2), Rundenzahl (Treat3), Gesamteinzahlungen (Total_Contributions), Gesamtverdienst vor Abzug der Strafkosten (Totalprofit_vorP), Gesamtverdienst nach Abzug der Strafkosten (Totalprofit_nachP), Spielertyp (0-6), Gruppentyp (1-7), erreichter prozentualer Anteil des maximal möglichen Profits (Proz_Max_Profit), erreichter prozentualer Anteil des maximal möglichen Profits pro Runde (Proz_Max_Profit_proRunde), erreichter prozentualer Anteil des sozialen Optimums (Proz_Soz_Opt), gegebene Strafpunkte (P), gesamt gegebene Strafpunkte (Total_P), erhaltene Strafpunkte (RP), gesamt erhaltene Strafpunkte (Total_RP), Straftyp, Bestrafttyp und Gruppenstraftyp.
 6. Die hinzugefügten Kennwerte wurden in SQL berechnet.
 7. Die Datenzusammenstellungen für die einzelnen Analysen wurden in SQL durchgeführt.
 8. Die Daten wurden nach MS Excel / OpenOffice Calc exportiert. Grafiken wurden in Open-Office erstellt und mit Gimp 2 (Version 2.8.10) bearbeitet.
 9. Die Daten wurden nach SPSS (Version 22) und R (Version 3.1.2) exportiert. Alle statistische Tests wurden damit durchgeführt. Grafiken wurden mit SPSS erstellt und mit Gimp 2 (www.gimp.org) bearbeitet.

Bei der Angleichung der Datensätze traten – durch die Umkodierung – einige Probleme auf. So ist im Datensatz von Fehr & Gächter (2000) jede Versuchsperson dreimal gelistet – diese Duplikate mussten gelöscht werden. Das wurde dadurch erschwert, dass alle Versuchspersonen pro Session die gleiche ID tragen – eine eindeutige Zuordnung musste gefunden werden. Zusätzlich sind die Verdienste nach Abzug der Strafkosten nicht korrekt kalkuliert. Dies betrifft allerdings nur Session zwei von fünf. Die Verdienste wurden nach der Formel

(Verdienst vor Strafkosten – (10% pro Strafpunkt) – Kosten eigene Bestrafung) korrigiert. Außerdem wird im Artikel angegeben, dass der Verdienst pro Runde nie negativ werden kann, d.h. negative Verdienste werden auf Null gesetzt. In den Daten wird dies allerdings nicht durchgeführt. Auch dies wurde korrigiert.

Zudem stellte sich der Datensatz von Fehr & Gächter (2000) als unvollständig heraus, d.h. in der letzten Partner-Session sind nicht alle Versuchspersonen erschienen. Dies wird allerdings im Artikel nicht erwähnt. Zudem sind die Gruppen-IDs nicht eindeutig, was ebenfalls zugeordnet werden musste.

Die umkodierte Datensätze wurden mit den Berechnungen in den Artikeln verglichen, um eventuelle Fehler zu finden und zu korrigieren. Ebenso wurde mit allen neu erzeugten Variablen verfahren: stichprobenartig wurden manuelle Überprüfungen in allen Datensätzen vorgenommen.

3.2 Ikariam

Für die Umwandlung der Daten und die Berechnungen zur Analyse wurde SQL (*structured query language*, <http://dev.mysql.com/>) benutzt (siehe Anhang G), für die Rohdaten R 3.1.2 (<http://www.r-project.org>). Einige hundert neue Variablen wurden angelegt, neue Werte berechnet, sowie Felder verbunden und geteilt. Die meisten Umwandlungen betreffen die Berechnung der Variablen (z.B. die Gesamtwertberechnung auf Basis der einzelnen Baumaterialien) und die Berechnung spezifischer Parameter (z.B. Spenden pro Insel; durchschnittliche Spendenhöhe pro Datensatz). Maximum, Minimum, Durchschnitt, Standardabweichung, Anzahl und Summe jeder genauer analysierten Variable werden berechnet. Der Export der Daten für Grafiken erfolgte über Open-Office Calc.

Um sicherzugehen, dass nur ernsthafte Spieler, die langfristige Ambitionen hegen und nicht-zufällige Strategien verfolgen, in die Berechnung aufgenommen werden, wurden jene Spieler, die weniger als 4 von 28 Malen (< 14% Teilnahmerate) spielten, gestrichen. Dieser Schritt verkleinerte den Datensatz auf etwa 89% der Ausgangsgröße. Wenn ein Spieler zwei oder mehr Städte auf einer Insel hatte, wurden diese als eine Stadt gewertet, da sonst eine eindeutige Zuordnung für weitere Berechnungen nicht mehr gegeben war. Es wurden keine anderen Schritte unternommen, um den Datensatz zu verändern. Es gibt eine kleine Unschärfe (0-3%) im Datensatz: Da es den Spielern gestattet ist, ihre Städte in der Zeit zwischen den Snapshots zu bewegen, ist für einige wenige Fälle nicht klar, welche Investition in das Gemeingut von wem stammt. Der verwendete Datensatz sammelt Informationen für jede Spieler-Insel-Kombi-

nation für jeden der 28 Snapshots. So wird eine klare Differenzierung zwischen Strategien eines Spielers in unterschiedlichen Kontexten ermöglicht.

Die Herkunft der Spieler wurde über ihre IP-Adresse bestimmt. Es gibt vier homogene Server: In der Türkei, Griechenland, Frankreich und Deutschland kommen 96%, 95%, 81% und 81% der Spieler aus dem jeweiligen Land. Das steht im deutlichen Gegensatz zum englischen Server, wo lediglich 11% der Spieler aus England, und je 10% aus den USA, Vietnam und den Philippinen kommen; die verbleibenden 59% verteilen sich auf weitere 37 Länder.

Alle spezielleren Kenndaten, die in direktem Bezug zu den in D 1, 2 und 3 durchgeführten Analysen stehen, sind diesen Kapiteln zugeordnet.

4 Methoden

Die verwendeten statistischen Prüfverfahren sind vor allem Tests auf *Unterschiede*, z. B. zwischen Spielertypen, Gruppentypen oder Ländern. Die Tests auf Unterschiede, die üblicherweise angewendet werden (T-Tests, ANOVA, *repeated measures* ANOVA) sind jedoch nur unter bestimmten Voraussetzungen anwendbar. Dazu gehören unter anderem multivariate Normalverteilung, Homogenität der Kovarianzmatrizen und Unabhängigkeit (Bortz & Schuster 2010).

Sowohl die Ikariam- als auch die PGG-Daten erfüllen diese Voraussetzung für die jeweiligen Erfolgsmaße *nicht*. Das betrifft Rathauslevel pro Spieldauer, Verdienst vor bzw. nach Abzug der Strafkosten pro Runde sowie den Gesamtverdienst vor und nach Abzug der Strafkosten. Keine dieser abhängigen Variablen ist normalverteilt, was sich sowohl in den zugehörigen Q-Q-Diagrammen als auch in den trendbereinigten Q-Q-Diagrammen deutlich visuell zeigt.

Zusätzlich sind Kolmogorov-Smirnov-Tests für jede genannte Variable auf dem 0,001-Niveau signifikant. Dies gilt bei den PGG für alle fünf Datensätze, bei Ikariam für alle fünf Länder. Daraus folgt, dass nicht-parametrische Testverfahren anzuwenden sind.

Q-Q-Diagramm von Normal von Investition in das Gemeingut in Prozent eigener Ressourcen

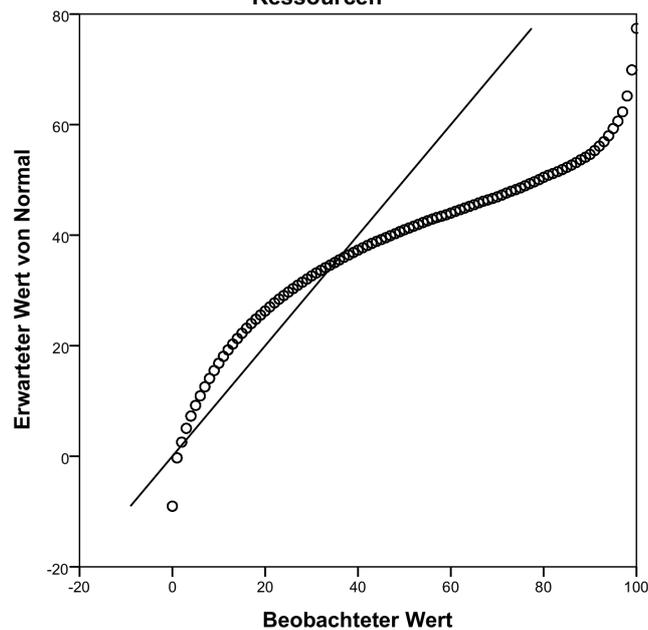


Abbildung 13: QQ-Diagramm für Investitionen in das Gemeingut in Prozent eigener Ressourcen über alle Länder für den letzten Snapshot (13.2)

Für Testverfahren, die zeitliche Zusammenhänge berücksichtigen und wegen ihrer Teststärke vorzuziehen sind (etwa *repeated measures* ANOVA), fällt zur Verletzung der Normalverteilungsannahme auch Mauchlys Test auf Sphärizität signifikant aus. Für Tests auf Unterschiede zwischen den Zeitpunkten verwende ich deshalb die nicht-parametrische Alternative zur wiederholten ANOVA, den Friedman Repeated Measures Analysis of Variance on Ranks-Test.

Studie 2 (Abschnitt D 1 und 1.2) differenziert Spieler in sieben Spielertypen, je nachdem, wie hoch deren Investition in das Gemeingut ist. Es wäre wünschenswert, diese Klassifizierung durch andere Verfahren zu bestätigen. Allerdings sind auch hier die gängigen Klassifikationsverfahren wie Diskriminanzanalyse und Clusterverfahren auf Grund der nicht normal verteilten abhängigen und unabhängigen Variablen (Investition in das Gemeingut in Prozent) nicht geeignet, da beide Analyseverfahren multivariate Normalverteilung voraussetzen. Zusätzlich arbeitet das *k-means* Clusterverfahren mit der quadrierten euklidischen Distanz: damit ist implizit vorausgesetzt, dass die unabhängigen Variablen etwa die gleiche Größenordnung haben – was bei der Vielzahl und Unterschiedlichkeit der Variablen in Ikariam auch durch Transformationen kaum erreicht werden kann. Weiterhin wird auf das Minimum für die quadrierte Fehlersumme (*error sum of squares*) optimiert, d.h. die Kovarianzmatrix zwischen Variablen innerhalb einer Klasse sollte etwa gleich sein. Auch diese Bedingung ist in den Ikariam-Daten nicht erfüllt. Von daher scheiden diese Verfahren für eine zusätzliche Absiche-

rung aus (für weitere inhaltliche Gründe siehe Abschnitt D 1).

Für Ikariam liegen die Daten in fünf unabhängigen Datensätzen getrennt nach Ländern vor. Für eine Erhöhung der Teststärke ist es wünschenswert, die Daten, wann immer möglich, zusammen zu legen. Das setzt voraus, dass es keine signifikanten Unterschiede zwischen ihnen gibt. Für diese Zwecke wird für den letzten Snapshot, den 13.2 für alle Länder auf Unterschiede (Kruskal-Wallis-Test) geprüft. Da diese Tests für alle verwendeten abhängigen Variablen (Spieldauer, Prozentsatz der investierten Ressourcen in das Gemeingut, Erfolgsmaß (Rathauslevel pro Spieldauer) und korrigiertes Erfolgsmaß (Rathauslevel pro Spieldauer pro Stadtanzahl) signifikante Unterschiede zwischen den Ländern aufweisen, ist eine Zusammenlegung aller Länder für keine abhängige Variable möglich (Kruskal-Wallis, $df = 4$, $n = 34451$, $p < 0,001$). In den korrigierten Post-hoc-Tests erweisen sich alle paarweisen Vergleiche für die Spieldauer mindestens auf dem 0,05-Niveau als signifikant, bis auf FR-GR; bei der Investition in das Gemeingut mindestens auf dem 0,03-Niveau, bis auf GR-DE, DE-EN, EN-FR und für die beiden Erfolgsmaße mindestens auf dem 0,001-Niveau, bis auf DE-TR. Von daher werden die Ergebnisse im Folgenden immer separat pro Land angegeben.

D Ergebnisse und Diskussion

Der folgende Teil stellt die Ergebnisse zu den Arbeitshypothesen aus Teil B vor und diskutiert sie. Aus Gründen der Lesbarkeit ist die Diskussion den jeweiligen Resultaten zugeordnet.

1 *Ergebnisse und Diskussion Studie 1: Differenzierung von Spielertypen*

1.1 Resultate für Hypothesen 1 und 2: Differenzierung von Spielertypen

Um über die bisherigen Ergebnisse hinauszugehen, die alle aus öffentlichen Gut-Spielen aus dem Labor stammen, präsentiere ich im Folgenden Daten zu mehr als 18 000 Personen aus fünf Ländern, die in einem Online-Spiel ein PGG spielen. Zudem kalkuliere ich die Spielertypen anhand möglichst vieler Definitionen aus der Literatur, um die Vergleichbarkeit zwischen den Forschungsergebnissen zu erhöhen. Damit hofft diese Arbeit, vor allem mehr Licht auf die Strategie konditional Kooperierender zu werfen.

Ich beginne mit einem Überblick über die bislang berichteten Häufigkeiten der Spielertypen in Tabelle 10. Diese beruht auf der Basis der Spielertyp-Definitionen (vgl. Tabelle 3).

Studie und Land	% Konditional Kooperierende	% Trittbrettfahrer	% Dreiecksform	% Kooperierende	% Andere (nicht klassifizierbar)
Fischbacher et al. 2001, Schweiz	50	30	14	N/A	6
Kurzban & Houser 2005, USA	63	20	N/A	13	4
Rustagi et al. 2010, Äthiopien	34	11	3	14	38
Fischbacher & Gächter 2010, Schweiz	55	23	12	N/A	10
Kocher et al. 2008, USA	81	8	0	N/A	11
Kocher et al. 2008, Japan	42	36	11	N/A	11
Kocher et al. 2008, Österreich	44	22	11	N/A	22
Chaudhuri & Paichayontvijit 2006, Neuseeland	62	16	9	7	6
Burlando & Guala 2005, Italien	35	32	N/A	18	15
Herrmann & Thöni 2009, Russland	56	6	7	N/A	31
Durchschnitt	52	20	8	15	13

Tabelle 10. Vergleich der Anteile an Spielertypen in Prozent aus 8 Studien zu konditionaler Kooperation (Messmethoden sind teilweise unterschiedlich, Typen sind für diesen Überblick teilweise zusammengefasst)

Für einen ersten Überblick über die Vergleichbarkeit der verwendeten Studien und die Spielertypen nach Studie gibt Abbildung 14 eine Darstellung der Spielertyp-Häufigkeiten in den genauer analysierten fünf PGG. Wie in Abschnitt C 1 beschrieben, wird ein einheitliches Merkmal gewählt ($< X$ % der Investitionen, siehe Tabelle 3 und Abschnitt C 2), um ein Kontinuum zwischen den Spielertypen zu gewährleisten. Zudem sind die sich ergebenden Häufigkeiten zwischen den Methoden vergleichbar. Die Spielertypen sind wie in der Legende beschrieben, als Kontinuum zu verstehen, das von 0% (T) bis hin zu über 50% Beitrag (K) reicht. Mit der Ausnahme von Nikiforakis (2008) kann man von einer guten Übereinstimmung der Spielertypen-Häufigkeiten über die Studien hinweg sprechen. In Nikiforakis (2008) sind die Typen dagegen sehr viel gleichmäßiger verteilt als in allen anderen Studien.

Schlüsselt man nun nicht nach Studien, sondern nach *Treatments* auf, die zwischen den Studien vergleichbar sein sollten, dann ergibt sich die folgende Verteilung (siehe Abbildung 15). Es werden vier *Treatments* unterschieden. Diese ergeben sich aus den Kombinationen von

Häufigkeiten Spielertypen in %

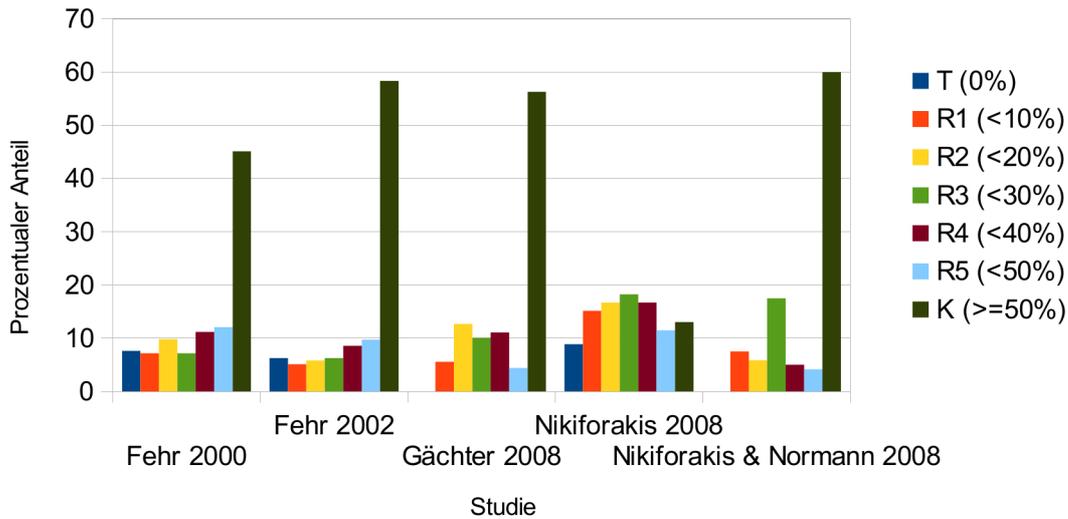


Abbildung 14. Häufigkeit von Spielertypen aggregiert über alle fünf Labor-PGG-Studien

Partner (Gruppenzusammensetzung bleibt gleich) und Stranger (Zusammensetzung wechselt) sowie mit Bestrafung (P) und ohne (NP). Es treten Treatment-Effekte auf: so führen sowohl Stranger-Treatments als auch straffreie Umgebungen zu einem vermehrten Auftreten egoistischer Strategien. Korrespondierend dazu werden kooperative Strategien nur dann eingesetzt, wenn sowohl die Gruppe vertraut ist als auch Bestrafungsmöglichkeiten gegeben sind.

Häufigkeiten Spielertypen in % nach Treatment

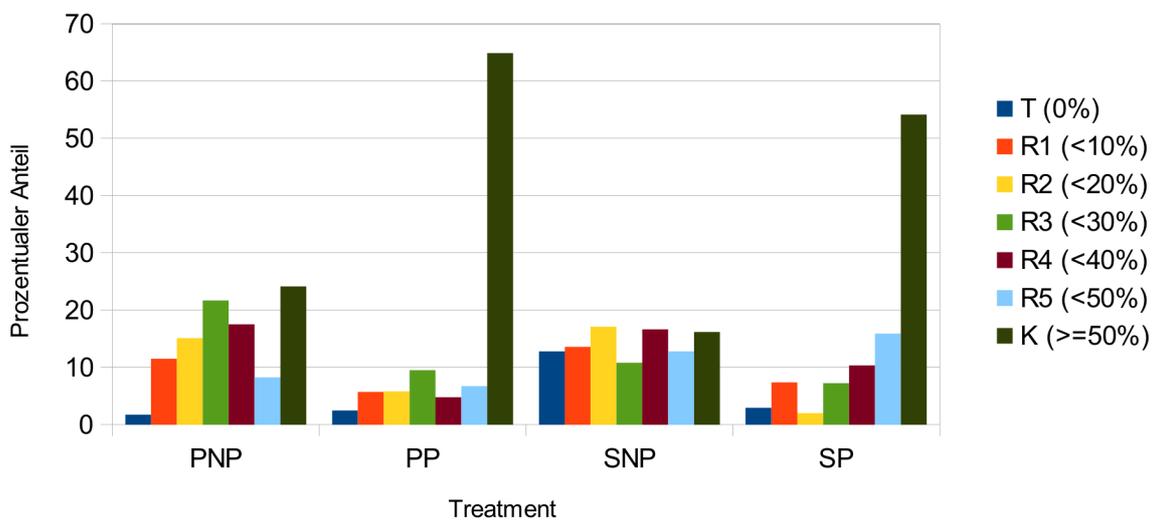


Abbildung 15. Häufigkeit von Spielertypen nach Treatment differenziert (PNP = Partner-No Punishment, PP = Partner-Punishment, SNP = Stranger-No Punishment, SP = Stranger-Punishment); Daten aus allen fünf Labor-PGG-Studien nach Treatment aggregiert

Zusammenfassend lässt sich sagen:

- die fünf Studien sind von ihren Parametern her gut vergleichbar,
- die Studie von Nikiforakis (2008) zeigt eine gleichmäßigere Verteilung von Spielertypen als die anderen Studien,
- für die letztgenannte Studie liegen keine Daten für den Verdienst *nach* Abzug der Bestrafungskosten vor; d.h. diese Studie wird bei allen Analysen bezüglich des Gesamtverdienstes ausgeschlossen.

Alle weiteren spezifischeren Analysen sind den jeweiligen Hypothesen zugeordnet.

Analog zu den Häufigkeitsverteilungen in den PGG werden diese in einem nächsten Schritt für Ikariam angegeben: wie verteilen sich die Spieler auf die insgesamt 5351 Inseln pro Welt in Hinblick auf die Anzahl an Städten pro Insel? Diese Zahlen haben einen wichtigen Einfluss auf die Erzeugung der Gemeingüter, die sich umso leichter erzeugen lassen, je mehr Personen daran beteiligt sind. Die folgende Tabelle 11 gibt dazu einen Überblick. Sie ist nach den möglichen Gruppenzusammensetzungen der Spielertypen aufgeteilt, die in Daten und Methoden (siehe Tabelle 4) definiert wurden.

Die heterogenste Gruppenzusammensetzung (KRT = Kooperierende, Reziprokatoren und Trittbrettfahrer) vereint länderübergreifend im Durchschnitt die meisten Städte pro Insel auf sich (12), gefolgt von den beiden Kombinationen KR mit 11 und RT ebenfalls mit 11. Reine Gruppen aus Kooperierenden K (1) bilden dagegen im Durchschnitt sehr viel kleinere Gruppen. T und KT-Gruppen existieren überhaupt nicht. KT-Gruppen setzen sich im Gegensatz zu den anderen beiden Gruppen aus zwei Extremstrategien zusammen (hohe Kooperation und reines Trittbrettfahren).

	K	KR	R	KRT	KT	RT	T
DE	0,51	10,46	8,73	11,62	0	10,49	0
EN	1,30	10,58	8,16	12,01	0	10,27	0
FR	0,67	11,16	9,39	11,98	0	10,71	0
GR	1,85	11,76	8,28	12,89	0	11,29	0
TR	0,54	10,00	8,64	11,78	0	10,41	0
Durchschnittliche Anzahl an Städten	0,97	10,79	8,64	12,06	0	10,63	0

Tabelle 11. Mittelwert Anzahl Städte pro Inseltyp im Ländervergleich ohne den ersten Snapshot (25.4) für Inseln mit mindestens zwei Spielern

Die folgende Abbildung 16 zeigt die Verteilungsmaße für die häufigste Gruppenzusammensetzung (KRT).

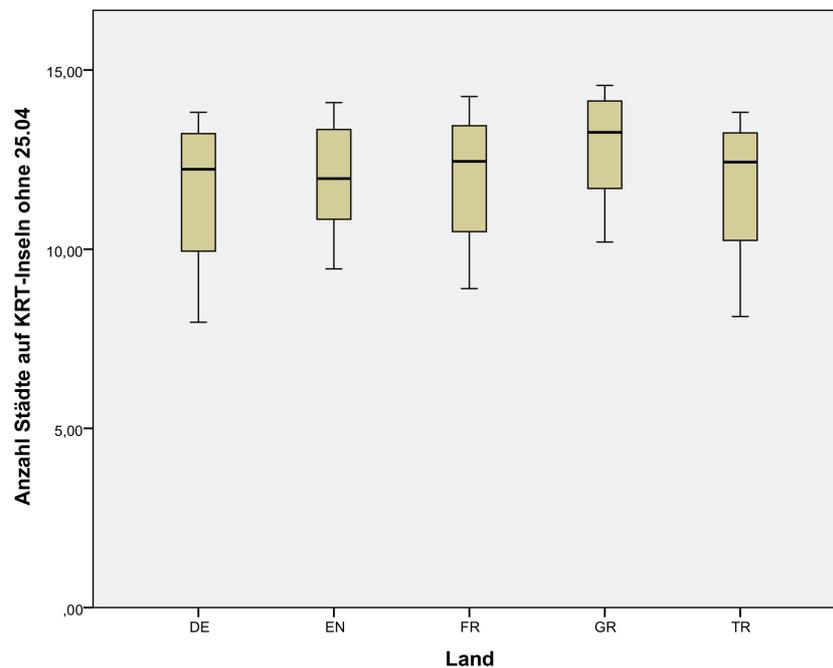


Abbildung 16: Verteilungsmaße für die Anzahl der Städte pro Land für KRT-Inseln

Wie verteilt sich nun die Häufigkeit der Inseltypen? Welche Gruppen sind besonders häufig anzutreffen? Die untenstehende Abbildung 17 zeigt deutlich, dass RT-Gruppen mit einem durchschnittlichen Anteil von 60% im Durchschnitt über alle Länder dominieren. Ihnen folgen KRT (21% Anteil) und R (14%).

Häufigkeit der Gruppentypen

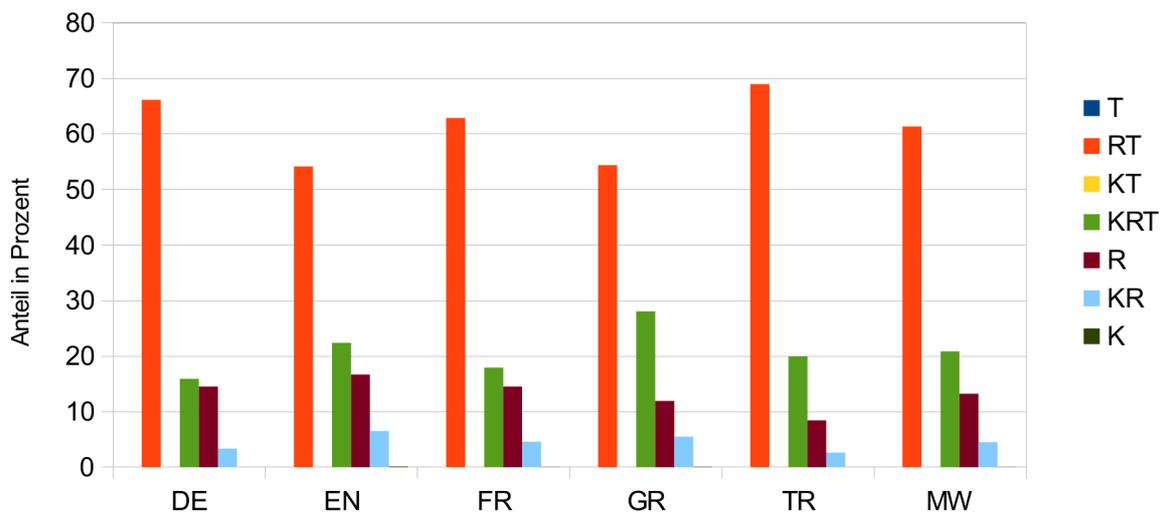


Abbildung 17. Häufigkeit der Gruppentypen für Inseln mit mindestens zwei Spielern für alle Snapshots

Die Häufigkeiten verändern sich deutlich (Tabelle 12), wenn auch Inseln zugelassen werden, die nur aus einem Spieler bestehen – hier kann man natürlich nicht mehr sinnvoll von Gruppen sprechen, weshalb sie in späteren Analysen auch stets ausgeschlossen bleiben.

Gruppentyp	Anteil für Spielerzahl > 1	Anteil für alle Spielerzahlen
K	0,1	1,3
KR	4,7	3,7
R	13,6	24,3
KRT	21,4	17,3
KT	0	0
RT	60,3	53,4
T	0	0

Tabelle 12. Vergleich des Anteils an Gruppentypen für Inseln mit und ohne Einzelspieler ohne den ersten Snapshot vom 25.4

Während KR und KRT bei Ausschluss von Einzelspielern fast unverändert bleiben, gibt es viele Inseln, auf denen nur ein Spieler des Typs K oder R ist. Diese Nischen entstehen etwa durch die aktive Suche von Einzelspielern (z. B. eines Spielertyps K). KT und T-Inseln existieren nicht.

Nach den Häufigkeiten und Gruppengrößen komme ich zu den Spielertypen. Man findet große

Unterschiede zwischen den einzelnen Strategien, d.h. Spielertypen können identifiziert und voneinander abgegrenzt werden. Die folgende Tabelle 13 zeigt diese Resultate im Detail für Reziprokatoren, Tabelle 14 für Trittbrettfahrer und Kooperierende. Die Spalten stehen dabei für die jeweilige Messmethode, die Tabelle ist also nur in Spalten zu lesen – Zeilen ergeben deshalb nicht 100%.

Land / Kriterium	Zunehmende oder abnehmende Richtung der Einzahlungen gleich der Gruppe	Positiver Spearman Korrelationskoeffizient & $p < 0.05$	Positiver Spearman Korrelationskoeffizient & $p < 0.01$	Positiver Spearman Korrelationskoeffizient & $p < 0.001$	max. -10% bis +10% Differenz zu Gruppeneinzahlungen
DE	58,84	40,96	30,23	20,16	4,38
EN	66,41	35,35	25,01	16,16	4,84
FR	56,58	38,94	28,07	18,33	4,74
GR	50,79	39,79	29,43	19,55	4,58
TR	67,20	39,41	28,45	18,32	3,94
Mittelwert	59,96	38,89	28,24	18,50	4,50

Tabelle 13. Häufigkeiten von Reziprokatoren in Prozent mit verschiedenen Messmethoden (siehe Methoden)

Trittbrettfahrer und Kooperierende sind wie folgt zusammengefasst:

Land / Definition	% Trittbrettfahrer (Mittelwert über alle 28 Snapshots)				% Kooperierende (Mittelwert über alle 28 Snapshots)			
	0 Einzahlungen	$\leq 3\%$ eigener Ressourcen im Mittel	$\leq 5\%$ eigener Ressourcen im Mittel	0 in mehr als 50% der Snapshots	$\geq 40\%$ eigener Ressourcen im Mittel	$\geq 50\%$ eigener Ressourcen im Mittel	$\geq 60\%$ eigener Ressourcen im Mittel	$\geq 50\%$ in mehr als 50% der Snapshots
DE	25,21	37,68	45,48	22,70	5,41	3,10	1,91	2,07
EN	21,56	33,27	40,42	19,11	9,41	6,18	4,28	4,53
FR	23,06	34,75	42,47	20,87	6,05	3,60	2,29	2,28
GR	24,69	38,53	46,61	19,67	9,09	6,69	5,07	4,86
TR	30,31	43,86	51,60	27,52	5,25	3,43	2,35	2,88
Mittelwert	24,97	37,62	45,32	21,97	7,04	4,60	3,16	3,32

Tabelle 14. Häufigkeiten von Trittbrettfahrern und Kooperierenden in Prozent mit verschiedenen Messmethoden (siehe Methoden)

Eine Übersicht über die Lagemaße der in den Studien 2 und 3 verwendeten Definitionen von Trittbrettfahrern (0% Investition in das Gemeingut) und Kooperierenden ($\geq 50\%$ Investition in das Gemeingut) geben die folgenden beiden Abbildungen 18 und 19. Zunächst für Trittbrettfahrer:

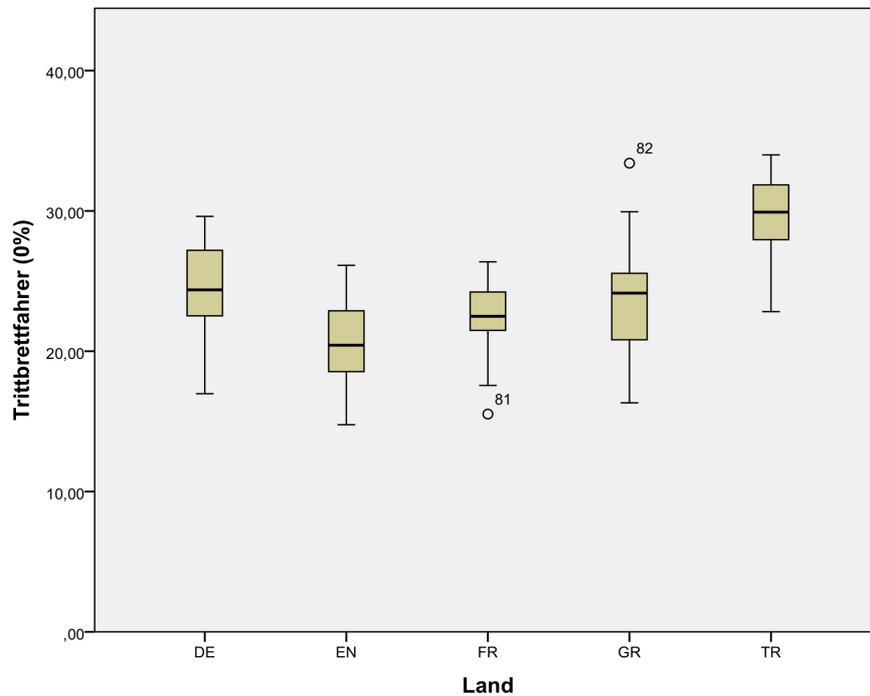


Abbildung 18: Lagemaße für Trittbrettfahrer (0%) im Ländervergleich

Und für Kooperierende:

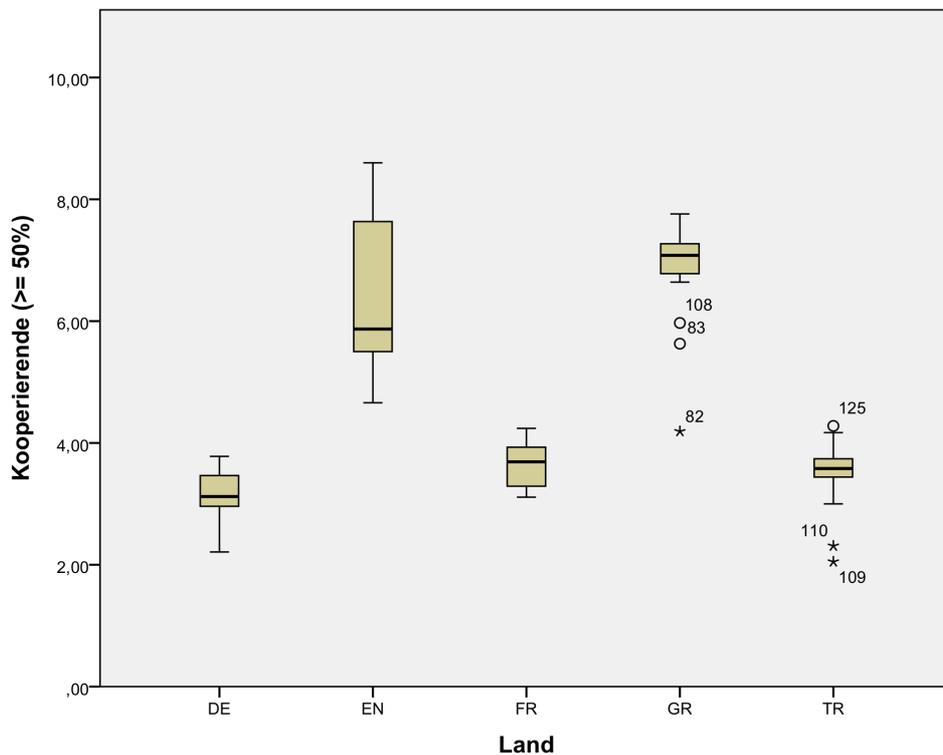


Abbildung 19: Lagemaße für Kooperierende ($\geq 50\%$) im Ländervergleich

Alle Länderunterschiede für Trittbrettfahrer und Kooperierende sind für alle Messmethoden, die in Tabelle 14 verwendet werden, für den letzten Snapshot (13.02) signifikant (Kruskal-Wallis-Test, $df = 4$, $n = 34451$, $p < 0.05$). Für die verwendeten Maße ($K \leq 50\%$ und $T = 0\%$) erweisen sich in den korrigierten Post-hoc-Tests alle paarweisen Vergleiche mindestens auf dem 0,001-Niveau als signifikant, bis auf EN-GR, DE-TR, DE-FR, TR-FR für Kooperierende und alle Ländervergleiche ohne die Türkei für Trittbrettfahrer.

Die Daten aus Tabelle 14 erlauben es sogar, zwischen pessimistischen Reziprokatoren und Trittbrettfahrern zu unterscheiden. Trittbrettfahrer sind Spieler, die nichts geben und der Meinung sind, dass dies andere auch nicht tun werden. Pessimistische Reziprokatoren sind hingegen Spieler, die deshalb nichts in das öffentliche Gut beisteuern, weil sie der Meinung sind, andere Spieler geben nichts – sie reagieren demnach auf die Einzahlungen der anderen, während Trittbrettfahrer davon unbeeindruckt bleiben. Diese fehlende Unterscheidung wurde verschiedentlich kritisch angemerkt, da sie mit der Methode der *eliciting beliefs*, die der Standard ist, nicht möglich ist (Kocher et al. 2008).

In dieser Arbeit können diese Typen allerdings unterschieden werden, indem nur Fälle ausgewählt werden, in denen die Einzahlung der Gruppe > 0 ist und die Einzahlung des untersuchten Spielers 0. Fälle, in denen sowohl die Gruppe als auch der Spieler nichts in das

Gemeingut investieren, werden nicht berücksichtigt, da sie sowohl durch pessimistische Reziprokatoren oder Trittbrettfahrer zustande kommen können. Diese Maßnahme hat zur Folge, dass einerseits alle pessimistischen Kooperierenden ausgeschlossen werden können. Dadurch wird der Anteil der Trittbrettfahrer und der Reziprokatoren präziser eingegrenzt. Andererseits wird der Anteil der Trittbrettfahrer sehr leicht unterschätzt (ca. 1-2%).

Ein weiteres wichtiges Ergebnis ist die Häufigkeit der Spielertypen im Spielverlauf. Denn es ist anzunehmen, dass unerfolgreiche Strategien angepasst oder sogar aufgegeben werden, es also zu Lerneffekten kommt. Die folgenden beiden Abbildungen (Abbildungen 20 und 21) zeigen den Anteil von Kooperierenden bzw. von Trittbrettfahrern über einen Verlauf von zehn Monaten.

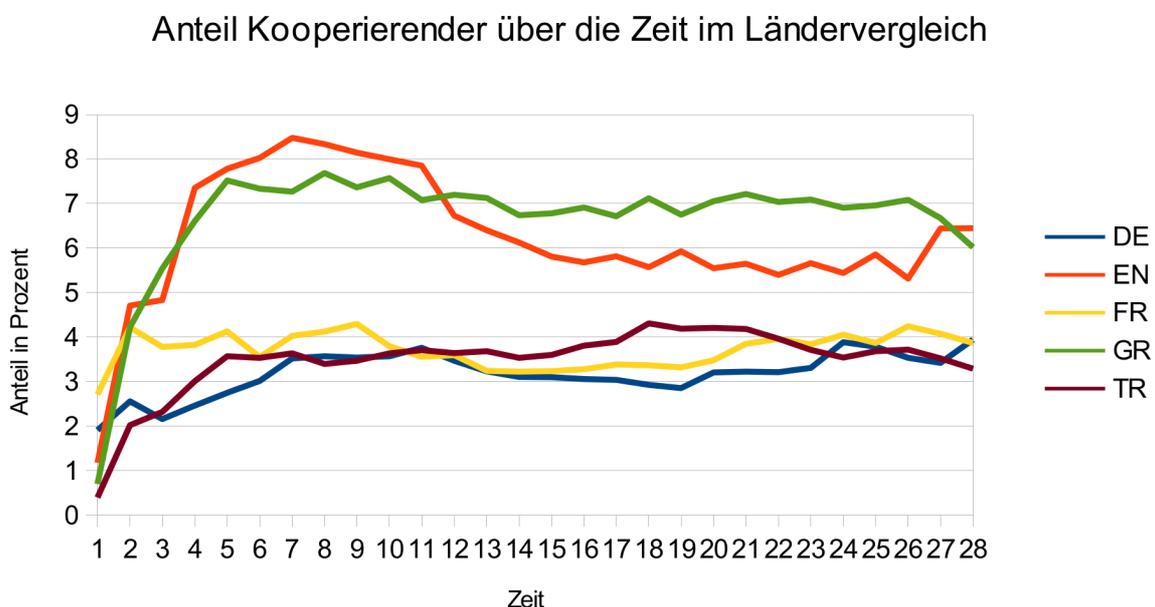


Abbildung 20. Anteil Kooperierender über die Zeit im Ländervergleich

Die Anteile der Spielertypen ändern sich aus mehreren Gründen über die Zeit. Zum einen kommen neue Spieler hinzu und bestehende Spieler hören auf. Zum anderen gründen Spieler im Spielverlauf ständig neue Städte. Da diese Arbeit allerdings von stabilen Spielertypen ausgeht, sollten diese Schwankungen minimal sein. Deshalb wird getestet, ob die Spielertypen zwischen den Zeitpunkten schwanken. Um Unterschiede zwischen den Zeitpunkten statistisch belastbar nachzuweisen, wird der Friedman-Rang-Test verwendet, und die Zugehörigkeit zu einem Spielertyp über die 28 Snapshots getestet. Für alle Länder fällt dieser Test zwar signifikant aus (Friedman-Rang-Test, $df = 27$, DE: $n = 92$, EN: $n = 92$, FR: $n = 135$, GR: $n = 137$, TR: $n = 167$, $p < 0,001$), bei den paarweisen Vergleichen fällt jedoch auf, dass die Unter-

schiede *ausschließlich* auf den Unterschied zum ersten Snapshot zurückzuführen sind. Alle anderen korrigierten Post-hoc-Vergleiche sind nicht signifikant. Da ich in den Methoden (C 2) bereits die Andersartigkeit der Anfangsphase belege, können diese nicht signifikanten Befunde als Beleg für die Stabilität der Spielertypen über die Zeit gewertet werden.

Es ist in Abbildung 20 gut zu sehen, dass extrem wenige Spieler zu Beginn sehr kooperativ auftreten – das tun länderübergreifend weniger als drei Prozent (in GR und TR sogar weniger als 1%) aller Teilnehmer. Dieser Anteil steigt jedoch rapide an und pendelt sich zwischen 3% und 6 % ein. Damit wird sichtbar, dass selbst nach über zehn Monaten kooperative Strategien als Randgruppe existieren. Solche Investitionen von mehr als 50 % der eigenen Ressourcen in die Gemeingüter über einen langen Zeitraum stellen eine extreme Strategie dar. Wie sieht es nun mit dem anderen Extrem, den Trittbrettfahrern aus? Die unten stehende Abbildung (Abbildung 21) zeigt dies analog. Hier sieht man einen gegenläufigen Trend. Ausgehend von Spitzenwerten von über 50% Anteil von Trittbrettfahrern an der Gesamtpopulation zu Beginn, fällt dies auf einen Wert zwischen 20% und 30 % ab, um sich am Ende in einem engen Korridor von 16% bis 23% für alle Länder wiederzufinden.

Anteil Trittbrettfahrer über die Zeit im Ländervergleich

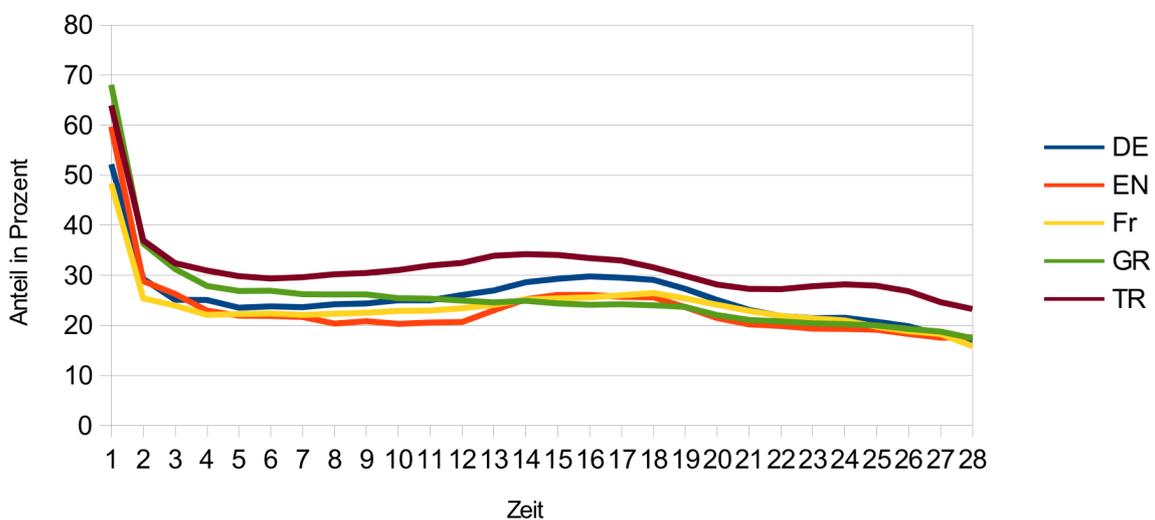


Abbildung 21: Anteil Trittbrettfahrer über die Zeit im Ländervergleich

Für die Spielertypen R2 bis R5 gelten ähnlich stabile Trends, als Beispiel R2:

Anteil Reziprokatoren R2 über die Zeit im Ländervergleich

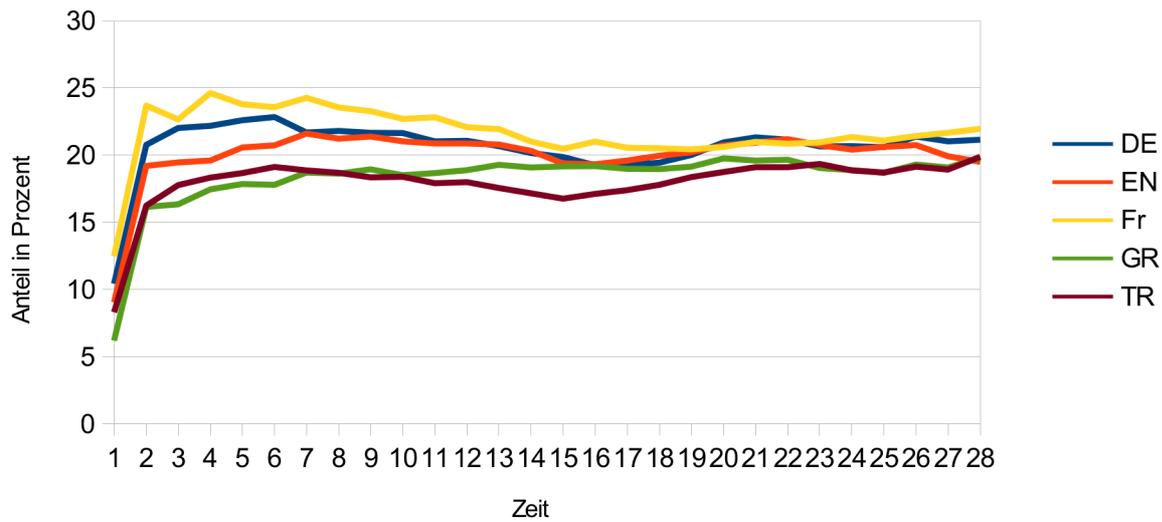


Abbildung 22: Anteil Reziprokatoren R2 über die Zeit im Ländervergleich

Eine Ausnahme macht R1 – hier ist ein Aufwärtstrend erkennbar (Abbildung 23):

Anteil Reziprokatoren R1 über die Zeit im Ländervergleich

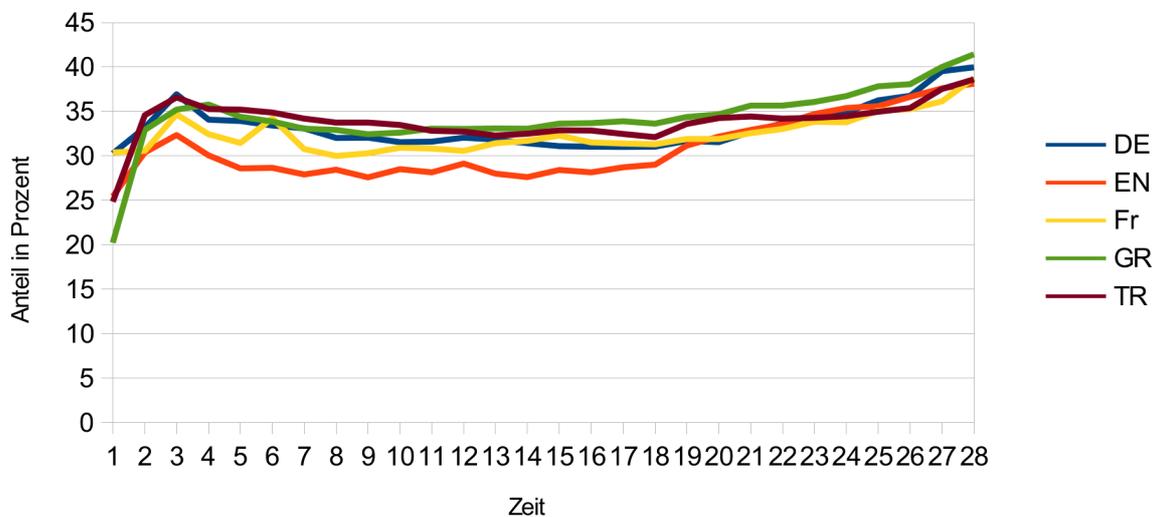


Abbildung 23: Anteil Reziprokatoren R1 über die Zeit im Ländervergleich

Damit sind R1-Spieler mit einem mittleren Anteil von 33% die zahlenmäßig häufigste Gruppe, noch vor den Trittbrettfahrern mit 26%, wenn für beide Spielertypen über alle Länder aggregiert wird.

Aus diesen recht stabilen Verläufen ergibt sich die Frage, wie stabil bzw. wie wiederholbar das Verhalten einzelner Spieler ist. Dazu wird zunächst die Standardabweichung der abhän-

gigen Variablen (Investition in das Gemeingut in Prozent) berechnet, da sie Auskunft darüber gibt, wie stark das Verhalten schwankt (ohne den ersten Snapshot): sie liegt für Deutschland bei 5,42%, für England bei 5,95%, für Frankreich bei 5,50%; für Griechenland bei 5,93% und für die Türkei bei 4,94%. Diese Werte stellen die Mittelwerte pro Land aller individuellen Standardabweichungen dar.

Eine Gesamtübersicht über alle Spielertypen nach meiner Definition gibt Tabelle 15:

Spielertyp/ Land	DE	EN	FR	GR	TR	Mittelwert
T	25,6	23,4	23,7	25,8	31,5	26,0
R1	33,3	30,8	32,5	34,3	34,0	33,0
R2	20,6	20,0	21,7	18,3	18,0	19,7
R3	10,0	10,4	10,8	8,3	7,7	9,4
R4	4,9	5,6	5,1	4,2	3,6	4,7
R5	2,4	3,4	2,6	2,5	1,9	2,6
K	3,2	6,3	3,7	6,7	3,5	4,7

Tabelle 15. Mittelwerte über alle 28 Snapshots der Anteile von Spielertypen in Prozent im Ländervergleich

Zum Vergleich und zur Abschätzung der Robustheit der Ergebnisse zeigt die folgenden Tabelle die Häufigkeitsverteilung der Spielertypen nach zehn Monaten (am 13.2):

Spielertyp/ Land	DE	EN	FR	GR	TR	Mittelwert
T	16,9	17,6	15,9	17,4	23,3	18,9
R1	40,0	38,1	38,8	41,4	38,6	39,4
R2	21,2	19,5	22,0	19,8	19,9	20,4
R3	9,9	9,8	11,2	8,5	8,8	9,5
R4	5,3	5,0	5,8	4,3	4,1	4,8
R5	2,9	3,5	2,6	2,6	2,1	2,6
K	3,9	6,4	3,9	6,0	3,3	4,5

Tabelle 16: Häufigkeitsanteile der Spielertypen in Prozent im Ländervergleich nach zehn Monaten

Die Ergebnisse zwischen den Mittelwerten über die gesamte Zeit und den letzten Snapshot sind vergleichbar, da die Spielertypen R2, R3, R4, R5 und K um maximal 0,7 abweichen. Eine Ausnahme besteht in der Abnahme der Trittbrettfahrer um 7% – und um diesen Prozentsatz nimmt die angrenzende Kategorie R1 zu.

1.2 Diskussion Hypothesen 1 und 2

1.2.1 Interpretation der Häufigkeiten von Spielertypen

Für Trittbrettfahrer sind die auftretenden Spannweiten recht niedrig und die Häufigkeiten über die Länder vergleichbar. Für Kooperierende (bis auf Kooperierende in England) sind die Spannweiten sogar noch kleiner, auch wenn es hier zu Unterschieden in den Häufigkeiten im Ländervergleich kommt. Diese Stabilität zeigt sich auch über die Zeit – bis auf Spielertyp R1. Dessen Häufigkeitszunahme kommt durch Zugewinne aus dem benachbarten Spielertyp Trittbrettfahrer zustande, die dann doch zumindest etwas (< 10%) beisteuern.

Es überrascht, wie häufig Versuchspersonen mehr als 50% ihres Einsatzes in das Gemeingut über den gesamten Spielverlauf einzahlen (was hier die Definition für Spielertyp K ist) – das sind zwischen 45% und 60% je nach Studie mit dem „Ausreißer“ von 13% für Nikiforakis (2008).

Weiterhin fällt auf, dass in zwei von fünf Studien (Gächter et al. 2008 und Nikiforakis & Normann 2008) der Spielertyp Trittbrettfahrer überhaupt nicht vorkommt. Dies hängt damit zusammen, dass beide Studien nur mit Partner-Treatments arbeiten und es eine große Anzahl Kooperierender gibt (Abbildung 14 und 15).

Während Kocher et al. (2008) große Länderunterschiede für die verschiedenen Spielertypen berichten – und zwar zwischen den USA, Österreich und Japan mit 8%, 22% und 36% für Trittbrettfahrer – verorten andere Studien den möglichen Anteil an Trittbrettfahrern zwischen 10% (Rustagi et al. 2010) und etwa 30% (Fischbacher et al. 2001) sowie 32% (Burlando & Guala 2005). Meine Resultate können letztere Ergebnisse stützen, da in allen fünf Ländern sehr stabile Werte in einem engen Korridor gefunden werden: 22% (DE), 23% (EN), 24% (FR), 25% (GR) und 30% (TR). Der Gesamtmittelwert liegt bei 25 %.

Kooperierende sind viel seltener als Trittbrettfahrer. Ihre prozentualen Häufigkeiten liegen zwischen 2% (Rustagi et al. 2010) und 13% (Kurzban & Houser 2005). Eine andere Studie, die bis zu 8% findet, verwendet allerdings ein sehr laxes Kriterium (Burlando & Guala 2005). Die Häufigkeiten im vorliegenden Datensatz bestätigen diese Ergebnisse, setzen jedoch die Häufigkeit relativ niedrig bei etwa 5% an (im Durchschnitt aller Länder).

Im Gegensatz zu den relativ stabilen und übereinstimmenden Zahlen ist sich die Literatur und auch die vorliegende Studie uneinig (siehe die Differenzen zwischen Tabelle 13 und 15), mit welcher Häufigkeit konditional Kooperierende auftreten. Allein durch die verschiedenen

Messmethoden variieren die Häufigkeiten bereits enorm (siehe Tabelle 13). Um dieses Problem zu umgehen, messe ich die Anteile dieser Spielertypen im Vergleich jeweils mit zu bisherigen Studien identischen Messmethoden. Nichtsdestotrotz findet diese Arbeit große Unterschiede: Im Vergleich zu Fischbacher et al. 2001, der von 50% Reziprokatoren spricht, finde ich nur 28%; im Vergleich zur Methode von Rustagi et al. 2010, der 34% sowie 12% (für schwach konditional Kooperierende), berechne ich 19% und 20%; im Vergleich zu der Metrik von Kocher et al. 2008, der 42 bis 81% berichtet, komme ich auf 51 bis 67%, obwohl diese Arbeit einen schwachen monotonen Anstieg nur emulieren kann, indem sie testet, ob die Versuchspersonen ihre Einzahlungen in das Gemeingut erhöht oder verringert, je nachdem was die Gruppe gerade tut.

Die vergleichsweise niedrigen Häufigkeiten (3,94 bis 4,84%) für Spieler, die im Bereich von -10 bis + 10% der Investitionen anderer Spieler in das öffentliche Gut einzahlen ergeben sich vor allem aus den zeitlichen Differenzen zwischen Spielern – je länger ein Spieler schon dabei ist, desto mehr kann er absolut einzahlen. Zu Anfängern sind diese Unterschiede sehr groß, weswegen ich den prozentualen Anteil der Einzahlungen der eigenen Ressourcen als Erfolgsmaß wähle. Dieser Unterschied ist also auf Unterschiede zwischen Labor-PGG und dem Kontext von Ikariam zurückzuführen.

Bislang ist fraglich, ob die Häufigkeiten der Spielertypen auch außerhalb der PGG-Laborsituation Gültigkeit besitzen. Wie eben diskutiert, ist die bisher gefundene Bandbreite zudem hoch: Man findet 34-81% für konditionale Kooperation, 8-36% für Trittbrettfahrer und 7-18% für Kooperierende. Es ist ungeklärt, ob diese Varianz unterschiedlichen Messmethoden, geringen Teilnehmerzahlen oder fehlenden Vergleichen geschuldet ist, um nur die wichtigsten Möglichkeiten zu nennen. Die vorliegende Studie ist von keinem dieser Defizite betroffen und findet zwar statistisch signifikante Unterschiede zwischen den Ländern, die Unterschiede in Prozent sind jedoch vergleichsweise klein.

Für *Trittbrettfahrer* finde ich zwischen 22% (EN) und 30% (TR). Der Mittelwert aller Länder beträgt hier 25%. Im Vergleich mit dem Mittelwert aus den in Tabelle 10 genannten Studien, der 20% beträgt, kann dieser Wert als bestätigt gelten.

Für *konditional Kooperierende* finde ich allerdings einen geringeren Prozentsatz für die am häufigsten verwendete Definition: in Ikariam liegt der niedrigste Wert bei 25% (EN) und der höchste bei 30% (DE), der Mittelwert bei 28%. Die Differenz zu den publizierten Studien (Mittelwert aus Tabelle 10: 52%) ist hier größer und liegt bei 24%.

Auch den Anteil an *Kooperierenden* berechne ich mit 3% (DE) bis 7% (GR) und einem Mittelwert von 5% im Vergleich geringer (Mittelwert aus Tabelle 10: 15%).

Allgemein gesprochen könnten diese Unterschiede in den Ergebnissen eine Ursache der künstlichen Situation im Kontext der Laborexperimente sein, sobald man versucht, Präferenzen abzugreifen (*eliciting beliefs*). Da bei dieser Methode für alle mögliche Einzahlungen der anderen Teilnehmer die eigene Einzahlung hypothetisch abgefragt wird, könnte es sein, dass viele Teilnehmer eine einfache Heuristik anwenden, indem sie bei steigenden Einzahlungen der anderen einfach ihre Einzahlungen synchron steigern – das Ergebnis wäre ein großer Prozentsatz an vermeintlichen Reziprokatoren; und genau das findet man. Ein zweiter Grund könnte in der Spielstruktur von *Ikariam* liegen: zwischen den Spielern gibt es große Unterschiede in der Spielzeit und demzufolge in den Investitionen.

Schließlich ist es wichtig, zu erwähnen, dass die Ergebnisse konsistent über die verschiedenen Messmethoden sind – zumindest für Trittbrettfahrer und Kooperierende. So hat zum Beispiel England für alle vier Messmethoden von Trittbrettfahrern immer den geringsten prozentualen Anteil und die Türkei immer den jeweils höchsten.

Diese Arbeit kann demnach die bisherigen Ergebnisse über kooperatives Verhalten mit einigen wichtigen Qualifikationen bestätigen:

- es gibt individuell verschiedene und gut unterscheidbare Kooperationsstrategien, was
- in Spielertypen resultiert, die über die in PGG üblichen Typen hinausgehen;
- diese Spielertypen treten kulturübergreifend auf,
- ihre relativen Anteile in der Gesamtbevölkerung sind relativ stabil, was vor allem für Trittbrettfahrer, die bei etwa 25% und Kooperierende, die bei 5 % liegen, gilt;
- für Reziprokatoren bzw. konditional Kooperierende finde ich eine weitaus größere Spannweite an Häufigkeiten.

Indem diese Arbeit die verschiedenen Spielertypen mit den häufigsten in der Literatur verwendeten Messmethoden misst, mache ich die Ergebnisse bisheriger Studien vergleichbar mit den vorliegenden Resultaten und erhöhe die Robustheit der Ergebnisse. Ich identifiziere in einer großen und sehr heterogenen Stichprobe, die sich über fünf Länder erstreckt, klar trennbare Spielertypen, nämlich Trittbrettfahrer, Reziprokatoren und Kooperierende. Das ist ein erster Schritt zu einer detaillierteren Analyse von Kooperation, die auch Persönlichkeitsmerk-

male im Sinne einer Persönlichkeitspsychologie als Erklärung berücksichtigt (Bierhoff & Fetchenhauer 2006).

Möglicherweise trägt dieses Ergebnis auch zur weiteren Klärung der Frage bezüglich der Fragilität von Kooperation bei. Da konditional Kooperierende ihre Beiträge zu Gemeingütern an die Investitionen der anderen anpassen, was bei vielen Trittbrettfahrern zu einer Abwärtsspirale führt, lässt sich spekulieren, ob Kooperierende und Reziprokatoren deshalb existieren, weil sie es in einigen Fällen schaffen, durch den Ausschluss von Trittbrettfahrern diese im Erfolg zu übertreffen (vgl. Burlando & Guala 2005; Page et al. 2005). Diese Frage wird auch in Teil D 3 aufgenommen.

1.2.2 Evolutionäre Interpretation der Spielertypen

Aus einer biologischen Perspektive ist das Ergebnis verschiedener Strategien keine Überraschung. Im Gegenteil, es gibt durch alle Taxa hindurch beliebig viele Beispiele dafür (Volland 2013). Das Auftreten von Trittbrettfahrern ist ebenfalls zu erwarten, da es sich um eine evolutionär stabile Strategie handelt. Eine Erklärung für konditionale Kooperation, die mit etwa 50% Einzahlung beginnt, ist dagegen komplexer. Axelrod (1984/2000) hat erfolgreiche Kooperationsstrategien über fünf Merkmale charakterisiert: ist sie freundlich (beginnt sie mit Kooperation?), ist sie zurückschlagend (vergilt sie Defektion mit Defektion?), ist sie vergebend (kann sie nach Defektion wieder zu Kooperation zurückkehren?), ist sie nicht-neidisch (ist sie auch mit einem zweiten Platz zufrieden?) und ist sie einfach (ist sie berechenbar?)? Diese fünf Merkmale sind an vielen erfolgreichen Strategien entdeckt worden. Konditionale Kooperation besitzt alle diese Merkmale einer solch potentiell erfolgreichen Strategiewahl. Da die Ähnlichkeit mit der verbreiteten Tit-for-Tat-Strategie ebenfalls gegeben ist, könnten diese Charakteristika für Häufigkeit und Erfolg verantwortlich sein.

Ein grundlegendes Merkmal dieser Strategie ist die Fähigkeit zur Diskriminierung zwischen Trittbrettfahrern und Kooperierenden, über die wir als Menschen verfügen (Barclay 2008; Oda 1997; Bell et al. 2010). Da unser ökologisches Habitat vor allem durch soziale Probleme gekennzeichnet war (Dunbar 1998), ist dies ebenfalls kongruent zu den Entstehungsbedingungen menschlicher Kooperation. Es passt auch zu Befunden, die nachweisen, dass Kooperation vor allem von reziproken Interaktionen bestimmt ist, und nicht nur von Verwandtschaftsverhältnissen (Hill 2001).

Der kleine Anteil (ca. 5%) an un konditionalen altruistischen Strategien lässt sich möglicher-

weise durch die Anfälligkeit gegenüber defektierenden Strategien erklären. Sollte diese Überlegung stimmen, würde folgen, dass wir bestimmte, prototypische Reaktionen erwarten dürfen, wenn eine Strategie auf eine andere trifft (Trittbrettfahrer auf Kooperierende, konditional Kooperierende auf Trittbrettfahrer, usw.). Die Existenz solcher Strategien ist möglicherweise auch darauf zurückzuführen, dass einige oder viele Dilemmasituationen in unserer evolutionären Vergangenheit nicht dem Gefangenendilemma entsprachen, sondern eher dem Snowdrift-Dilemma oder dem Hirschjagd-Dilemma (*Stag-Hunt*) ähnelten. Für diese Vermutung liegen allerdings keine Daten vor.

Zwei zentrale Ergebnisse dieser Analyse – der hohe Anteil an Strategien, die je nach Interaktionspartner mit Kooperation oder Defektion reagieren und der kleine Anteil an Individuen, die hoch kooperieren, führen zur naheliegenden Frage: Wie können kooperationsbereite Individuen Trittbrettfahrer abwehren? Die vielleicht wichtigste Maßnahme ist die Bestrafung derjenigen, die nicht oder nicht genug in ein Gemeingut einzahlen. Obwohl darüber viel Literatur existiert, wird dabei üblicherweise nicht zwischen den Spielertypen differenziert. Die folgenden Analysen greifen die vorangegangenen Ergebnisse auf und untersuchen das Strafverhalten je nach Spielertyp.

2 Ergebnisse und Diskussion Studie 2: Effiziente Bestrafung

Anmerkung: Da in dieser Studie fünf Arbeitshypothesen bearbeitet werden, ist aus Lesbarkeitsgründen die Diskussion direkt an die jeweilige Hypothese angeschlossen und nicht wie bei Studie 1 im Block danach.

2.1 Resultate für Hypothese 1: Kooperative Spielertypen strafen häufiger

Hypothese 1a: Kooperativere Spielertypen strafen häufiger und mehr als egoistische.

Um diese Hypothese zu testen, werden die durchschnittlichen Strafpunkte pro Spielertyp (Abbildung 24), die gesamten Strafpunkte (Abbildung 25) und die Anzahl der Strafen (Abbildung 26) ermittelt. Spielertyp T (Trittbrettfahrer) ist zwar der Vollständigkeit halber aufgeführt, tritt aber nur in der Studie von Fehr & Gächter (2002) auf (jedoch nicht als Bestrafter, sondern nur als bestrafter Spielertyp).

Es zeigt sich kein übergreifendes Muster über alle drei Studien (Abbildung 24). Während in Fehr & Gächter (2000) und Nikiforakis (2008) tendenziell kooperativere Spielertypen mehr bestrafen, ist dies in Fehr & Gächter (2002) nicht der Fall – hier strafen Typ R1 und R2 am meisten.

Strafen pro Spielertyp

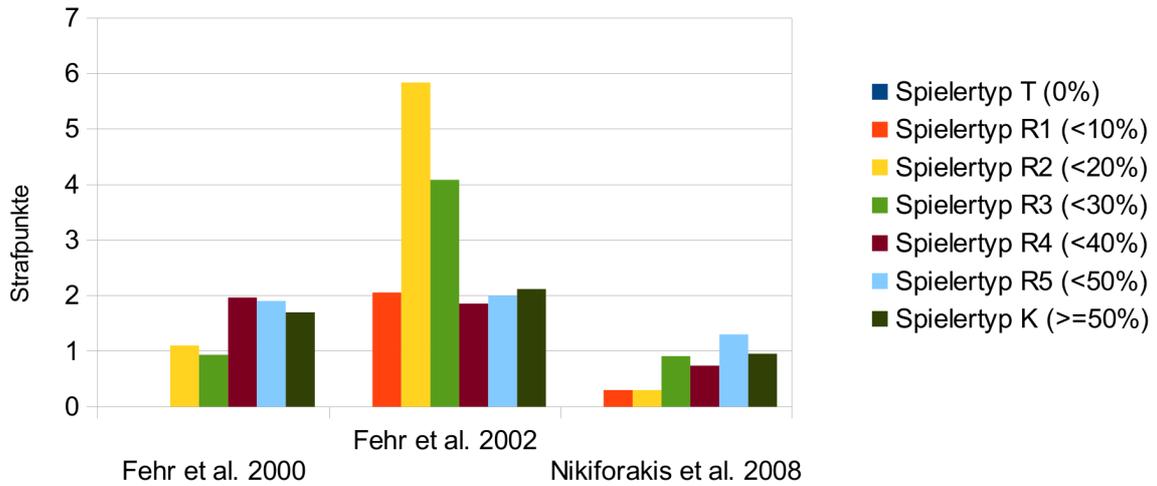


Abbildung 24: Summe aller Strafpunkte pro Spielertyp (durch Anzahl Spieler pro Spielertyp)

Auch die gesamt gegebenen Strafpunkte unterstützen diese Aussage, da die Werte stark streuen, und keine eindeutige Tendenz erkennen lassen:

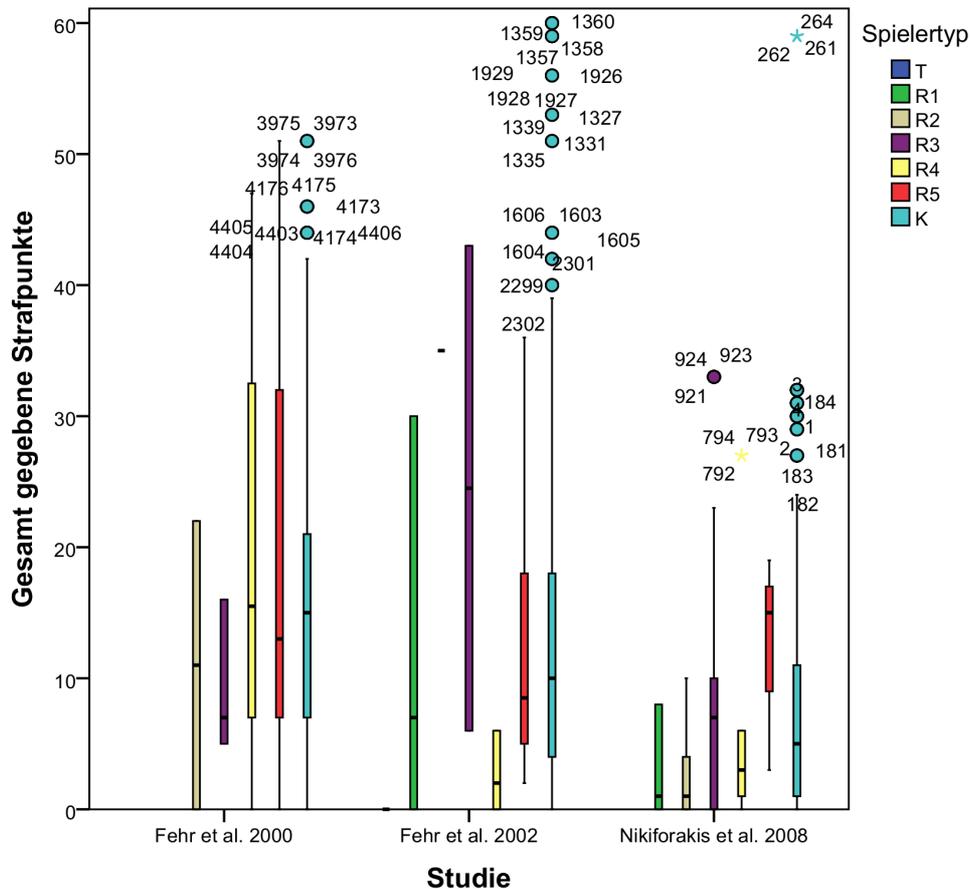


Abbildung 25: Gesamt gegebene Strafpunkte pro Spielertyp und Studie

Ein Test auf Unterschiede zwischen den Spielertypen in Bezug auf die gegebenen Strafpunkte ergibt ein signifikantes Ergebnis (Kruskal-Wallis-Test, $df = 5$, $n = 3370$, $p < 0,001$). Am meisten bestraft der Spielertyp R5, gefolgt von Kooperierenden, danach folgen R4, R3, R2 und R1. In den korrigierten Post-hoc-Tests erweisen sich alle paarweisen Vergleiche mindestens auf dem 0,04-Niveau als signifikant, bis auf R1-R2, R1-R3, R2-R3, R3-R4 und R4-K.

Neben der durchschnittlichen Summe an Strafpunkten ist auch die *Häufigkeit* von Interesse, also die Anzahl der Runden und Spieler, an die Strafen verteilt werden. Die nächste Abbildung (26) zeigt diese Häufigkeit.

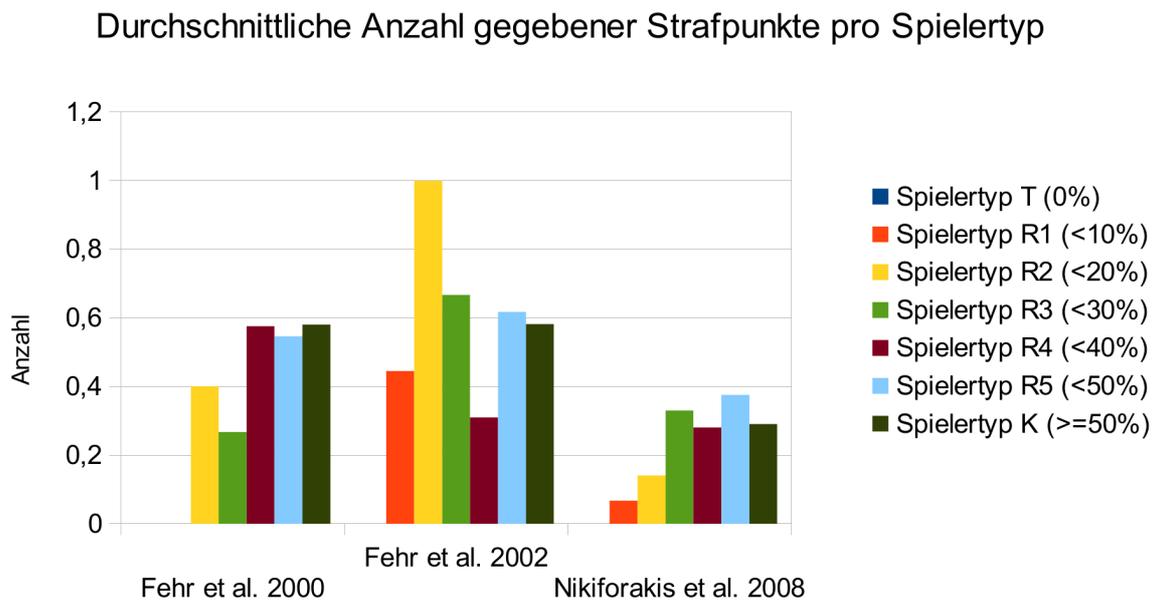


Abbildung 26: Anzahl aller Strafen pro Spielertyp (durch Anzahl Spieler pro Spielertyp)

Diese Daten zeigen für Fehr & Gächter (2000) und Nikiforakis (2008) die Tendenz kooperativerer Spieler auch mehr zu strafen, während in Fehr & Gächter (2002) die Spielertypen R2 sowie R3 am häufigsten bestrafen.

Der zweite Teil der Hypothese zielt auf den Erhalt von Strafen ab:

Hypothese 1b: Kooperativere Spielertypen werden seltener und weniger als egoistische bestraft.

Auch für diese Hypothese gebe ich Anzahl und Häufigkeit an, diesmal der *erhaltenen* Strafpunkte pro Spielertyp (Abbildung 27), der gesamt erhaltenen Strafen (Abbildung 28) sowie die Häufigkeit der Bestrafung (Abbildung 29).

Summe erhaltener Strafpunkte

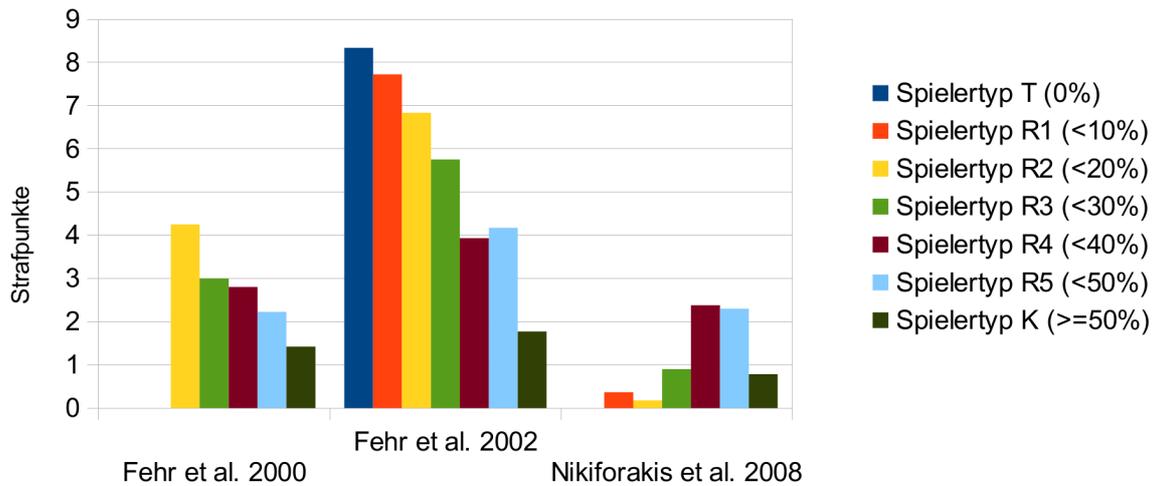


Abbildung 27: Summe aller erhaltenen Strafpunkte pro Spielertyp (durch Anzahl Spieler pro Spielertyp)

Während sich in den Studien mit Stranger-Treatments (Fehr & Gächter 2000 und Fehr & Gächter 2002) ein klarer negativer Zusammenhang zwischen Investition und Bestrafung abzeichnet, ist dies im Partner-Treatment bei den Daten von Nikiforakis (2008) nicht der Fall. Man findet dabei durchschnittlich über 2 Strafpunkte pro Runde und Spielertyp für Individuen, die mehr als 30% bzw. 40% in das Gemeingut einzahlen. Im Boxplot wird auch die Breite der erhaltenen Strafpunkte pro Spielertyp sichtbar:

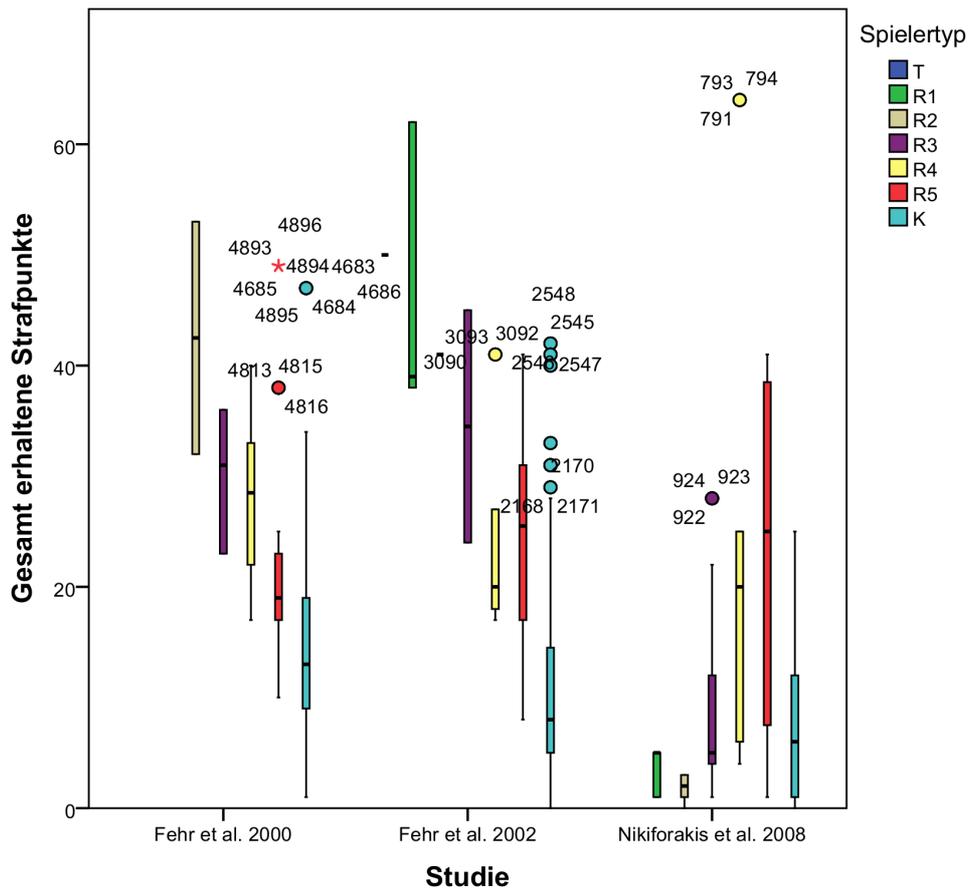


Abbildung 28: Gesamt erhaltene Strafpunkte pro Spielertyp und Studie

Hier folgen die beiden Studien im Stranger-Design (Fehr & Gächter 2000 und Fehr & Gächter 2002) wiederum einer klaren Abstufung der Bestrafung je nach Höhe der Einzahlungen – je mehr man in das öffentliche Gut investiert, desto weniger wird man auch bestraft. Für die Studie von Nikiforakis (2008) gilt das nicht – R4 und R5 ziehen gesamt die meisten Strafen auf sich.

Ein Test auf Unterschiede zwischen den Spielertypen zeigt ein signifikantes Ergebnis (Kruskal-Wallis-Test, $df = 6$, $n = 3376$, $p < 0,001$). Am meisten werden die seltenen Trittbrettfahrer in Fehr et al. (2002) bestraft, gefolgt von R4 und R5. Niedrigere mittlere Ränge als diese beiden haben R3, R1 und R2 sowie schließlich Kooperierende (K). In den korrigierten Post-hoc-Tests erweisen sich alle paarweisen Vergleiche mindestens auf dem 0,03-Niveau als signifikant, bis auf die in Tabelle 17 gekennzeichneten (siehe auch Diskussion):

	T	R1	R2	R3	R4	R5	K
T					n.s.		
R1				n.s.			
R2		n.s.					
R3							
R4						n.s.	
R5	n.s.						
K		n.s.	n.s.	n.s.			

Tabelle 17: Paarweise Vergleiche der Post-hoc-Tests für Unterschiede in der Bestrafung zwischen Spielertypen in PGG

Die Häufigkeiten der Bestrafung (Abbildung 29) bestätigen die obigen Beobachtung, da auch hier in Fehr & Gächter (2000) und Fehr & Gächter (2002) pro Spielertyp ein Spieler weniger oft bestraft wird, je mehr er oder sie in das öffentliche Gut einbezahlt hat:

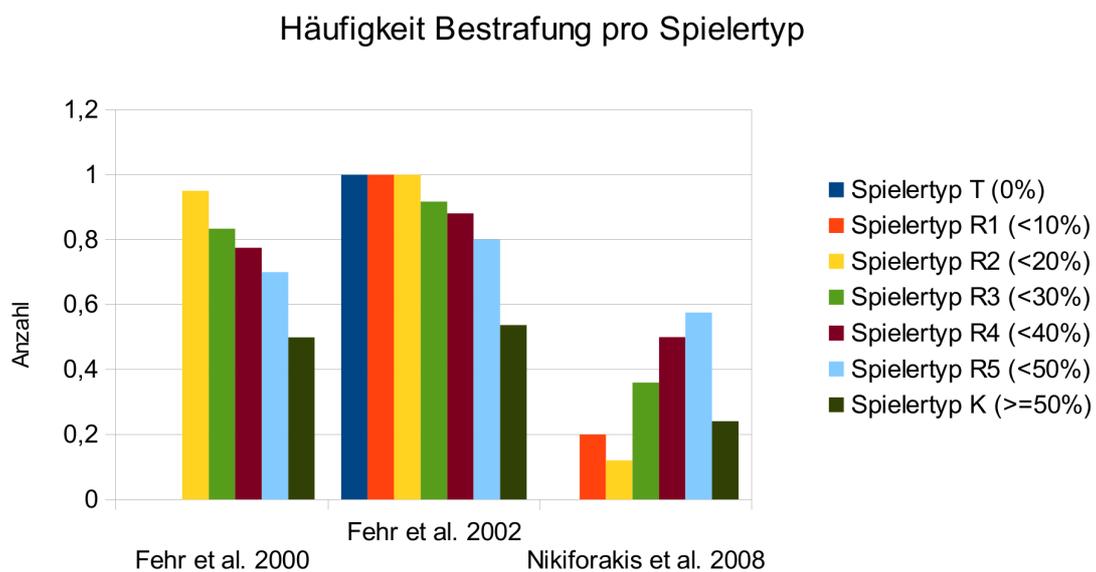


Abbildung 29: Häufigkeit der Bestrafung pro Spielertyp (durch Anzahl Spieler pro Spielertyp)

In der Studie im Partner-Design (Nikiforakis 2008) werden dagegen vor allem die Spielertypen R3, R4 und R5 bestraft, sowohl was Höhe, als auch Häufigkeit betrifft.

Dies bestätigen auch Korrelationsanalysen, die nur zwischen dem Spielertyp (der ja ordinal nach Höhe der Investition aufsteigt) und der erhaltenen Bestrafung einen negativen Zusammenhang feststellen können, nicht aber zwischen Spielertyp und den ausgeteilten Sanktionen.

	Fehr & Gächter 2000, n = 112		Fehr & Gächter 2002, n = 216		Nikiforakis 2008, n = 96	
	Gesamt gegebene Strafe (P)	Gesamt erhaltene Bestrafung (RP)	P	RP	P	RP
Spielertyp	n.s.	-0,51**	n.s.	-0,46**	n.s.	n.s.
Gesamt gegebene Strafe (P)	-	n.s.	-	n.s.	-	n.s.

Tabelle 18: Spearman-Korrelationen (zweiseitig) zwischen Spielertyp, Strafen und erhaltener Bestrafung für drei Studien (n = 112; ** = auf dem 0,01-Niveau signifikant)

2.2 Diskussion Hypothese 1

Während Hypothese 1a (Kooperierende strafen mehr und häufiger als Trittbrettfahrer) nur teilweise bestätigt werden kann, zeigen sich doch Tendenzen in Richtung der Bestätigung: so ist der Test auf Unterschiede signifikant und in der Reihenfolge der Strafpunkte ordnen sich die Spielertypen fast genau wie theoretisch erwartet an, nämlich R5, K, R4, R3, R2, R1. Dies macht es wünschenswert, diese Hypothese genauer zu beleuchten, da die Abhängigkeit dieser Verhaltensdimensionen noch kaum untersucht ist.

Für eine Bestätigung der Hypothese 1b (Trittbrettfahrer werden mehr und häufiger bestraft als Kooperierende) finden sich dagegen reichlich Belege. Dieses Ergebnis einer absteigenden Bestrafung mit ansteigenden Einzahlungen ist theoriekonform und ist in der Literatur (ohne Differenzierung nach Spielertypen) bereits ausführlich diskutiert worden (Herrmann et al. 2008; Fehr & Gächter 2000; Fehr & Gächter 2002). Die vergleichsweise vielen nicht signifikanten paarweisen Vergleiche in den Post-hoc-Tests (Tabelle 17) sind darauf zurückzuführen, dass zwar die Studien von Fehr & Gächter (2000) und Fehr & Gächter (2002) genau den erwarteten negativen Zusammenhang zwischen Höhe der Einzahlungen und Höhe der Bestrafung zeigen (Abbildung 28), in Nikiforakis (2008) dagegen vor allem Spieler mit hohen Einzahlungen bestraft werden.

Diese hohen Strafen für Reziprokatoren des Typs 3, 4 und 5 in Nikiforakis (2008) überraschen. Sie könnten ein Ausdruck von Gegenbestrafung oder antisozialer Bestrafung sein, da Menschen nicht immer rational bestrafen. Die theoretische Erwartung ist zwar, dass Sanktionen gegen diejenigen gerichtet werden, die am wenigsten kooperieren. Allerdings findet man auch das Gegenteil, nämlich die Bestrafung von Individuen, die sich sozial verhalten,

d.h., die gleich oder sogar mehr als der Strafende in das Gemeingut einzahlen (Herrmann et al. 2008).

Antisoziale Bestrafung, d.h. Bestrafung von Personen, die mehr als der Strafende einzahlen) können möglicherweise bis zu 21 % aller Bestrafungskosten zugerechnet werden (Nikiforakis 2008). Auch eine zweite Art von Bestrafung – *Gegenbestrafung* – wird gefunden. Hier werden besonders diejenigen bestraft, die den Bestrafenden bestraft haben. Die Gründe dafür könnten einerseits im Rachedenken liegen, andererseits auch strategische Erwägung sein, um weitere Bestrafung für sich selbst zu vermeiden (Nikiforakis 2008). Ein solches Verhalten zerstört einen Großteil der Motivation zu kooperieren und gleichzeitig einen großen Teil der Wohlfahrtsgewinne (vgl. Frey & Rusch 2012). Das obige Ergebnis zeigt, dass Strafen auch unabhängig von der Kooperationsbereitschaft eingesetzt werden, und hier kein Automatismus vorliegt – viele Investitionen in das Gemeingut bedeuten eben nicht automatisch Straffreiheit.

2.3 Resultate für Hypothese 2: Spielertypen und Straftypen treten in bestimmten Kombinationen auf

Hypothese 2: Die beiden Strategie-Dimensionen Spielertyp und Straftyp treten bevorzugt in bestimmten Kombinationen auf.

Die Definition der Straftypen (siehe Tabelle 5) erfolgt analog zu den Spielertypen, d.h. Straftyp 1 bestraft überhaupt nicht, während Straftyp 7 sehr viel in Strafen investiert. Diese Reihung gilt ebenso für die Bestrafttypen (siehe ebenfalls Tabelle 5) – Bestrafttyp 0 wird überhaupt nicht, Bestrafttyp 7 sehr stark bestraft.

Die Resultate bezüglich Hypothese 1 weisen bereits auf eine weitgehende Unabhängigkeit zwischen den Dimensionen Spielertyp und Straftyp hin. Dies bestätigen auch die folgenden drei Abbildungen (30, 31 und 32). Sie zeigen für jeweils eine Studie (Fehr & Gächter 2000; Fehr & Gächter 2002; Nikiforakis 2008) relativ gleichmäßige Anteile an Straftypen für Spielertypen. Das gilt insbesondere für den Spielertyp K (Kooperierender) und mit Einschränkungen auch für R4 und R5. Deutlich wird auch, dass Trittbrettfahrer nicht strafen, da für diesen Spielertypus, falls existent, nur der Straftyp 1 (keine Strafen) auftritt (Abbildung 31).

Im Gegensatz zu Fehr & Gächter (2000) treten bei Fehr & Gächter (2002) (Abbildung 30) alle Spielertypen auf.

Straftypen nach Spielertyp aufgeschlüsselt (Fehr et al. 2000)

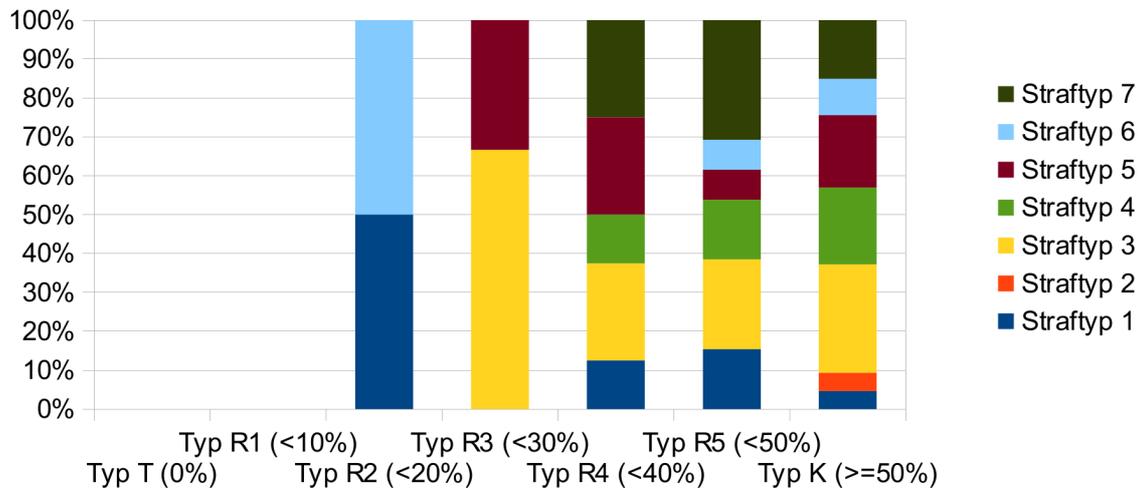


Abbildung 30: Anteil der Straftypen an den sieben Spielertypen in Prozent (Fehr et al. 2000)

Kooperierende (K) sind dabei der einzige Spielertyp, bei dem alle Straftypen vorkommen. Im Gegensatz dazu bestrafen Trittbrettfahrer überhaupt nicht (nur Straftyp 1) und Reziprokatoren R2 bestrafen sehr stark (nur Straftyp 7).

Straftypen nach Spielertyp aufgeschlüsselt (Fehr et al. 2002)

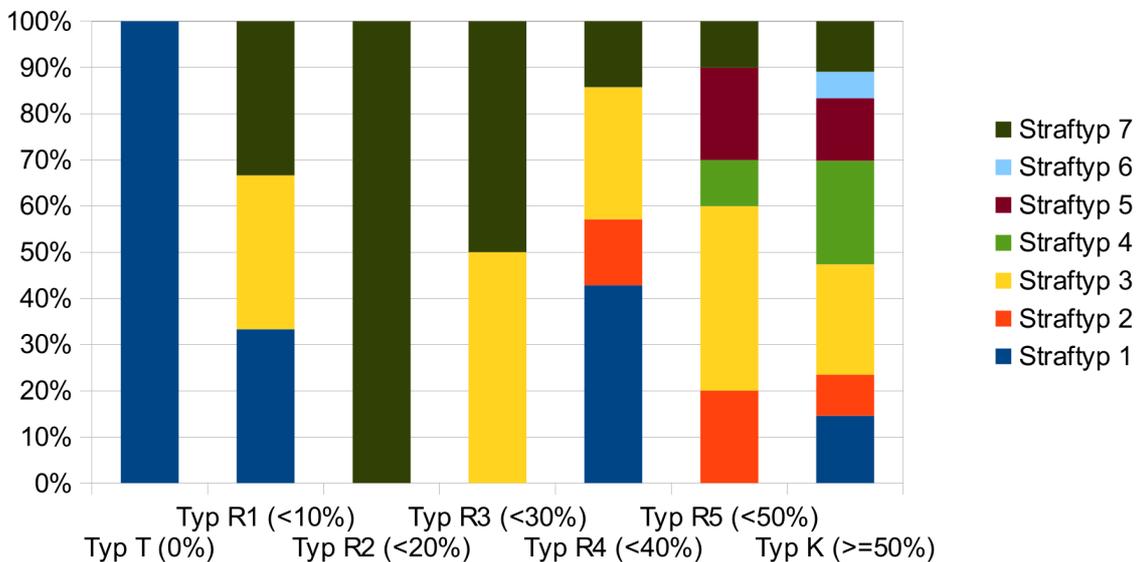


Abbildung 31: Anteil der Straftypen an den sieben Spielertypen in Prozent (Fehr et al. 2002)

Dabei ist zu beachten, dass die Zahlen für T und R1-5 nicht belastbar sind, da sie gesamt nur aus 144 (24%) Individuen bestehen. Trittbrettfahrer und R2 sind beispielsweise jeweils nur sechs Spieler. Den Hauptteil machen Kooperierende aus (n = 1152, 76%).

Wie in der ersten Studie kommt auch bei Nikiforakis (2008) der Typ Trittbrettfahrer nicht vor:

Straftypen nach Spielertyp aufgeschlüsselt (Nikiforakis et al. 2008)

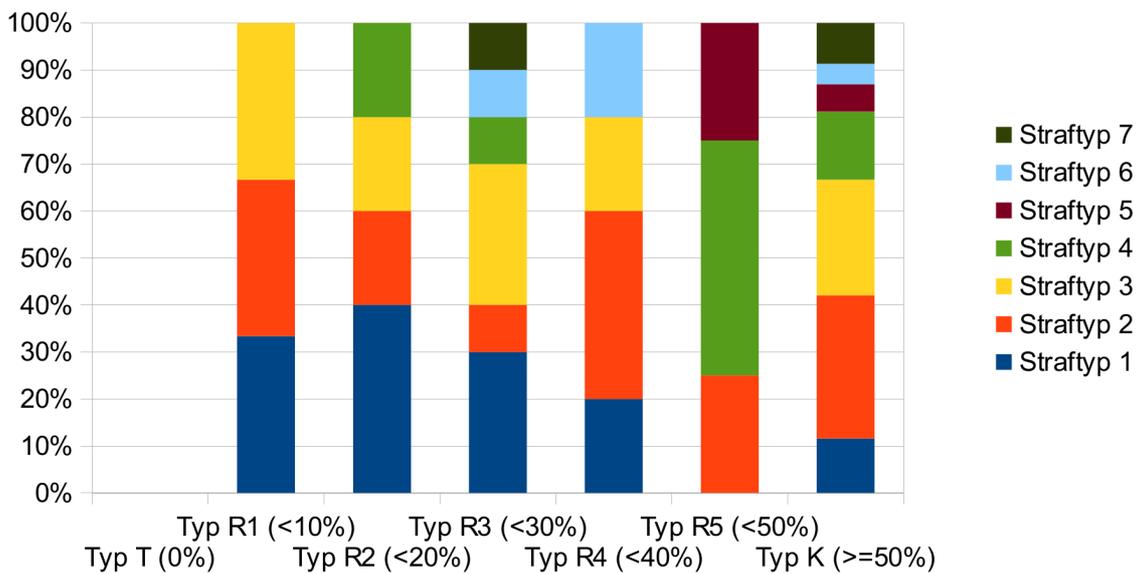


Abbildung 32: Anteil der Straftypen an den sieben Spielertypen in Prozent (Nikiforakis et al. 2008)

In Abbildung 32 zeigen sich abnehmende Tendenzen für Straftyp 1 mit zunehmender Kooperationsbereitschaft. Auch höhere Strafen (Straftyp 5, 6 und 7) kommen erst ab Spielertyp R3 vor.

Spearman-Korrelationen zwischen Spielertypen und Straftypen zeigen nur schwache Zusammenhänge: in der Studie von Fehr & Gächter (2000) ist kein Zusammenhang feststellbar (n.s.), bei Fehr & Gächter (2002) ist er praktisch kaum ausgeprägt (Spearman's rho = 0,065, $p < 0,001$) und nur bei Nikiforakis (2008) strafen kooperativere Spielertypen auch mehr (Spearman's rho = 0,13, $p < 0,001$). Über alle Studien ist dieser Zusammenhang dann auch entsprechend schwach (Spearman's rho = 0,082, $p < 0,001$).

Dieser Befund einer weitgehenden Unabhängigkeit von Spielertyp und Straftyp wird auch dadurch erhärtet, dass sich 45,3% der Varianz des Strafverhaltens (gesamt vergebene Strafpunkte) allein durch die unabhängigen Variablen *Gesamteinzahlungen* und *Verdienst in Prozent des maximal möglichen Verdienstes* erklären lässt. Fügt man dieser linearen Regressionsanalyse *Spielertyp* als weitere unabhängige Variable hinzu, so „verbessert“ sich die Modellgüte lediglich um 0,1%.

2.4 Diskussion Hypothese 2

Hypothese 2 – Spielertyp und Straftypen treten bevorzugt in einigen Kombinationen auf – muss daher abgelehnt werden. Es sind – von den Häufigkeiten ihres gemeinsamen Auftretens her – keine Kombinationen zu erkennen, die typisch nur bei bestimmten Kombinationen auftreten würden. Lediglich die Kombination Trittbrettfahrer gekoppelt mit keinerlei Austeilen von Strafen ist in einer Studie (Fehr & Gächter 2002) deutlich vertreten. Auf Grund des Mangels an Trittbrettfahrern in den anderen Studien ist dieses Ergebnis jedoch nicht reproduzierbar und kann deshalb nicht als robust gelten. Ein weiteres Ergebnis zeigt sich im Partner-Design von Nikiforakis (2008): hier scheint es Tendenzen zu mehr Strafen bei Kooperierenden zu geben, die sich statistisch auch erhärten lassen, der Zusammenhang ist jedoch insgesamt schwach.

Eine mögliche Interpretation dieser Ergebnisse ist, dass Spielertypen und Straftypen verschiedene Dimensionen eines Strategieraumes sind, die nicht unbedingt verbunden sein müssen. Mit anderen Worten, das Kooperationsniveau ist weitgehend unabhängig von der Art und Weise zu strafen. Tendenziell bestrafen allerdings Trittbrettfahrer und Reziprokatoren des Typs 1 im Vergleich zu anderen Spielertypen weniger. Dies stimmt mit den theoretischen Erwartungen überein, da mit diesen Strategien vor allem Kosten minimiert werden – und dazu gehören eben sowohl Investitionen ins öffentliche Gut als auch Sanktionen.

2.5 Resultate für Hypothese 3: Gruppenzusammensetzung beeinflusst das Strafverhalten

Hypothese 3: Je nach Gruppenzusammensetzung ändert sich das Strafverhalten.

Da sowohl Fehr & Gächter (2000) als auch Fehr & Gächter (2002) das Stranger-Design benutzen, ist eine Untersuchung bezüglich Hypothese 3 bei ihnen nicht sinnvoll: da sich jede Runde die Gruppenzusammensetzung ändert, kann keine Rede von konstanten Gruppentypen und möglichen Strategien sein. Diese beiden Studien bleiben daher bei der Analyse bezüglich Hypothese 3 außen vor. Zusätzlich ist bei den folgenden Daten zu berücksichtigen, dass 15,6% der Teilnehmer nie strafen, obwohl sie die Möglichkeit dazu haben, und dass 11,5% nie bestraft werden.

Die folgenden beiden Abbildungen (33 und 34) zeigen für alle vorkommenden Gruppentypen ohne Trittbrettfahrer (da diese nicht vorkommen) die Strafen nach Spielertyp aufgeschlüsselt. Dabei ist berücksichtigt, dass Spielertypen unterschiedlich häufig vorkommen: die Summe der Bestrafung wird durch die Anzahl Spieler dieses Spielertyps geteilt.

Sowohl Gruppentyp 1 als auch 3 bestehen ausschließlich aus einem Spielertyp (K bzw. R1-R5), während Gruppentyp 2 eine Mischung aus Kooperierenden (K) und Reziprokatoren (R1-R5) ist. Während in reinen Kooperierenden-Gruppen definitionsgemäß nur Kooperierende bestrafen und im Gruppentyp R nur R1 bis R5, sind dies in heterogenen Gruppen (hier Gruppentyp 2) sowohl K als auch R-Spielertypen.

Gegebene Bestrafung per Spielertyp und drei Gruppentypen

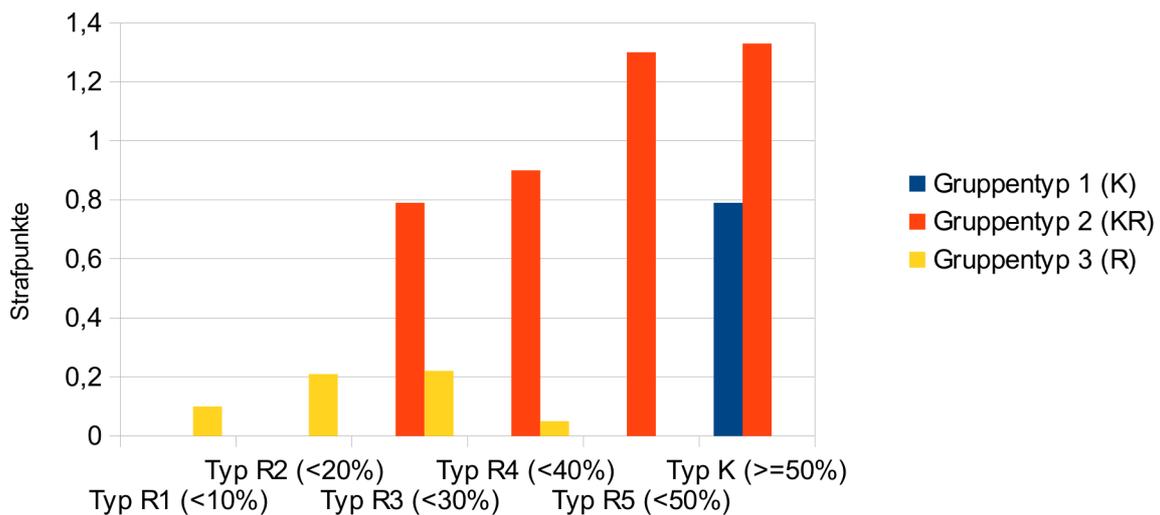


Abbildung 33: Bestrafung pro Spielertyp und Gruppentyp (pro Häufigkeit Spielertyp)

Die Strafhöhe pro Spieler pro Typ steigt für gemischte Gruppen (rote Balken) monoton an, d.h. Kooperierende strafen hier am meisten. Bei den Reziprokatoren (Gruppentyp 3, gelbe Balken) übernehmen vor allem Spielertyp 2 und 3 die Bestrafung. In reinen Kooperierenden-Gruppen (blaue Balken) ist bei Kooperierenden die Strafhöhe deutlich geringer als in heterogenen Gruppen (Gruppentyp KR). Dieser Unterschied ist signifikant (Mann-Whitney-Test, $n = 690$, $p < 0,001$).

Wer wird nun bestraft? Abbildung 34 zeigt, dass dies in reinen Reziprokatoren-Gruppen (Typ 3) vor allem Spielertyp 4 ist. In gemischten Gruppen trifft es die Spielertypen 4 und 5.

Erhaltene Bestrafung per Spielertyp und drei Gruppentypen

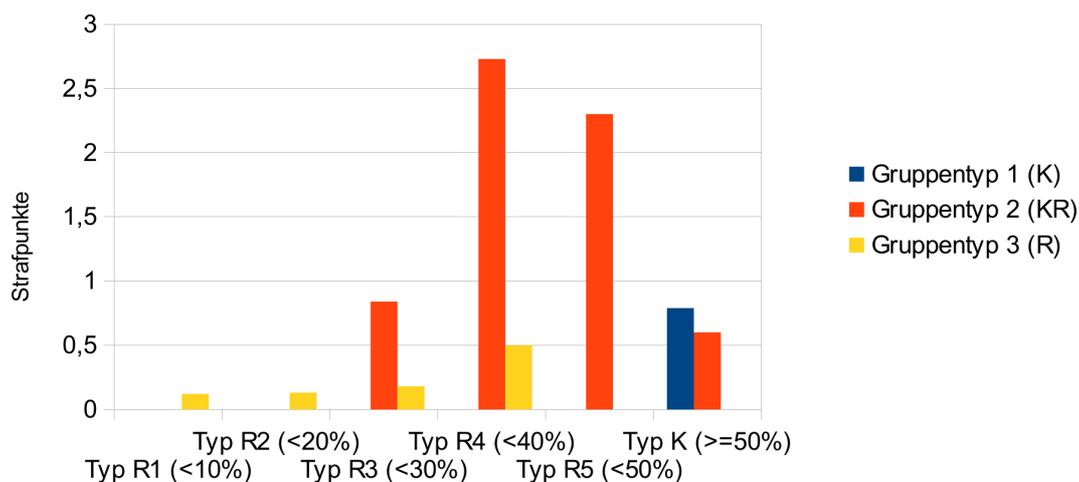


Abbildung 34: Erhaltene Bestrafung pro Spielertyp und Gruppentyp (pro Häufigkeit Spielertyp)

Dabei ist wiederum zu berücksichtigen, dass Kooperierende bei diesen Daten 71,9% aller Teilnehmer stellen und reine Kooperierenden-Gruppen (K) schon 58,3% aller Gruppentypen ausmachen. Kooperierende werden nicht signifikant mehr in KR-Gruppen als in reinen Kooperierenden-Gruppen (K) bestraft (Mann-Whitney-Test, $n = 690$, n.s.).

2.6 Diskussion Hypothese 3

Abbildung 33 zeigt, dass diejenigen Spielertypen, die weniger investieren (R1 bis R4), sich lediglich in den reinen Reziprokatoren-Gruppen an der Bestrafung beteiligen. In den gemischten Gruppen (KR) übernehmen das zwar überlappend ebenfalls R3 und R4, aber vor allem R5 und Kooperierende (K). Hier zeigt sich ein Zusammenhang zwischen Bestrafung, Spielertypen und Gruppentypen: je höher die Kooperationsbereitschaft, desto höher auch die Bereitschaft, Strafkosten in heterogenen Gruppen zu übernehmen. In homogenen Gruppen verschiebt sich die Strafbereitschaft dagegen hin zu Spielertypen, die weniger in das Gemeingut investieren.

Dabei trifft es Spielertyp R4 am stärksten – gegen ihn sind sowohl in homogenen als auch heterogenen Gruppen (Gruppentyp KR (2) und R (3)) die meisten Strafen gerichtet. Die Tatsache, dass Kooperierende in homogenen Gruppen mehr bestraft werden als in gemischten Gruppen kann in zweifacher Hinsicht interpretiert werden: Entweder verschieben sich die Ziele von Bestrafenden in gemischten Gruppen hin zu Spielertypen, die weniger investieren, oder es kommt auch in Gruppen mit sehr hohen Einzahlungen zu Gegenbestrafung.

Damit kann die Hypothese 3 – je nach Gruppenzusammensetzung ändert sich das Strafver-

halten – unter Berücksichtigung der erwähnten Einschränkungen bezüglich Gruppenanzahl und Untersuchung nur an einem Datensatz – als bestätigt gelten.

2.7 Resultate für Hypothese 4: Sanktionen sind in ihrem zeitlichen Verlauf unabhängig vom Spielertyp und treten vor allem zu Beginn von Interaktionen auf

Hypothese 4: Unabhängig vom Spielertyp wird anfangs im Vergleich zum späteren Verlauf überproportional viel bestraft.

Die angesprochene Zeitkomponente sollte relativ unabhängig vom Spielertyp sein, da bisherige Ergebnisse eine sehr deutlich abfallende Tendenz von Strafen in den ersten Runden zeigen. Die folgenden Abbildungen (35, 36 und 37 für die drei Studien) bestätigen diese Vermutung. In allen drei Studien ist ein klarer Abwärtstrend für alle Spielertypen zu erkennen. Jeder der sieben Spielertypen – falls er überhaupt bestraft – zeigt in allen drei Studien (egal ob Stranger- oder Partner-Treatment) einen Abwärtstrend. Es ist lediglich ein Ausreißer zu erkennen (Abbildung 37 in Nikiforakis 2008, Spielertyp R2 in Runde 6).

Strafe pro Spielertyp und Runde (Fehr et al. 2000)

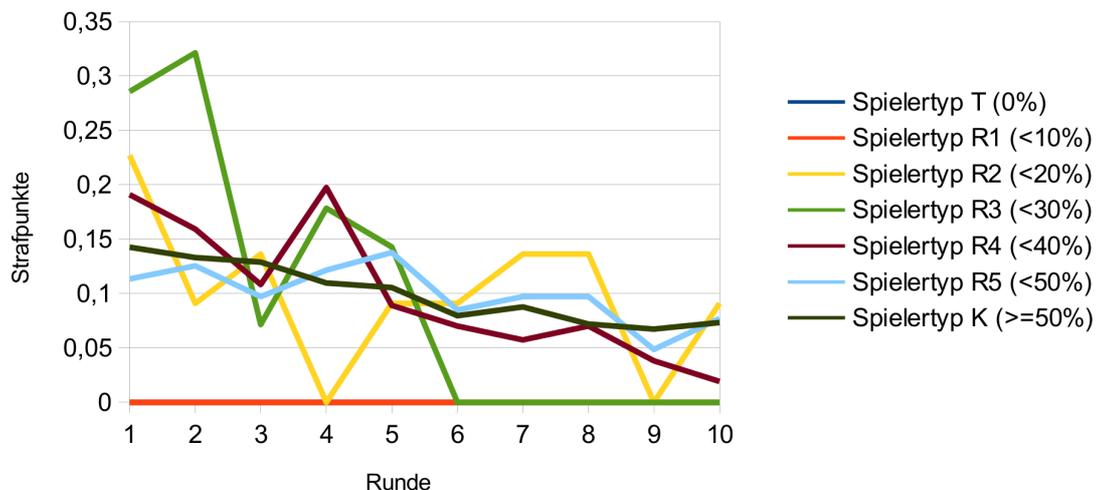


Abbildung 35: Strafpunkte pro Spielertyp und Runde (Fehr et al. 2000)

Wie Abbildung 35 zeigt, vergeben die Spielertypen R1 und R3 sogar ab Runde sechs überhaupt keine Strafpunkte mehr.

In Abbildung 36 liegt jeder Spielertyp nach kurzen Anstiegen in der zweiten und dritten Runde teilweise deutlich unter dem Anfangsniveau.

Strafe pro Spielertyp und Runde (Fehr et al. 2002)

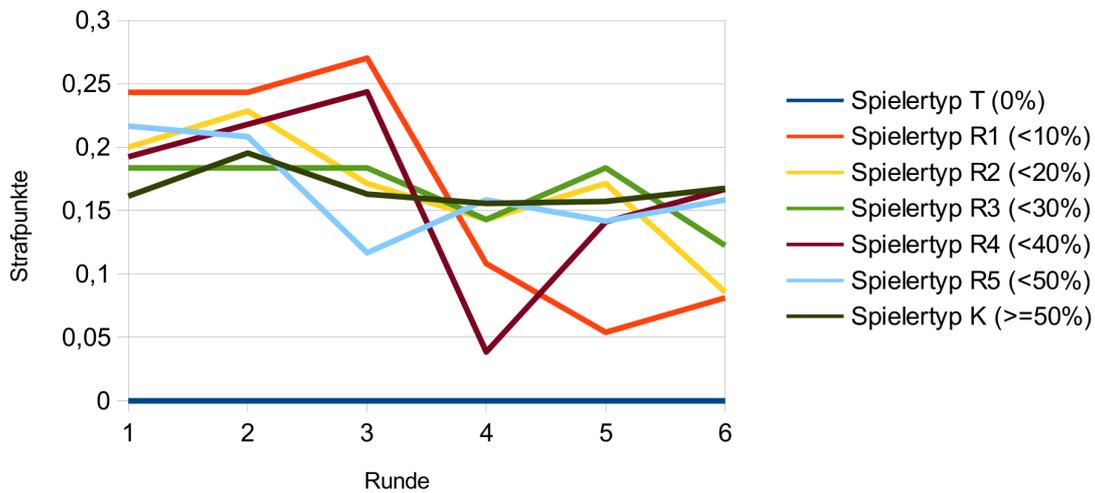


Abbildung 36: Strafpunkte pro Spielertyp und Runde (Fehr et al. 2002)

Die folgende Abbildung 37 belegt diese Muster ein weiteres Mal – hier straft R1 sogar ab Runde vier nicht mehr, alle übrigen bewegen sich bereits ab Runde fünf auf einem durchgängig niedrigen Niveau.

Strafe pro Spielertyp und Runde (Nikiforakis et al. 2008)

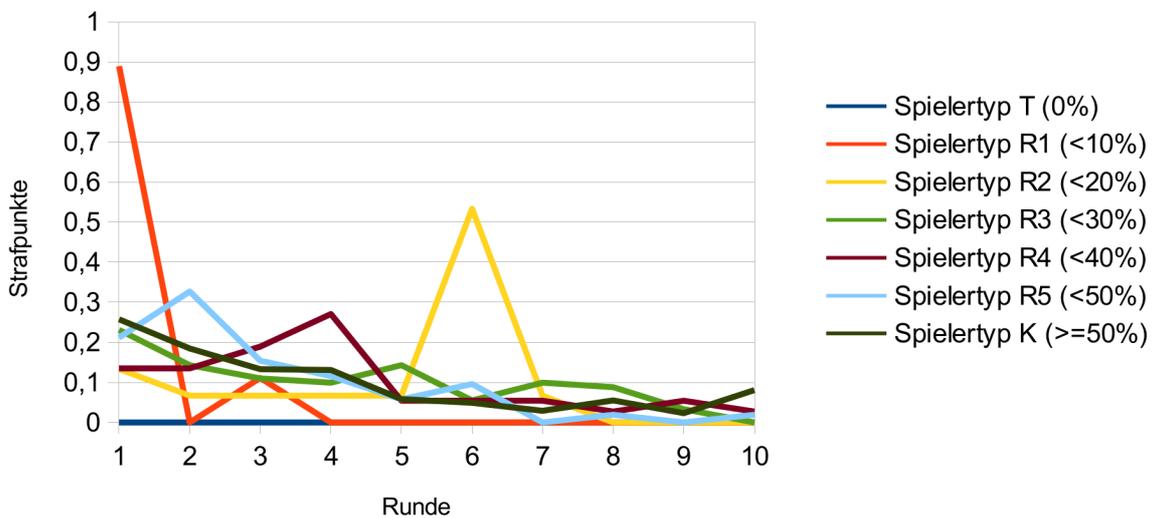


Abbildung 37: Strafpunkte pro Spielertyp und Runde (Nikiforakis et al. 2008)

Dies wird durch Korrelationsanalysen bestätigt. Es besteht über alle drei Studien ein negativer Zusammenhang zwischen Bestrafung und Runde (Spearman's rho = -0,22, p < 0,01).

2.8 Diskussion Hypothese 4

Bisherige Studien konnten demonstrieren, dass es zu einer starken Abnahme von Strafkosten kommt. Die folgende Abbildung (Abbildung 38) demonstriert diese typische Abnahme der

Kosten für Strafen nach Treatment, nicht nach Spielertyp (Daten aus Nikiforakis & Normann 2008, eigene Kalkulationen und Abbildung des Autors, übernommen aus Frey & Rusch 2012).

Abnehmende Kosten für Bestrafung

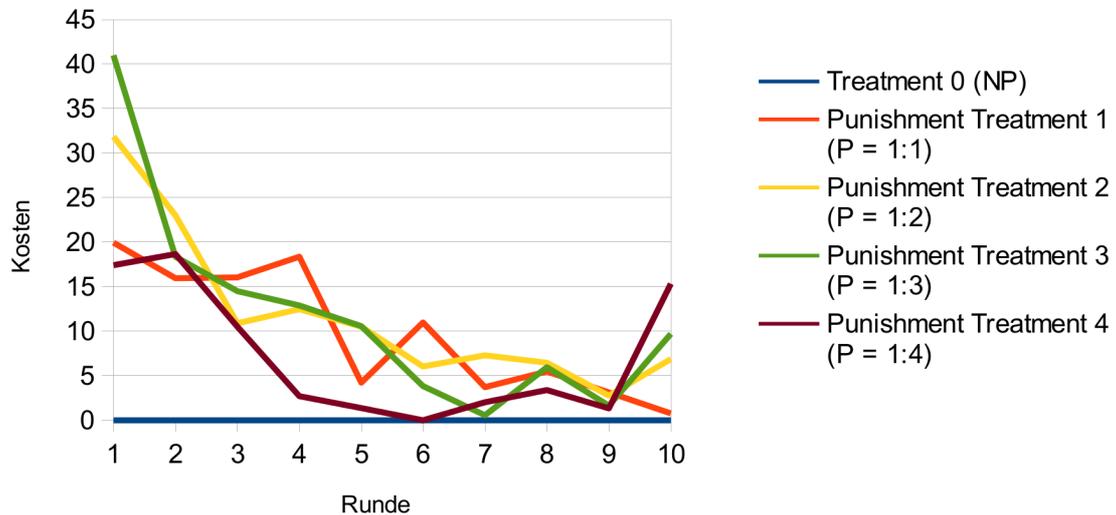


Abbildung 38: Abnehmende Kosten für Bestrafung im Zeitverlauf nach Effektivität der Sanktionen

Dieses Resultat kann generalisiert werden. Wie die nächste Abbildung zeigt (Abbildung 39, Daten aus Fehr & Gächter 2000; Sefton et al. 2007 und Nikiforakis & Normann 2008; eigene Kalkulationen und Abbildung des Autors, übernommen aus Frey & Rusch 2012), sind die Kosten der Bestrafung in den ersten Runden sehr hoch.

Abnehmende Ausgaben für Bestrafung

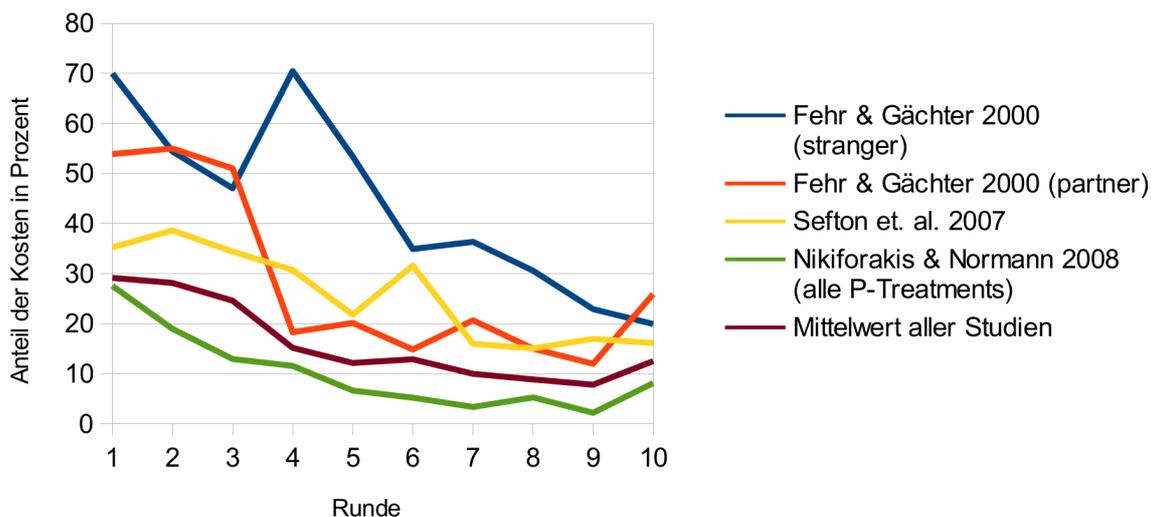


Abbildung 39: Abnehmende Ausgaben für Bestrafung im Zeitverlauf im Studienvergleich

Bis zu 70% (27-70%) aller Verdienste werden in der ersten Runde zerstört, der Mittelwert für die ersten fünf Runden liegt bei 22 %. Der Mittelwert für alle Runden in allen Bestrafungs-Treatments für alle Studien ist 16 %. In den letzten fünf Runden müssen dagegen nur noch 10% der Verdienste für Strafen ausgegeben werden. Aus einer evolutionären Perspektive sind zehn Runden eines Laborexperiments bezüglich der Entwicklung von Kooperation eine sehr kurze Zeitspanne. Zusätzlich sind die Sanktionen im Vergleich zu evolutionären Szenarien extrem teuer.

Der in Abschnitt C 1 vorgestellte Datensatz zu PGG kann diese Belege für effiziente Sanktionen mit einer weiteren Berechnung bestätigen: Im Durchschnitt der Studien verlieren Trittbrettfahrer (T) 62% ihres Verdienstes, wenn man diesen vor und nach Bestrafung vergleicht. Vor Abzug der Strafkosten setzt sich der Verdienst aus dem privat behaltenen Betrag und der Auszahlung aus dem Gemeingut zusammen. Dieser Verdienst wird nach Bestrafung durch die eigenen Kosten für Strafen und gegebenenfalls der Reduktion durch die Bestrafung anderer vermindert. Für die übrigen Spielertypen (siehe C 2.1) lauten die Zahlen entsprechend 41,3% (R1), 41,5% (R2), 28,7% (R3), 24,6% (R4), 29,3% (R5) und 16,3% (K). Diese Aufschlüsselung nach Spielertypen zeigt, dass die Bestrafung stark von den Einzahlungen in das Gemeingut abhängt, da die Verluste von T (62%) absteigend über R1 bis R5 zu K (16%) laufen.

Auf der Basis der eben diskutierten zusätzlichen Ergebnisse (Frey & Rusch 2012) zeigt sich demnach eine vollständige Bestätigung von Hypothese 4 *auch nach Aufschlüsselung nach Spielertypen*: Bestrafung ist an die ersten Runden gekoppelt – hier werden die Normen festgelegt und Normüberschreitungen geahndet. Durch die hohen Kosten der Bestrafung ändern die meisten Spieler schnell ihr Verhalten und kooperieren fortan. Dies macht wiederum Bestrafung überflüssig – zumindest bis Trittbrettfahrer wiederum versuchen, Kooperierende auszunutzen. Muss die Drohung einer Strafe dann wiederum in die Tat umgesetzt werden, bleiben – wenn dies selten genug sein muss – Strafkosten niedrig und das Kooperationsniveau hoch. Typischerweise sind Sanktionen in der letzten Runde notwendig – hier versuchen viele bislang reziprok Spielende und „gezähmte“ Trittbrettfahrer Einnahmen mitzunehmen; dies wird allerdings nicht geduldet, was durch einen scharfen Anstieg in den Strafen sichtbar wird. Diese Bereitschaft ist nicht selbstverständlich: sind die Parameter so, dass keiner oder wenige bereit sind, zu strafen – beispielsweise auf Grund von hohen Kosten oder der Angst vor Gegenbestrafung – so nehmen die Trittbrettfahrer überhand und die Einzahlungen sinken.

2.9 Resultate für Hypothese 5: Die Kombination aus Sanktionen und hohen Einzahlungen in ein Gemeingut führt zu Wettbewerbsvorteilen für Kooperierende

Hypothese 5: Hohe Einzahlungen führen zusammen mit Bestrafung der Trittbrettfahrer zu einem Wettbewerbsvorteil für kooperative Individuen.

Die untenstehende Tabelle 19 zeigt für fünf PGG-Studien, soweit Daten verfügbar sind, den Zusammenhang zwischen Einzahlungen in das öffentliche Gut (Gesamteinzahlungen, *total contributions*) und dem Verdienst vor und nach Abzug der Strafkosten (sei es durch eigene Strafen oder durch erhaltene Bestrafung verursacht). In allen Studien ist hier ein deutlicher positiver Zusammenhang erkennbar – die hochsignifikanten Korrelationskoeffizienten reichen von 0,39 (stranger, 6 Runden) bis 0,95 (partner, 50 Runden).

Studie	Pearson-Korrelationskoeffizient Gesamte Einzahlungen / Gesamtverdienst vor Strafe	Pearson-Korrelationskoeffizient Gesamte Einzahlungen / Gesamtverdienst nach Strafe	Pearson-Korrelationskoeffizient Gesamtverdienst vor Strafe / Gesamtverdienst nach Strafe
Fehr & Gächter 2000, n = 112	n.s.	0,47**	0,46**
Fehr & Gächter 2002, n = 216	-0,76**	0,39**	-0,22**
Nikiforakis 2008, n = 96	0,64**	0,59**	0,83**
Gächter et al. 2008, n = 96	0,92**	0,95**	0,99**
Nikiforakis 2008, n = 192	-0,40**	nicht verfügbar	nicht verfügbar

*Tabelle 19: Pearson-Korrelationen (zweiseitig) zwischen Einzahlungen und Verdienst für drei Studien (** = auf dem 0,01-Niveau signifikant)*

Der Zusammenhang zwischen Gesamteinzahlungen und Verdienst vor Bestrafungskosten ist dagegen nicht eindeutig. Man findet sowohl negativ als auch positiv signifikante Korrela-

tionen, sowie einen nicht signifikanten Zusammenhang. Der Zusammenhang zwischen Gesamtverdienst vor und nach Bestrafungskosten ist tendenziell positiv. Bis auf ein Stranger-Design mit 6 Runden (Fehr & Gächter 2002) sind alle Korrelationen signifikant positiv.

2.10 Diskussion Hypothese 5

In allen hier untersuchten Datensätzen zeigen sich hohe Korrelationen zwischen Einzahlungen und dem Verdienst *nach* Abzug der Strafkosten. Dieser Zusammenhang besteht allerdings nicht zwischen Einzahlungen und Verdienst vor Abzug der Strafkosten. Zwischen Verdienst vor und nach Abzug der Kosten könnte ein positiver Zusammenhang bestehen, der allerdings noch nicht als robust zu bewerten ist.

Die untenstehende Abbildung (Abbildung 40) zeigt sogar, dass die Strafen bei ausschließlicher Betrachtung der Stranger-Treatments die Spielertypen perfekt reflektieren.

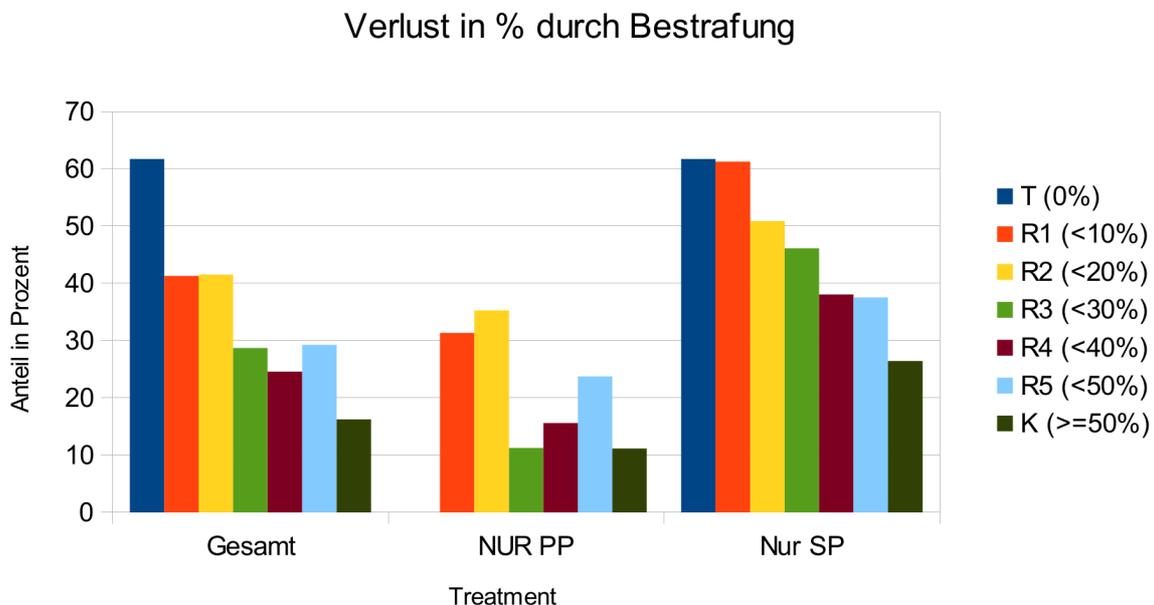


Abbildung 40. Verlust des Verdienstes in Prozent vor und nach den Bestrafungskosten für Spielertypen (getrennt nach Treatment und als aggregierte Werte)

Auffällig ist, dass selbst Personen, die mehr als 50% in das Gemeingut investieren, bis zu 25% (SP) ihres Verdienstes durch Bestrafung (eigene Strafen und Bestrafung) verlieren.

Die Ausnahme hierbei ist immer die letzte Runde, in der die Einzahlungen in fast allen Treatments stark abnehmen. In Treatments mit Bestrafung wird dies von gleichzeitig stark steigenden Sanktionen begleitet. Dieses scheinbar irrationale Verhalten der Bestrafung in der letzten Runde stellt ein bis dato nicht gelöstes Rätsel dar – denn wenn Bestrafung auf Verhaltensänderung abzielt, ist es sinnlos in der letzten Runde noch zu bestrafen. Ich mache einen Erklärungsvorschlag in Abschnitt E 2.

Das bedeutet zunächst, dass Trittbrettfahrer durchaus mit Kooperierenden im Verdienst mithalten könnten – allerdings nur, bis sie von den teils empfindlichen Strafen (Strafverhältnis 1:4 bzw. 10% des Einkommens für jeden vergebenen Strafpunkt) getroffen werden. Auch wenn Strafkosten den Verdienst schmälern – es sieht so aus, als wäre die Kombination aus hohem Kooperationsniveau durch die Verwandlung von Trittbrettfahrern in Spieler, die ihren Beitrag leisten, die erfolgreichste Strategie, was den Verdienst betrifft.

Der evolutionäre Bezug dieser ökonomischen Daten zeigt sich auch in den ansteigenden Korrelationskoeffizienten, die von kurzen Interaktionen (6 Runden) in wechselnden Gruppen (0,39) über wechselnde Gruppen und 10 Runden (0,47) und konstanter Gruppenzusammensetzung und 10 Runden (0,59) hin zu fast perfekter Korrelation (0,95) bei ausreichender Zeit (50 Runden) und konstanter Gruppenzusammensetzung laufen. Hier wird sowohl die Trennung von *in-* und *outgroup* deutlich, als auch der längere Zeithorizont kooperativer Interaktionen in unserer Jäger- und Sammler-Vergangenheit. Die Hypothese 5 – die Kombination aus homogenen Gruppen mit Bestrafungsoption macht Kooperation egoistischen Strategien überlegen – kann damit bestätigt werden.

3 Ergebnisse Studie 3: Der Erfolg Kooperierender

Anmerkung: Da in dieser Studie sieben Arbeitshypothesen bearbeitet werden, ist aus Lesbarkeitsgründen die Diskussion direkt an die jeweilige Hypothese angeschlossen und nicht wie bei der Studie 1 im Block danach.

3.1 Resultate PGG für Hypothese 1: Egoistischere Strategien sind – ohne zusätzliche Randbedingungen – kooperierenden Strategien überlegen

Die folgende Abbildung (Abbildung 41) zeigt den tatsächlich erreichten Anteil am maximalen Profit pro Spielertyp (für eine ausführliche Beschreibung dieser Kenngröße siehe C 1) über alle Studien (ohne Nikiforakis 2008, der keine Daten für den Verdienst nach Bestrafung angibt). Konsistent mit den theoretischen Erwartungen sind die durchschnittlichen Verdienste für Trittbrettfahrer (T) mit am höchsten und liegen bei einem Median von 58,79% (Mittelwert 58,41%, Standardabweichung 6,76). Der Gesamtverdienst ist – nicht völlig durchgängig – umso niedriger, je mehr man in das Gemeingut einzahlt (R3, R1, R2, R4, R5). Diejenigen Teilnehmer, die mehr als 50 % ihrer Einnahmen in das öffentliche Gut investieren (K), liegen mit 60,68% (MW 58,30%, Stabw. 12,37) auf Rang eins und damit noch vor R1 bis R5. Die Unterschiede sind signifikant (Kruskal-Wallis-Test, $df = 6$, $n = 12182$, $p < 0,001$). In den korrigierten Post-hoc-Tests erweisen sich alle paarweisen Vergleiche mindestens auf dem

0,001-Niveau als signifikant, bis auf R1-R2, R1-R3, R3-R2 und K-T. Wird die Studie von Gächter et al. (2008) nicht berücksichtigt, da das 50-Runden Treatment zu Verzerrungen im Vergleich führen kann, so ergibt sich – nach Median gestaffelt – fast exakt die theoretisch erwartete Reihenfolge von T, 1, 2, 3, 4, K, 5. Dieses Überlegung wird im Folgenden mit der Trennung nach Treatment fortgeführt.

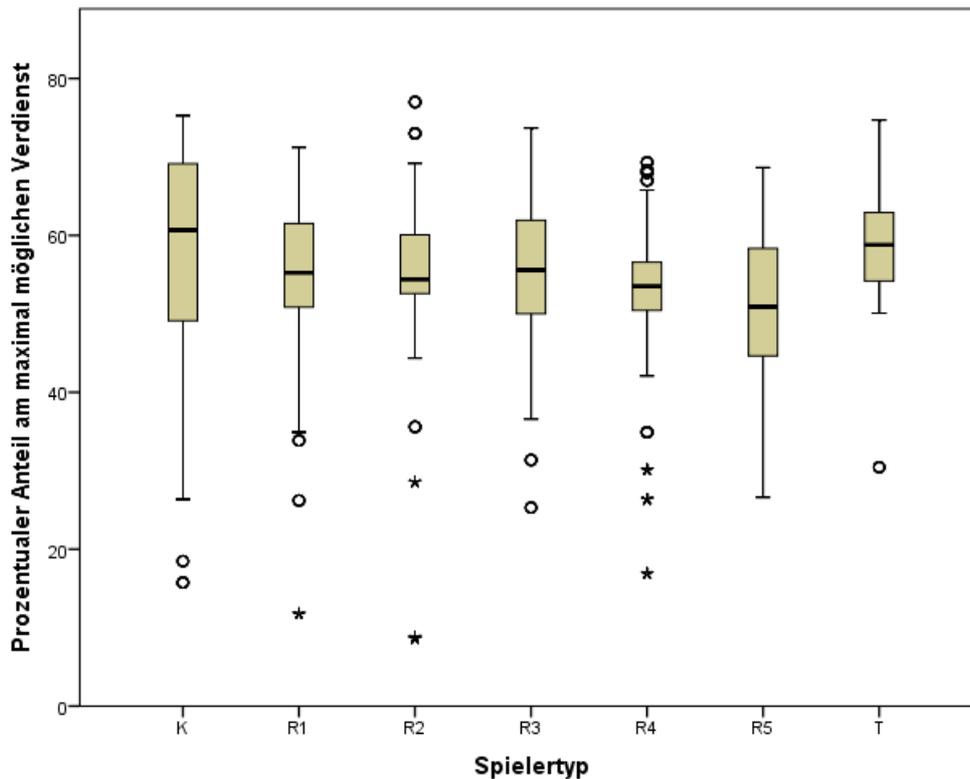


Abbildung 41: Durchschnittlicher prozentualer Anteil am maximal möglichen Verdienst für alle Spielertypen in PGG

Um herauszufiltern, welcher Anteil auf mögliche Treatment-Effekte zurückzuführen ist, differenziere ich im Folgenden nach dem Treatment (Abbildung 42). Hier bringt die Aufschlüsselung nach Treatments (Partner mit Bestrafung (PP), Stranger mit Bestrafung (SP), Partner ohne Bestrafung (PNP) und Stranger ohne Bestrafung (SNP)) einen klaren Zugewinn an Erkenntnis, da jetzt zwei Muster deutlich zu erkennen sind. In beiden Treatments mit Bestrafung zeigt sich eine klar ansteigende Tendenz des Erfolgs von Egoismus über R1 bis R5 hin zu Kooperierenden K. Für Stranger mit Bestrafung (SP) sind die Mediane für das Erfolgsmaß 47,73 (MW 46,68, Stabw. 7,72) für K; 44,37 (MW 38,44, Stabw. 7,90) für R2; 43,65 (MW 42,15, Stabw. 5,60) für R5; 43,02 (MW 42,60, Stabw. 9,59) für R4; 42,77 (MW 40,29, Stabw. 7,34) für R3 und 33,86 (MW 31,66, Stabw. 3,99) für R1. Für das Treatment Partner mit Bestrafung sind die Mediane für das Erfolgsmaß 66,33 (MW 63,66, Stabw. 10,31) für K;

56,26 (MW 54,42, Stabw. 11,45) für R5; 54,40 (MW 48,27, Stabw. 15,22) für R2; 53,50 (MW 50,55, Stabw. 7,45) für R4; 51,77 (MW 50,86, Stabw. 9,76) für R3 und 49,27 (MW 40,23, Stabw. 15,79) für R1.

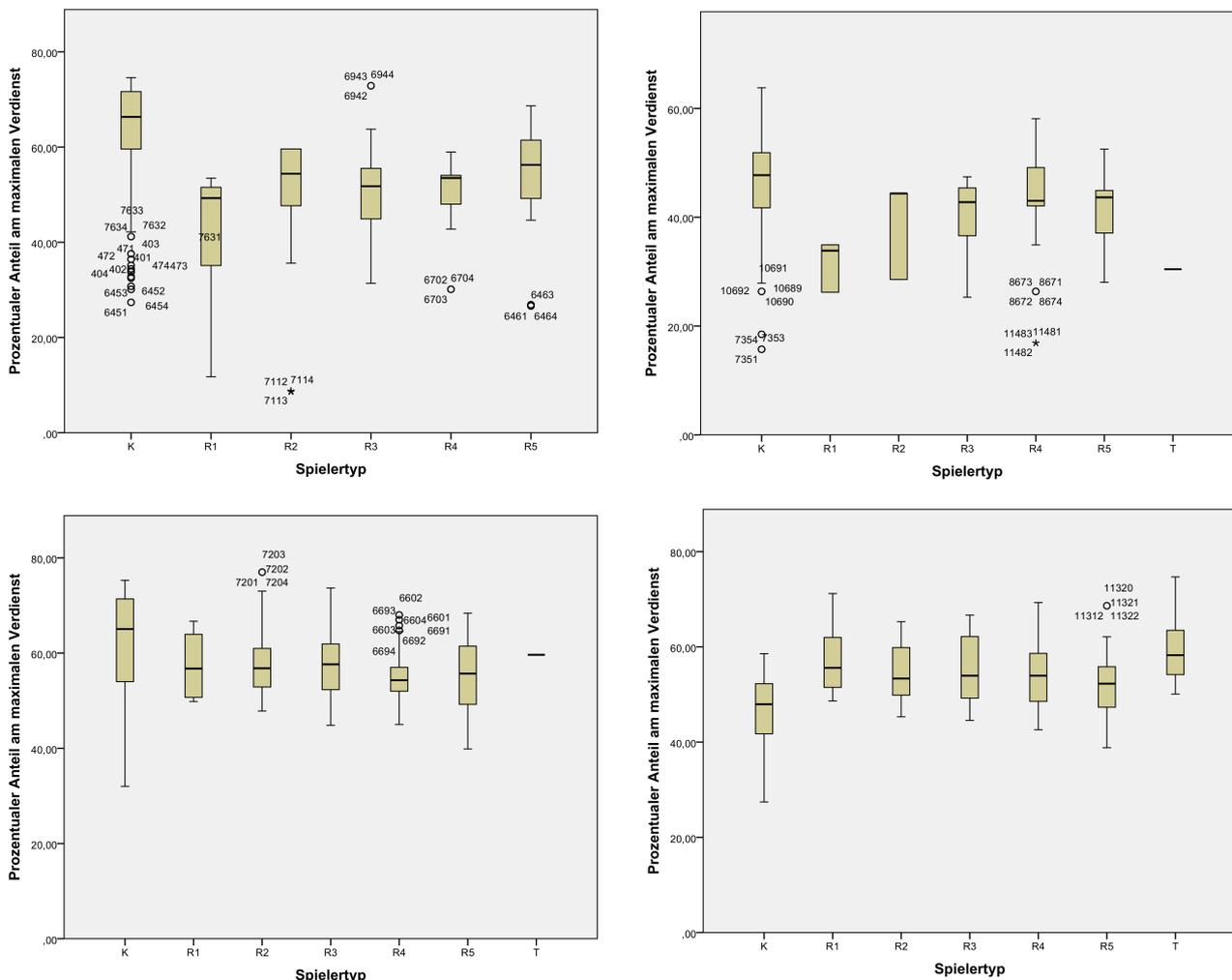


Abbildung 42: Durchschnittlicher prozentualer Anteil am maximal möglichen Verdienst für alle Spielertypen nach Treatment in PGG (im Uhrzeigersinn: Partner-P; Stranger-P; Stranger-NP; Partner-NP)

In den beiden Treatments ohne Bestrafung (PNP und SNP) kehrt sich dieses Muster exakt um, d.h. hier ist Trittbrettfahren am erfolgreichsten. Vor allem wenn keine Bestrafung möglich ist und die Gruppenzusammensetzung wechselt (SNP), zeigt sich ein sinkender Zusammenhang zwischen Erfolg und Investition in das öffentliche Gut. Die Mediane sind hier 58,26 (MW 58,90, Stabw. 5,74) für T; 55,59 (MW 56,89, Stabw. 6,08) für R1; 53,35 (MW 54,66, Stabw. 5,83) für R2; 53,94 (MW 55,26, Stabw. 6,64) für R3; 53,94 (MW 53,81, Stabw. 6,91) für R4; 52,27 (MW 51,99, Stabw. 6,15) für R5 und schließlich 47,95 (MW 46,42, Stabw. 7,46) für Kooperierende (K).

Die entsprechenden Tests auf Unterschiede sind für Partner mit Bestrafung (PP, n = 4360) und

ohne Bestrafung (PNP, n = 3790) sowie für Stranger mit Bestrafung (SP, n = 2016) und ohne Bestrafung (SNP, n = 2106) auf dem 0,001-Niveau signifikant (Kruskal-Wallis-Test, $df = 6$; $df = 5$ für PP, da hier T nicht vorkommen). In den korrigierten Post-hoc-Tests erweisen sich alle paarweisen Vergleiche zwischen Kooperierenden und Trittbrettfahrern mindestens auf dem 0,04 (PP), 0,007 (SNP) bzw. 0,05-Niveau (SP) als signifikant. Lediglich in PNP ist kein signifikanter Unterschied zwischen T und K festzustellen, was an der geringen Zahl der Trittbrettfahrer liegt. In PP, SP und PNP haben Kooperierende höhere mittlere Ränge, in SNP sind es die Trittbrettfahrer. Reziprokatoren werden hier nicht näher analysiert.

3.2 Diskussion H1 PGG

Diese Ergebnisse bestätigen die erste Hypothese. Kooperative Strategien scheinen nach diesen Daten flankierende institutionelle Maßnahmen (P) zu benötigen, um erfolgreich zu sein. Stehen solche Mittel nicht zur Verfügung, so machen Trittbrettfahrer das Rennen. Interessanterweise sind Partner-Treatments, d.h. Kooperation mit „bekanntem“ und gleichbleibenden Gruppen in der Lage, sowohl Ergebnisse mit als auch ohne institutionelle Maßnahmen zu beeinflussen. Ist Bestrafung möglich, so führt das Partner-Treatment (Partner P) dazu, dass die Gewinne noch unterschiedlicher ausfallen; ist Bestrafung nicht möglich (Partner NP), so führt das Partner-Treatment immer noch zu den höchsten Gewinnen für Kooperierende im Vergleich zum parallelen Stranger-Treatment ohne Bestrafung (Stranger NP).

Interessanterweise wird durch eine konstante Gruppenzusammensetzung mit Bestrafungsoption (PP) wirksam das Auftreten von Trittbrettfahrern verhindert – dies ist das einzige Treatment, bei dem über vier Studien keine einzige Gruppe zu Stande kommt, die nur aus Trittbrettfahrern besteht. Die Hypothese 1 – Überlegenheit für Trittbrettfahrerstrategien ohne Randbedingungen, Überlegenheit für Kooperationsstrategien mit Randbedingungen – ist damit für Labor-PGG bestätigt.

3.3 Resultate Ikarium für Hypothese 1: Egoistischere Strategien sind – ohne zusätzliche Randbedingungen – kooperierenden Strategien überlegen

Zunächst wird die Häufigkeitsverteilung der Spielertypen dargestellt:

Häufigkeiten Spielertypen im Ländervergleich für alle 28 Snapshots

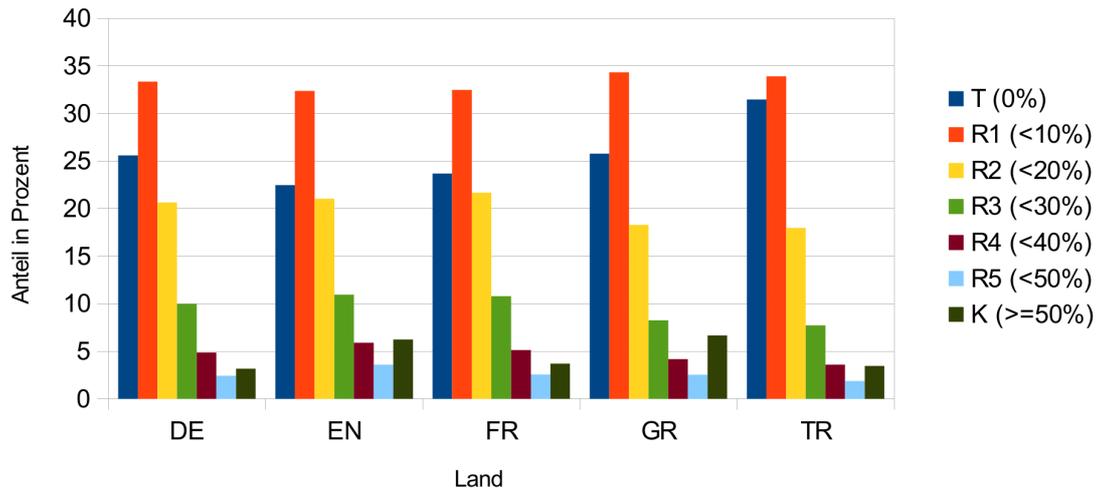


Abbildung 43. Häufigkeitsverteilung der Spielertypen in Ikariam im Ländervergleich für alle 28 Snapshots (10 Monate)

Die Entwicklung dieser Spielertypen über die Zeit ist in den Abbildungen 20, 21, 22 und 23 dargestellt, die Verteilungsmaße in den Abbildungen 18 und 19.

Es fällt auf, dass Spieler, die weniger als 10% in das öffentliche Gut investieren (R1), der am häufigsten auftretende Spielertyp sind, gefolgt von Trittbrettfahrern, R2 und R3. Dieser Befund ist über alle Länder konsistent. Kooperierende (K) sind zahlenmäßig selten vertreten (Mittelwert 4,65%), treten aber in allen Ländern häufiger als R5 auf und in EN sowie GR sogar häufiger als R4.

Wenn man die Definition von Trittbrettfahrern (üblicherweise durchgängig 0% Investitionen) als zu unrealistisch ansieht, so lässt sich dieses Kriterium etwas aufweichen, indem etwa statt 0% als Obergrenze 3% oder 5% festgesetzt werden. Dadurch erhöht sich der Prozentsatz der Trittbrettfahrer noch einmal deutlich (+13% bzw. +20%, siehe Tabelle 14). Da in dieser Arbeit neben diesen Definitionen auch der Spielertyp R1 (Investitionen $\leq 10\%$) zur Verfügung steht, ist auch damit ein Vergleich möglich: R1 stellt mit 39% über alle Länder den größten Spielertyp – so finden sich etwa alle Trittbrettfahrer, die versuchsweise winzige Anteile ihrer Ressourcen in das Gemeingut investiert haben, in dieser Kategorie (eben die genannten 13% für $\leq 3\%$). Wenn man also Trittbrettfahrer als diejenigen charakterisiert, die weniger als 5% investieren, so müsste man bei Ikariam fast die Hälfte der Gesamtpopulation (45%) so verstehen, in den fünf untersuchten Labor-PGG dagegen nur 2,9% – 19,3% (Mittelwert: 8,7%).

Im Folgenden interessiert uns der Erfolg der verschiedenen mehr oder weniger kooperativ

ausgerichteten Strategien. Die untenstehende Abbildung (Abbildung 44) fasst das Abschneiden der einzelnen Spielertypen nach Ländern getrennt zusammen. Wie in Teil C 2 beschrieben, wird die Anfangsphase (Snapshot 1) nicht aufgenommen, da es sich um eine in vielerlei Hinsicht anomale Phase handelt.

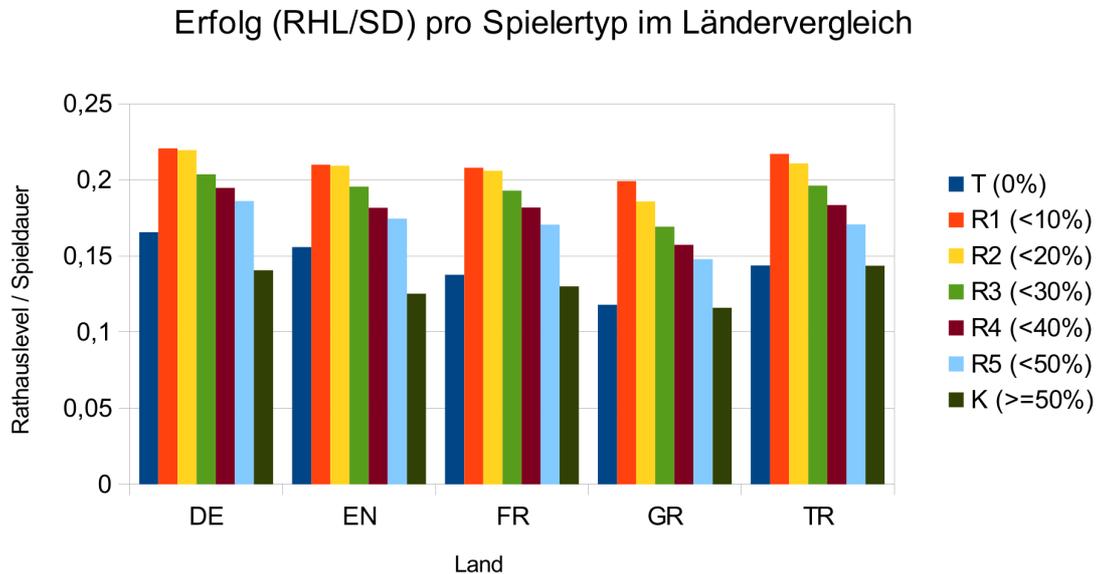


Abbildung 44. Erfolgsmaß (mittlerer Rathauslevel pro Spieldauer) im Ländervergleich für sieben Spielertypen ohne den ersten Snapshot in Ikariam

Es werden nur Inseln berücksichtigt, auf denen mindestens zwei Spieler spielen, da dies nach der Definition für PGG die Mindestanzahl darstellt. Die Abbildung zeigt einen robusten Zusammenhang zwischen Erfolg und Einzahlungen für alle nicht-extremen Strategien – fast identische Abstufungen über alle Länder hinweg jeweils von Spielern, die weniger als 10% geben, bis hin zu denjenigen, die zwischen 40 und 50% geben. Sowohl zu Trittbrettfahrern als auch Kooperierenden bestehen größere Unterschiede. Beide Strategien fallen klar im Erfolg ab, wobei keine der beiden als erfolgreicher als die andere gelten kann, wie die folgenden statistischen Analysen zeigen. Diese Analysen können durch das zweite verwendete Erfolgsmaß, nämlich Rathauslevel durch Spieldauer (RHL/SD) *korrigiert um die Zahl der Städte* bestätigt werden (Analyse hier nicht gezeigt) – hier ergibt sich ein fast identisches Bild.

Da der langfristige Erfolg interessiert, werden die statistischen Tests für dieses Erfolgsmaß für den letzten Snapshot (nach zehn Monaten) durchgeführt. Diese Daten bestätigen obige Analyse, da in allen Ländern die mittleren Ränge von R1 nach R5 laufen, gefolgt von T und K (in DE und FR liegen Kooperierende vor Trittbrettfahrern). Diese Unterschiede sind signifikant (Kruskal-Wallis-Test, $df = 6$, DE: $n = 5724$, EN: $n = 4202$, FR: $n = 6409$, GR: $n = 7544$,

TR: $n = 10572$, $p < 0,001$). Von den 105 paarweisen Vergleichen erweisen sich in den korrigierten Post-hoc-Tests alle mindestens auf dem 0,009 (DE), 0,02 (EN), 0,03(FR), 0,05 (GR) bzw. 0,007 (TR)-Niveau als signifikant, bis auf die in Tabelle 20 mit dem jeweiligen Länderkürzel gekennzeichneten nicht signifikanten Vergleiche:

	T	R1	R2	R3	R4	R5	K
T						TR	DE/EN/FR/GR/TR
R1			DE/EN				
R2				EN			
R3							
R4				DE/EN/FR/GR/TR			
R5				EN	DE/EN/FR/GR/TR		
K						EN/FR	

Tabelle 20: Paarweise Vergleiche der Post-hoc-Tests für Unterschiede im Erfolgsmaß zwischen Spielertypen in Ikariam

Die Unterschiede zwischen Kooperierenden und Trittbrettfahrern sind in England ($n = 1012$), Frankreich ($n = 1264$) und Griechenland ($n = 1770$) signifikant (Mann-Whitney-U-Test, $p < 0,001$) mit höheren mittleren Rängen für Kooperierende, in Deutschland ($n = 1193$) und der Türkei ($n = 2807$) jedoch nicht. Dieses Ergebnis sieht keine größeren Unterschiede zwischen diesen beiden Strategien. Selbst R5, die einer hoch kooperativen Strategie nahe steht, und die schlechteste der „gemäßigten“ Strategien ist, unterscheidet sich noch signifikant sowohl von Kooperierenden als auch von Trittbrettfahrern (Mann-Whitney-U-Test, $p < 0,001$). Diese Resultate sind statistisch robust, und bleiben unverändert, auch wenn verschiedene Parameter verändert werden. So bleibt die Reihenfolge im Erfolg stabil, auch wenn a) der erste Snapshot hinzugenommen wird, wenn b) nicht auf die Anzahl der Städte korrigiert wird und wenn c) nur die zweiten fünf Monate getestet werden.

3.4 Diskussion H1 Ikariam

Bevor der Erfolg diskutiert wird, wird die Frage nach den auftretenden Häufigkeiten der Spielertypen noch einmal aufgenommen, da sie in einer evolutionären Logik ein dynamisches Gleichgewicht darstellen, wobei vermehrtes Auftreten auch für evolutionären Erfolg spricht. Hierzu analysiere ich die Ergebnisse sowohl der PGG- als auch der Ikariam-Daten im Vergleich.

Auf den ersten Blick weichen die Ergebnisse aus den PGG im Labor und den PGG in Ikariam voneinander ab, da im ersteren die Häufigkeiten für Kooperierende zwischen 13% und 60%

(Mittelwert der fünf Studien: 46,2%) liegen, bei letzterem zwischen 3,1% und 6,7% (Mittelwert: 4,6%). Der Grund liegt in der *Definition*, was als kooperierender Spielertyp gilt: hier wird – um beide Datensätze vergleichen zu können – als gemeinsame Definition für Kooperierende „Investments \geq 50% des Einsatzes/der Ressourcen“ gewählt. In der Literatur zu PGG im Labor wird dagegen meist „Investments = 100% des Einsatzes“ verwendet. Wird diese Definition verwendet, dann lauten die Häufigkeiten für den Labor-PGG-Spielertyp Kooperierender wie folgt:

Studie	Häufigkeit in %
Fehr & Gächter 2000	2,7
Fehr & Gächter 2002	7,2
Gächter et al. 2008 (50 Runden)	4,9
Gächter et al. 2008 (10 Runden)	2,9
Nikiforakis 2008	0
Nikiforakis & Normann 2008	7,5

Tabelle 21. Häufigkeit der Spielertypen bei geänderter Definition (siehe Text)

Vergleicht man die in Tabelle 21 verwendete Definition mit den in der Literatur berichteten Häufigkeiten, so stellt sich eine Übereinstimmung in der Größenordnung ein: Rustagi et al. (2010) finden 2% Kooperierende, Kurzban & Houser (2005) 13% und Keser & van Winden (2000) 3% im Partner-Treatment sowie 0% im Stranger-Treatment.

Weitere Labor-PGG-Studien, die sich explizit mit Spielertypen beschäftigen, messen eine höhere Häufigkeit von Trittbrettfahrern. Dies gilt selbst für Studien mit dem strikteren Kriterium des 0%-Investments: Kocher et al. (2008) finden größere Unterschiede zwischen Ländern – USA mit 8%, Österreich mit 22% und Japan mit 36%, Rustagi et al. (2010) berichten 10% für Äthiopien, Fischbacher et al. (2001) 30% für die Schweiz und Burlando & Guala (2005) 32% für Italien. Für einen Überblick siehe Tabelle 14, die einen Mittelwert von 25% für Trittbrettfahrer und 4,6% für Kooperierende errechnet. Dies zeigt, dass es stark von der Definition abhängt, was als Trittbrettfahrer gilt, sich aber der gefundene Wert von 25% für Trittbrettfahrer in Ikariam gut in die Literatur einordnet, wie auch die 5% für Kooperierende.

Der vergleichsweise geringe Anteil (0 – 8,9%) an Trittbrettfahrern in den fünf PGG-Studien könnte daran liegen, dass es sich in vier von sieben gespielten Treatments um Partner-Treatments handelt. Damit wurde ein Kontext gewählt, in dem Kooperation nachweislich höher ist

und demzufolge Trittbrettfahrer seltener auftreten sollten. So treten beispielsweise bei der Studie von Gächter et al. (2008) im Partner-Design überhaupt keine Trittbrettfahrer auf.

Zum Erfolg der Spielertypen: Überraschend dabei ist das schlechte Abschneiden der Trittbrettfahrer, die sowohl aus theoretischen Überlegungen als auch durch die absteigende stabile Reihung der Strategien R1 bis R5 als Gewinner vermutet werden können. Es bieten sich verschiedene Erklärungsmöglichkeiten an. Erstens, Trittbrettfahrerverhalten wird von den anderen Teilnehmern nicht hingenommen und es wird versucht, sie zu zwingen, sich an den Investitionen in das öffentliche Gut zu beteiligen. Das führt zu drei Möglichkeiten: entweder sie wandeln sich zu Beitragenden, oder sie verlassen solche Inseln. Oder das Gegenteil geschieht – die kooperativeren Teilnehmer verlassen Inseln mit überdurchschnittlich vielen Trittbrettfahrern. Solche Selbstsortierungsprozesse werden in der PGG-Literatur als *endogenous assortment* oder *voluntary association* bezeichnet und zeichnen sich durch hohe Unterschiede im Profit zwischen den – in Hinsicht auf Spielertypen – relativ homogenen Gruppen aus (Page et al. 2005; Burlando & Guala 2005). Zweitens, die Effizienzgewinne der öffentlichen Güter, die ja von Trittbrettfahrern nicht oder nur in geringem Maße erzeugt werden, führen zu besseren Ergebnissen der kooperativeren Teilnehmer. Diese Vermutungen können nur durch differenziertere Analysen gestützt oder verworfen werden (siehe dazu auch Abschnitt D 3.7).

Während also in Ikariam die beiden Extremstrategien die beiden unerfolgreichsten sind, stellen sie in den Labor-PGG die erfolgreichsten dar. Im Folgenden wird dargelegt, woran das liegt. Aufgrund der absteigenden Reihenfolge von egoistischen Strategien hin zu kooperativen scheint das Stranger-NP Treatment mit den Ergebnissen von Ikariam am ehesten vergleichbar zu sein – andere Spieler in Ikariam bleiben relativ anonym, Kommunikation erfolgt praktisch nicht und Bestrafung ist nur indirekt möglich. Damit bestätigen sich zumindest diese Ergebnisse wechselseitig. Dies gilt nur für „Umgebungen“, die Kooperierenden nicht entgegenkommen (SNP). Wie man bei Hypothese 2 zeigen kann (nächster Abschnitt), sind K-Spieler auch in Ikariam unter bestimmten Umständen sehr erfolgreich. Warum Trittbrettfahrer in Labor-PGG vergleichsweise erfolgreicher sind, könnte etwa der abstrakten Einkleidung, der minimalen Handlungsmöglichkeiten, der künstlichen Situation oder der Kürze des Spiels geschuldet sein.

3.5 Resultate PGG Hypothese 2: Wenn Kooperierende unter sich sind, so sind sie Trittbrettfahrergruppen überlegen

Die zweite Hypothese – Kooperierende sollten erfolgreicher als Trittbrettfahrer sein, wenn sie

unter sich sind – folgt der grundsätzlichen Idee, dass sich Kooperierende, gemäß der Logik sozialer Dilemmata, näher am sozialen Optimum befinden als Trittbrettfahrer. Wenn sie gleichzeitig nicht von Trittbrettfahrern ausgebeutet werden, folgt daraus, dass sie die höchsten Effizienzgewinne durch das Zustandekommen des öffentlichen Gutes einfahren sollten. Im Folgenden wird geprüft, ob das der Fall ist. Für diese Analysen müssen sie in die unterschiedlichen Gruppen aufgetrennt werden, um sie hinsichtlich ihres Erfolges unterscheiden zu können. Folgende Gruppen sind logisch möglich (siehe Teil C, Tabelle 4):

- reine Gruppen aus Kooperierenden (nur K)
- nur Kooperierende und R1-R5 (KR)
- nur R1-R5 (R)
- Kooperierende, R1-R5 und Trittbrettfahrer (KRT)
- nur Kooperierende und Trittbrettfahrer (KT)
- nur R1-R5 und Trittbrettfahrer (RT)
- nur Trittbrettfahrer (T)

Die obenstehende Liste der Gruppentypen ist bereits nach dem erwarteten Erfolg sortiert. So sollten beispielsweise reine Gruppen aus Kooperierenden (K) näher am sozialen Optimum spielen als Gruppen, die aus Kooperierenden *und* R1-R5 bestehen. Zwischen anderen Gruppentypen ist die Reihung allerdings nicht unmittelbar klar – so könnten reine R1-R5-Gruppen (R) durchaus schlechter dastehen als heterogene Gruppen, die aus allen Spielertypen bestehen (KRT). Gleichmaßen unklar ist, ob reine Rezipkatoren-Gruppen (R) in ihrem Erfolg besser als Gruppen sind, die sowohl aus Kooperierenden und Trittbrettfahrern zusammengesetzt sind (KT).

Eine Prüfung auf diese Gruppentypen bei den analysierten Labor-PGG-Studien trifft auf das Problem, dass diese Gruppen zufällig entstehen und nicht durch interne (*endogenous group formation*) oder externe Mechanismen (*exogenous group formation*) gesteuert sind. Das macht die Daten auf der einen Seite authentischer, da Gruppenzusammensetzung ein reines Nebenprodukt der durchgeführten Experimente zu Bestrafung ist, auf der anderen Seite ist diese Arbeit auf die zufällig entstandenen Gruppen angewiesen, die nicht in allen Studien in gleichbleibender Häufigkeit vorkommen. Trotzdem ist bereits die Häufigkeit des Auftretens von

Gruppentypen je nach Treatment ein erster Hinweis auf die Dynamik, also die Wechselwirkungen und Reaktionen von Spielertypen auf jeweils andere Typen. Ich berichte deshalb zunächst die Häufigkeit ihres Auftretens über alle Treatments (Abbildung 45). Sowohl Gruppen, die nur aus Trittbrettfahrern (T, 0%) als auch Gruppen aus Kooperierenden und Trittbrettfahrern (KT, 0,2%) kommen nicht bzw. praktisch nicht vor. Überhaupt sind die Gruppen – wohl durch den hohen Anteil aus Kooperierenden in vier von fünf Studien (siehe Abschnitt C 1) bedingt – stark hin zu eher kooperierenden Gruppentypen verschoben.

Häufigkeit Gruppentypen über alle Labor-PGG

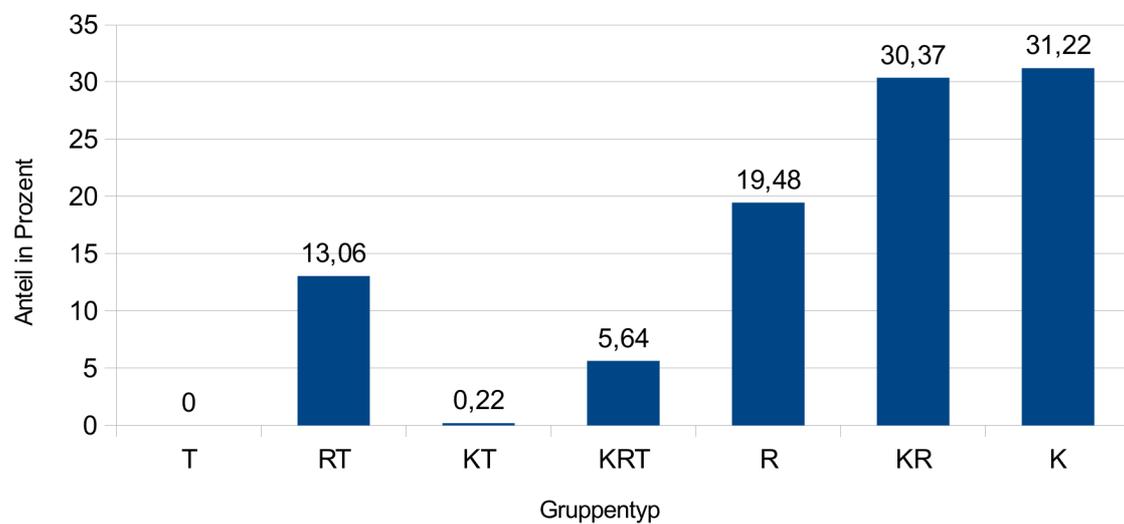


Abbildung 45. Häufigkeit Gruppentypen aggregiert über alle Labor-PGG

Gruppen, in denen alle Typen vorkommen (KRT) kommen relativ selten vor (6%), während KR mit 30% Anteil häufig auftreten. Reine Gruppen aus Kooperierenden (31%) sind häufiger als reine R1-R5-Gruppen (19%), obwohl letztere ja ein viel breiteres Spektrum an Strategien umfassen.

Da diese Aggregation allerdings wichtige Informationen verbergen könnte, schlüsselt die untenstehende Abbildung (Abbildung 46) die Gruppentypen wiederum nach Treatments auf.

Häufigkeit Gruppentypen nach Treatment

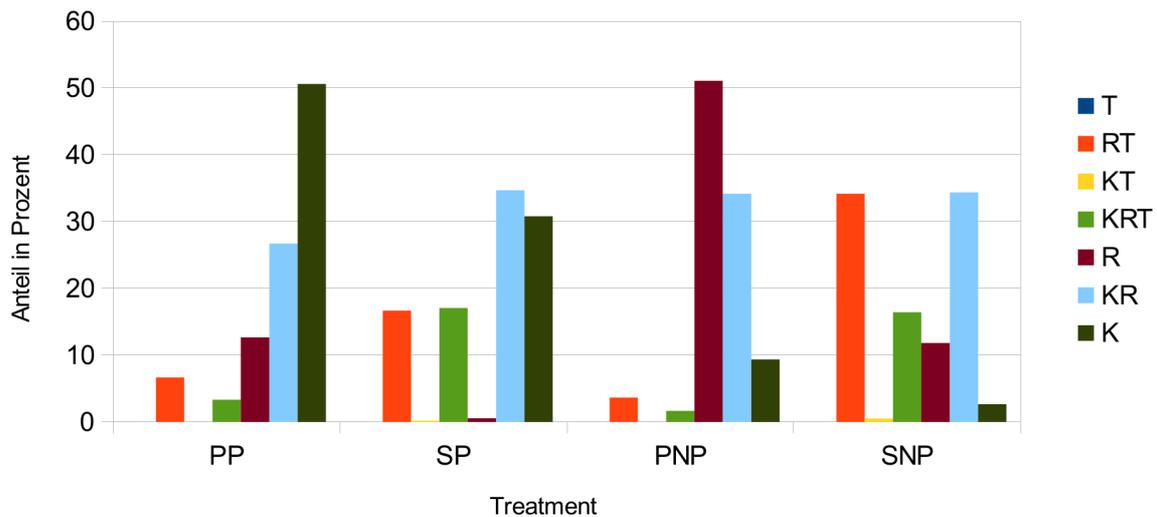


Abbildung 46. Häufigkeit der Gruppentypen nach Treatment in % über alle Labor-PGG aggregiert nach Treatment

Auf Basis dieser Verteilung der Gruppentypen lassen sich weitere Beobachtungen treffen. Erstens, Kooperierenden-Gruppen (K) treten deutlich häufiger in Partner-Treatments auf, wobei die Bestrafungsoption die Häufigkeit von 2,7% (SNP) auf 50,7% (PP) dramatisch ansteigen lässt. Zweitens, KR-Gruppen sind in allen Treatments ungefähr gleich oft zu finden. Drittens, KRT-Gruppen kommen in beiden Partner-Treatments kaum vor (1,7% bzw. 3,3%), in Stranger-Treatments jedoch schon (16% in SNP bzw. 17% in SP). Viertens, RT-Gruppen sind in Stranger-Treatments häufiger als in Partner-Treatments zu finden.

Nachdem die Häufigkeiten behandelt worden sind, interessiert uns im Folgenden primär der Erfolg dieser Gruppentypen, der über die tatsächlichen Auszahlungen an die Versuchspersonen gemessen wird (siehe Abschnitt D 3.1). Dieser Gesamtverdienst – gemessen in Prozent des maximal möglichen Verdienstes – verteilt sich pro Gruppentyp wie folgt:

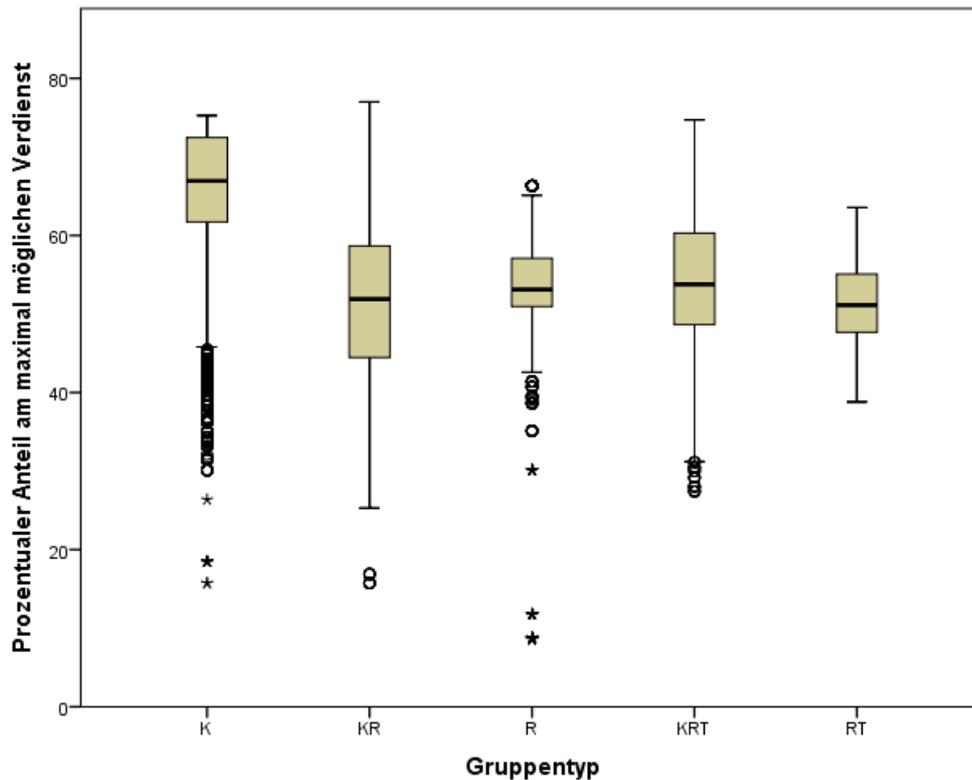


Abbildung 47: Erreichte prozentualen Anteile des maximalen Verdienst über alle Runden und vier PGG-Studien pro Gruppentyp

Abbildung 47 belegt, dass kooperierende Gruppen (K) über alle Studien hinweg mit 64% den höchsten durchschnittlichen Wert bezüglich des maximal möglichen Verdienstes erreichen. KT und T treten als Gruppentypen nicht auf. Alle übrigen Gruppen liegen relativ nahe zusammen, die mittlere Standardabweichung aller Gruppen liegt bei 8,43%. Diese Unterschiede sind nichtsdestotrotz signifikant (Kruskal-Wallis-Test, $df = 4$, $n = 12182$, $p < 0,001$). Die mittleren Ränge laufen absteigend von K und KRT über R und KR bis zu RT. In den korrigierten Post-hoc-Tests erweisen sich alle paarweisen Vergleiche mindestens auf dem 0,001-Niveau als signifikant, bis auf R-KRT und RT-KR.

Teilt man wiederum nach Treatment ein, so ergeben sich untenstehende Ergebnisse (Abbildung 48). Zunächst fällt auf, dass Partner-Treatments (PNP und PP) sehr wirksam alle nicht-kooperativen Gruppenzusammensetzungen außer K und KR, R und KRT verhindern. In diesen Treatments kehrt sich der Verdienst dieser beiden Gruppen je nach Bestrafungsoption um, wobei Bestrafung den Gruppen zu Gute kommt, die ausschließlich aus Kooperierenden bestehen. Diese (K) erreichen in konstanten Gruppen mit Bestrafung (PP) 63% des maximalen Verdienstes. K-Gruppen erzielen auch den höchsten Wert im SP-Treatment mit 47%, was

einen Effizienzverlust von 16% darstellt! Im Stranger-Treatment in wechselnden Gruppen (SNP) erzielen gemischte Gruppen (KRT) mit 55% die höchsten Gewinne.

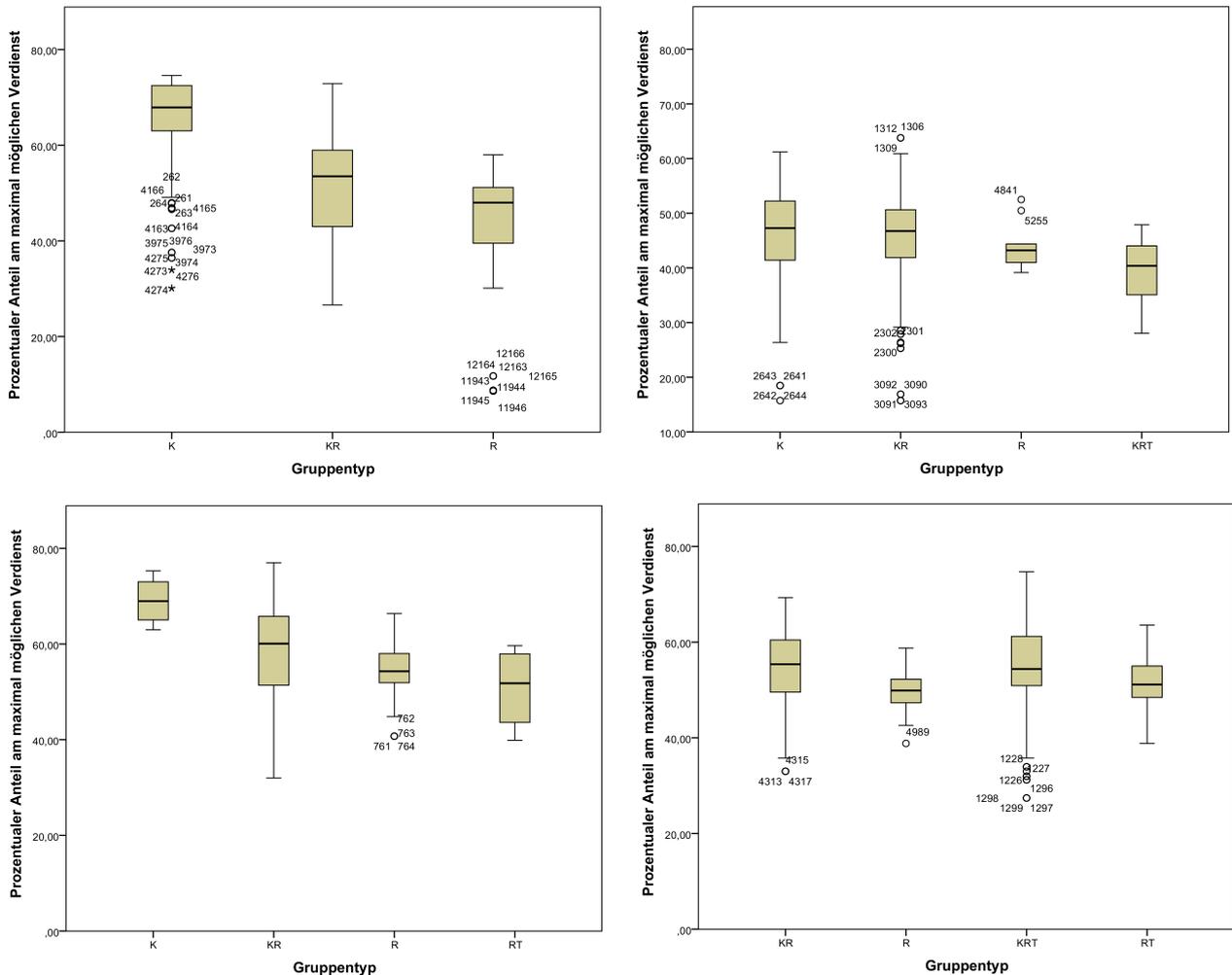


Abbildung 48: Durchschnittlicher Verdienst pro Gruppentyp über vier PGG-Studien nach Treatment (im Uhrzeigersinn: Partner-P; Stranger-P; Stranger-NP; Partner-NP)

Diese Unterschiede sind jeweils signifikant (Kruskal-Wallis-Test, $p < 0,001$):

- PP: $df = 2$, $n = 4360$; absteigende mittlere Ränge: K, KR, R.
- SP: $df = 3$, $n = 2016$; absteigende mittlere Ränge: K, KR, R, KRT
- PNP: $df = 3$, $n = 3790$; absteigende mittlere Ränge: K, KR, R, RT
- SNP: $df = 3$, $n = 2016$; absteigende mittlere Ränge: KRT, KR, RT, R

In den korrigierten Post-hoc-Tests erweisen sich alle paarweisen Vergleiche mindestens auf dem 0,001-Niveau (PP, SNP) bzw. dem 0,002-Niveau (SP, PNP) als signifikant, bis auf R-RT in SNP und K-KR in SP.

Selbst ohne Bestrafungsmöglichkeit bieten konstante Gruppen Kooperierenden immer noch gute Möglichkeiten zu hohen Verdiensten – hier beträgt der Median des maximal erreichbaren Verdienstes 68% (Mittelwert 69%) des Maximums, während die Mediane und Mittelwerte aller anderen Gruppen im Vergleich deutlich niedriger liegen.

3.6 Diskussion H2 PGG

Es zeigt sich demnach in Übereinstimmung mit der Literatur (Fehr & Gächter 2002), dass eine stabile Gruppenzusammensetzung (Partner-Treatment) und die Möglichkeit, Trittbrettfahrer zu strafen (Bestrafungs-Treatment) das für Kooperation geeignetste Milieu ist, da sich hier häufiger Gruppen finden, die zusammenarbeiten (K und KR).

Da viele Gruppen überhaupt nicht vorkommen, können allerdings nur wenige Folgerungen aus den Daten gezogen werden. Es scheint aber der Fall zu sein, dass Gruppen aus Kooperierenden im Vergleich zu anderen Gruppen in einigen Treatments gut abschneiden. Gerade auf Grund der fehlenden Daten und der Artifizialität der Laborumgebung, die echte Gruppenbildung nur sehr eingeschränkt, wenn überhaupt, zulässt, sind eher die Daten aus Ikariam ein echter Test für diese Hypothese.

3.7 Resultate Ikariam Hypothese 2: Wenn Kooperierende unter sich sind, so sind sie Trittbrettfahrergruppen überlegen

Die Natur der Daten ist besonders geeignet, diese Frage (nach der Überlegenheit Kooperierender, so sie unter sich sind) zu beantworten, da die Welt in Inseln geteilt ist; diese stellen die Gruppen dar, die miteinander interagieren. Eine erste Annäherung an die Frage ist daher – wie bei den Labor-PGG auch – die Häufigkeit, in der Inseltypen (K, KR, R, KRT, KT, RT, T) existieren, auch weil sich evolutionärer Erfolg oft in der Häufigkeit des Auftretens widerspiegelt. Es werden die Daten aller Snapshots bis auf den ersten berücksichtigt. Ebenfalls nicht berücksichtigt sind alle Inseln, auf denen niemand oder nur ein Spieler siedelt (siehe Tabelle 12 für daraus resultierende Unterschiede in den Häufigkeiten).

Die Abbildung 49 verdeutlicht, dass R1-R5 gemischt mit Trittbrettfahrern (RT) die mit Abstand häufigste Gruppenzusammensetzung sind (53%-68%). Dahinter kommen Inseln mit allen Spielertypen (KRT, 16%-29%) und reine Reziprokatoren-Inseln (R, 9%-17%). Sehr selten kommen Inseln vor, die nur aus Kooperierenden bestehen (K, 0,01%-0,2%), während Kooperierende und Reziprokatoren (KR) immerhin zwischen 3% und 7% ausmachen. Reine Trittbrettfahrerinseln (T) und Inseln, auf denen nur Kooperierende und Trittbrettfahrer vorkommen (KT) kommen überhaupt nicht vor.

Häufigkeit der Gruppentypen

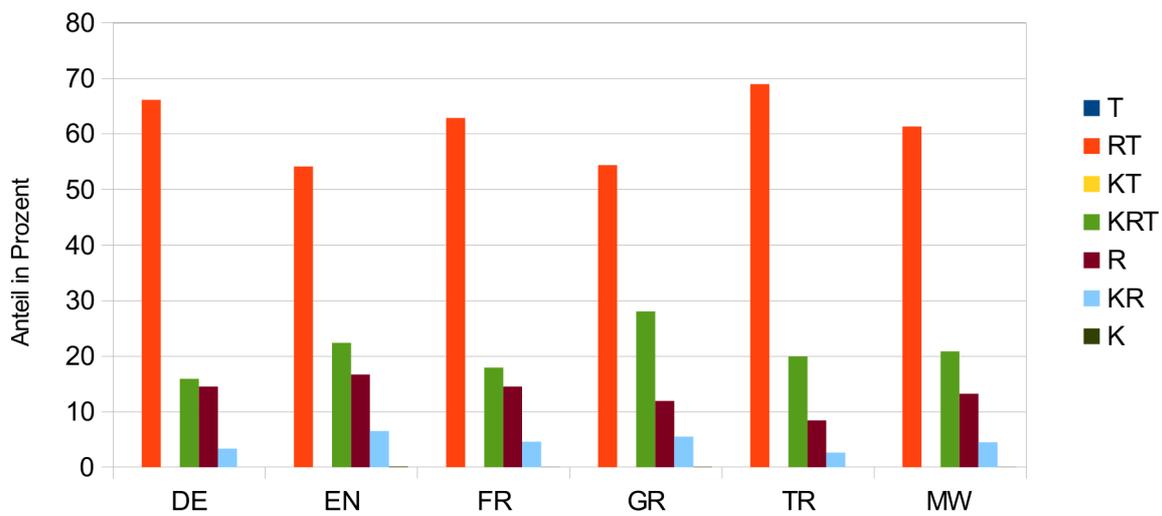


Abbildung 49. Häufigkeit der Gruppentypen in Ikariam im Ländervergleich ohne den ersten Snapshot

Ein weiterer Aspekt ist die Entwicklung der Gruppenzusammensetzung über die Zeit. Eine Frage, die mit den vorliegenden Daten beantwortet werden kann, ist, ob es ohne Anreize von außen zu einer Gruppenbildung kommt (*endogeneous group formation*, vgl. z. B. Page et al. 2005). Die untenstehende Abbildung (Abbildung 50) belegt, dass dies zumindest in Ansätzen der Fall ist.

Mittlerer Anteil der Gruppentypen über alle Länder

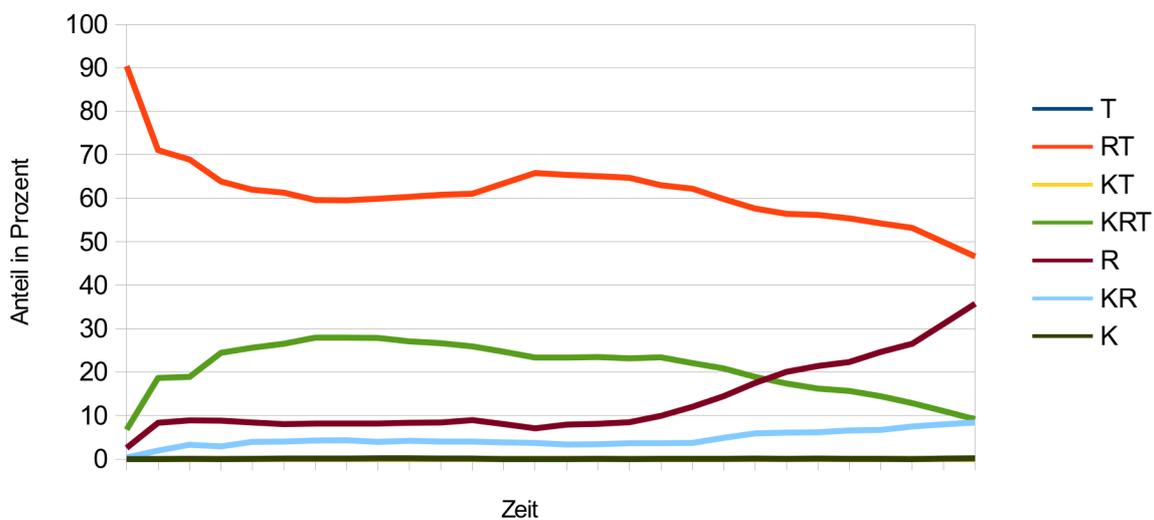


Abbildung 50: Durchschnittliche Häufigkeit der Gruppentypen im Zeitverlauf für alle Länder (Mittelwerte)

Der Anstieg reiner R1-R5-Gruppen bei gleichzeitiger Abnahme von RT und KRT-Gruppen zeigt, dass die Homogenität ansteigt. Ob dies durch den Ausschluss von Trittbrettfahrern oder die „Umwandlung“ von Trittbrettfahrern in R1-R5-Typen geschieht – sei es durch Anpassung oder sei es durch Zwang – kann hier nicht weiter verfolgt werden. Da die Spielertypen allerdings sehr konstant sind, ist eine Selbstsortierung am wahrscheinlichsten. Ich wende mich wiederum dem Erfolg zu, der durch die abhängige Variable Rathauslevel pro Spieldauer gemessen und für alle Inseln mit mehr als einer Stadt für alle Snapshots außer dem ersten erhoben wird (Abbildung 51). Zur Erinnerung: Die Gruppenzusammensetzung RT ist mit Abstand die häufigste (53-68%). Vom Erfolg her ist sie im Mittelfeld zu verorten. Bis auf Frankreich zeigen sich KR und R-Inseln als Gewinner, während Inseln, die nur aus Kooperierenden bestehen (K) abgeschlagen sind.

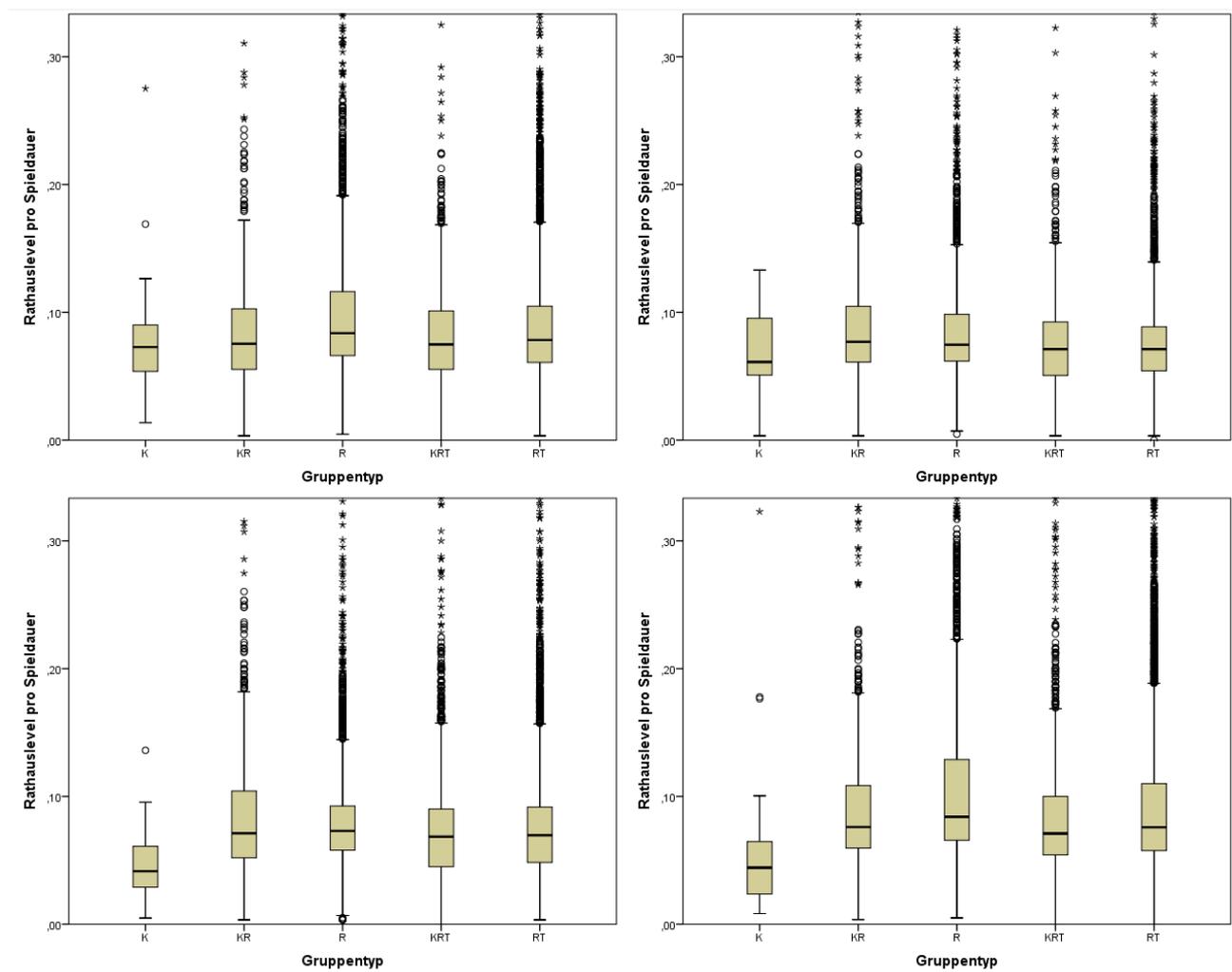


Abbildung 51: Durchschnittlicher Erfolg (Rathauslevel / Spieldauer) nach Gruppentyp im Ländervergleich für den letzten Snapshot (im Uhrzeigersinn: DE, EN, TR, GR)

Die statistische Prüfung auf Unterschiede ergibt zwischen den Gruppentypen ein hochsignifikantes Ergebnis (Kruskal-Wallis-Test, $df = 4$, $n = 34451$, $p < 0,001$). Bei den paarweisen Vergleichen (nach Anzahl der Tests korrigiert) fallen alle Vergleiche zwischen den Gruppentypen signifikant aus (bis auf RT und KR), wobei sich für den mittleren Rang die Reihenfolge R, KR, RT, KRT und K ergibt. Insgesamt machen also die Reziprokatoren das Rennen, die Extremstrategie K ist weit abgeschlagen, die Gruppentypen T und KT kommen nicht vor.

Um zu berücksichtigen, dass kleinere Gruppen Gemeingüter schwerer erzeugen können als große, kann man um die Anzahl der Städte korrigieren. Teilt man also den Erfolg pro Spieldauer um die durchschnittliche Anzahl der Städte pro Inseltyp und Snapshot, dann ergibt sich ein anderes Bild (Abbildung 52). Das liegt vor allem daran, dass Kooperierende (K), Trittbrettfahrer (T) und Kooperierende und Trittbrettfahrer (KT) mit deutlich kleineren Gruppen zu kämpfen haben (siehe Tabelle 11 in Abschnitt B 1).

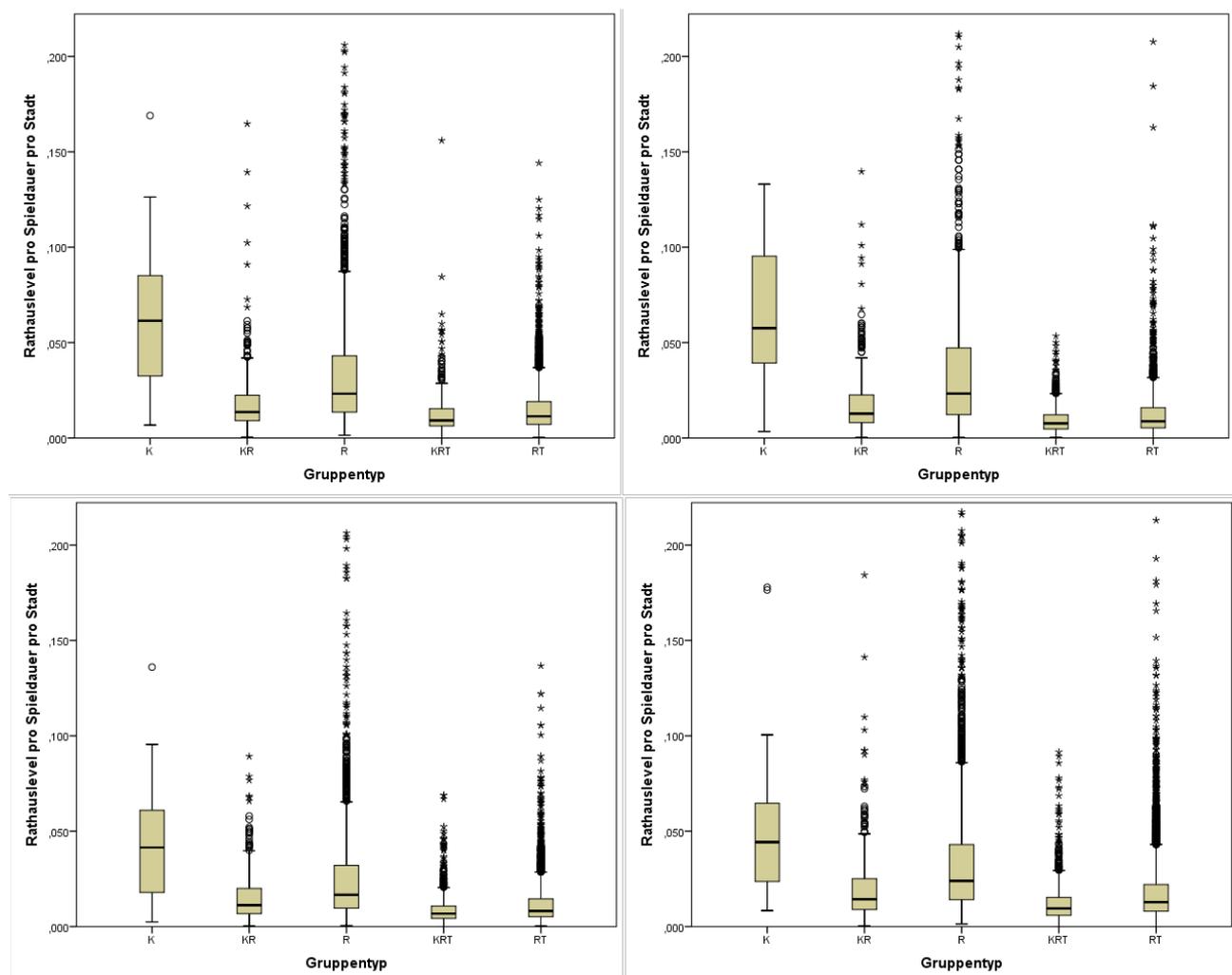


Abbildung 52: Durchschnittlicher Erfolg (Rathauslevel / Spieldauer) pro Anzahl der Städte nach Gruppentyp im Ländervergleich (im Uhrzeigersinn: DE, EN, TR, GR)

Die Werte, auf denen diese Abbildung basiert, sind in der folgenden Tabelle 22 zu finden. Kooperierende Gruppen verzeichnen mit Abstand den höchsten Erfolg (0,0582) vor Gruppen, die nur aus Reziprokatoren bestehen (0,0331). Dies ist durchgängig für alle Länder der Fall.

RHL / SD /Anzahl Städte	K	KR	R	KRT	RT
DE	0,0696	0,0187	0,0347	0,0131	0,0153
EN	0,0633	0,0180	0,0365	0,0100	0,0134
FR	0,0488	0,0178	0,0318	0,0110	0,0146
GR	0,0444	0,0149	0,0268	0,0085	0,0116
TR	0,0650	0,0200	0,0355	0,0128	0,0178
MW	0,0582	0,0179	0,0331	0,0111	0,0145

Tabelle 22. Durchschnittlicher Erfolg (Rathauslevel / Spieldauer) pro Anzahl der Städte nach Inseltyp im Ländervergleich nach zehn Monaten

Auch hier ergibt die Prüfung auf Unterschiede für den letzten Snapshot ein hochsignifikantes Ergebnis (Kruskal-Wallis-Test, $df = 4$, $n = 34451$, $p < 0,001$):

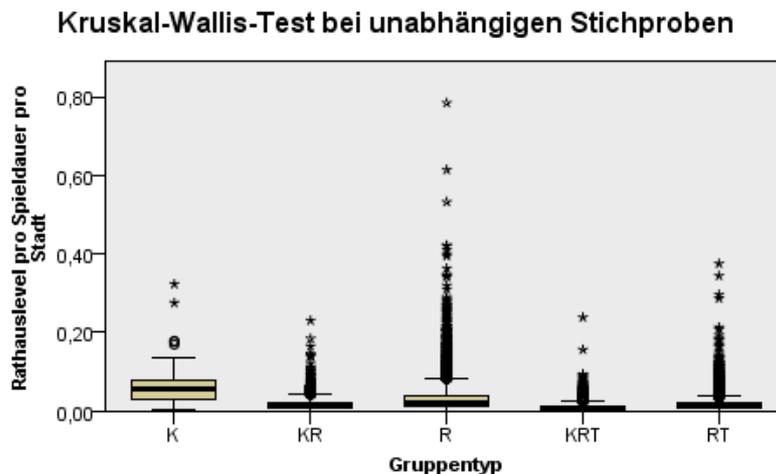


Abbildung 53: Test auf Unterschiede im Erfolg (Rathauslevel / Spieldauer) pro Anzahl der Städte nach Gruppentyp für alle Länder am 13.2

Für die mittleren Ränge ergibt sich die Reihenfolge Kooperierende (K) vor Reziprokatoren (R) gefolgt von Kooperierenden und Reziprokatoren (KR) sowie Reziprokatoren und Trittbrettfahrern (RT) und schließlich von allen Spielertypen (KRT). Alle paarweisen Vergleiche sind mindestens auf dem 0,001-Niveau signifikant (nach Anzahl der Tests korrigiert). Es macht demnach einen erheblichen Unterschied für den Erfolg, mit welcher Gruppenzusammensetzung

zung man arbeiten muss. Die durchschnittliche Anzahl der Städte ist dabei mit dem Erfolg (RHL pro Spieldauer) zwar signifikant, aber ohne ausgeprägten Korrelationskoeffizienten korreliert (Pearson $r = -0,04$, $p < 0,001$).

3.8 Diskussion H2 Ikariam

In den *Häufigkeiten* werden deutliche Unterschiede zu den Labor-PGG sichtbar. Dort lautete die Reihenfolge der Häufigkeit K-KR-R-RT-KRT-KT (siehe Abbildung 45), wobei RT mit 14% nur einen kleineren Anteil ausmachen. Bei Ikariam sind RT mit Abstand am häufigsten, gefolgt von R-KRT-KR und K. Tatsächlich scheinen die Zahlen von Ikariam realistischer zu sein, wenn man die in der Literatur berichteten Anteile von Reziprokatoren (ca. 50%), Trittbrettfahrern (ca. 20%) und Kooperierende (ca. 15%) in PGG heranzieht und daraus auf Gruppenzusammensetzungen schließt (vgl. Kocher et al. 2008; Rustagi et al. 2010; Fischbacher et al. 2001).

Hinsichtlich des Erfolges decken sich die Datensätze: Kooperierende können, wenn sie unter sich sind, klare Vorteile im Erfolg für sich verbuchen. Reine Gruppen aus Kooperierenden stehen sowohl in PGG als auch in Ikariam an der Spitze, gefolgt von Gruppen, die auch Reziprokatoren enthalten.

Der Ausreißer für Kooperierende bei Frankreich ist möglicherweise ein zusätzlicher Beleg für die Hypothese, dass Kooperierende unter geeigneten Bedingungen, also ohne Trittbrettfahrer, durch die Effizienzgewinne des erzeugten öffentlichen Gutes zu sehr großen Wettbewerbsvorteilen im Gruppenvergleich kommen können, da hier eine Insel mit extrem hohen Werten das Ergebnis von sehr wenigen anderen Inseln nach oben verzerrt. Bei diesen Daten ist *nicht* auf die Anzahl der Städte korrigiert.

Eine darauf korrigierte Reihenfolge ist aussagekräftiger als die unkorrigierte, da letztere vor allem die Zahl der Städte widerspiegelt (siehe Tabelle 11 in Abschnitt C). Da es natürlich billiger für den Einzelnen wird und einfacher ist, die Grenzen (*thresholds*) für die Erzeugung des öffentlichen Gutes zu erreichen, je mehr Personen sich an den Investitionen beteiligen, sind kleine Gruppengrößen von Nachteil. Anders ausgedrückt: wenn die Erzeugung eines Gemeingutes auf mehr Individuen verteilt werden kann, wird die Minimalinvestition jedes Beteiligten kleiner. Da bei Ikariam und auch in der realen Welt die absoluten Kosten für ein Projekt oft unabhängig von der Anzahl der Beteiligten sind, etwa die Kosten für ein Krankenhaus, das durch lokale Spenden finanziert werden soll, ist die Korrektur um die Anzahl der Beteiligten wichtig, um die Investitionsbereitschaft unabhängig vom Erreichen des Zieles in den Fokus zu

rücken.

Beide Betrachtungsweisen – mit und ohne Korrektur um die Inselgröße – haben allerdings ihre Berechtigung: wie evolutionär zu erwarten ist, können Kooperierende nur sehr selten Bedingungen schaffen, in denen keine Interaktionen mit Trittbrettfahrern stattfinden; analog benötigen Trittbrettfahrer ein „Trittbrett“ in Form anderer, investierender Spieler und müssen diese deshalb aktiv aufsuchen.

Da vor der Besiedlung einer Insel nur die Ausbaustufe der Ressource und die Zahl der Spieler bekannt ist, nicht aber deren strategische Ausrichtung, greifen diese Überlegungen vermutlich bei den meisten Spielern; außer örtlicher Nähe gibt es keine weiteren zu optimierenden Parameter. Ein Grund, warum auf K-Inseln kleine Gruppen zu finden sind, obwohl doch dorthin bevorzugt *hingewechselt* werden sollte, besteht in der Seltenheit von K-Spielern (ca. 5%). In der Tat hat der Inseltyp KR mit im Mittel 11 Spielern die zweithöchste Anzahl an Spielern; hier hat also möglicherweise genau dieser Hinzugseffekt stattgefunden.

Evolutionäre Strategien müssen robust gegen andere Strategien sein, d.h. in Populationen mit verschiedenen Strategien bestehen können, was sich hier statistisch belegen lässt – die drei Mischformen (KR, KRT, RT) sind häufig (siehe auch Abbildung 49) und besitzen auch die meisten Städte.

Als Fazit: Die reinen Kooperierenden-Inseln sind nach Korrektur um Inselgröße nicht mehr hinter allen anderen Inseltypen, sondern im Gegenteil im Ausbau ihrer Städte pro Zeit weit vorne. Die Hypothese, dass Kooperierende unter sich anderen Gruppen überlegen sind, kann damit als bestätigt gelten, auch wenn der Erfolg der einzelnen Gruppen nicht genau der theoretisch erwarteten Reihenfolge gehorcht (K-R-KR-RT-KRT). Eine weitere Möglichkeit besteht darin, dass in kleinen Gruppen die Notwendigkeit jedes Einzelnen zur Investition stärker ausgeprägt ist, da niemand sonst die Kosten für den Aufbau des Gemeinguts übernehmen kann. Die weiteren Analysen werden diese Befunde erneut aufnehmen und diskutieren.

3.9 Resultate PGG Hypothese 3: Die Staffelung des Gruppenerfolgs sollte K, KR, R, KRT, KT, RT, T sein

Aus theoretischer Sicht ist im Gruppenvergleich die Reihenfolge K, KR, R, KRT, KT, RT, T zu erwarten. Diese Reihenfolge ergibt sich aus der absteigenden Nähe zum sozialen Optimum. Es ist zwar nicht ausgeschlossen, dass KRT im Einzelfall erfolgreicher als R-Gruppen sein können, aber im Regelfall sollten die Trittbrettfahrer die Reziprokatoren in der Gruppe zu fallenden Investitionen in das Gemeingut veranlassen.

Die Resultate sind bereits in den obenstehenden Abbildungen 47 und 48 zusammengefasst. In der Gesamtaggregation ergibt sich die Reihenfolge K, KRT, R, KR und RT. Nach Treatment aufgeteilt (Abbildung 48) sind Kooperierenden-Gruppen (K) und KR-Gruppen stets vorne (bis auf SNP, was zu erwarten ist). Erschwerend ist hierbei, dass nicht alle Gruppentypen vorkommen. Deshalb sind die Daten für belastbare vergleichende Aussagen nur bedingt auswertbar. Es zeigen sich jedoch zumindest Tendenzen, welche die theoretisch erwartete Reihenfolge unterstützen. Im Gegensatz dazu können die Ikariam-Daten ihre Stärke – die großen Spieler- und infolgedessen auch Gruppenzahlen – wiederum voll ausspielen.

3.10 Resultate Ikariam H3: Die Staffelung des Gruppenerfolgs sollte K, KR, R, KRT, KT, RT, T sein

Der Erfolg der Gruppentypen ist in der folgenden Abbildung 54 zusammengefasst und zeigt beide Erfolgsmaße im Vergleich. Ein Test auf Unterschiede für den Erfolg nach zehn Monaten zeigt signifikante Unterschiede zwischen den Gruppentypen (Kruskal-Wallis-Test, $df = 4$, DE: $n = 5724$, EN: $n = 4202$, FR: $n = 6409$, GR: $n = 7544$, TR: $n = 10572$, $p < 0,001$). Die paarweisen korrigierten post-hoc-Vergleiche zeichnen für die Länder ein uneinheitliches Bild, da für Deutschland viele Vergleiche nicht signifikant ausfallen, für Griechenland und die Türkei dagegen praktisch alle:

	K	KR	R	KRT	RT
K		DE/EN	DE/EN	DE/EN	DE/EN
KR			GR		
R					
KRT		DE/FR/TR			DE/EN/FR/GR
RT		DE/FR/TR			

Tabelle 23: Paarweise Vergleiche der Post-hoc-Tests für Unterschiede im Erfolgsmaß zwischen Gruppen in Ikariam

Dabei ergeben die mittlere Ränge für Deutschland eine Staffelung von R, RT, KR, KRT und K, für England von KR, R, KRT, RT und K, und für Frankreich, Griechenland und die Türkei von R, KR, RT, KRT, und K. Gruppen von Kooperierenden (K) sind hier jeweils weit abgeschlagen. Reine Trittbrettfahrergruppen und KT-Inseln kommen nicht vor.

Korrigiert man um die Anzahl der Städte, so ergibt sich nach zehn Monaten eine andere Erfolgsreihenfolge, die für alle Länder gleich ist: Kooperierende (K) vor Reziprokatoren (R) und deren Kombination (KR). Schlusslichter sind hier Gruppen aus Reziprokatoren und Trittbrettfahrer (RT) und Gruppen mit allen Typen (KRT). KT und T-Gruppen kommen in keinem

Erfolg der Gruppentypen im Vergleich der zwei Erfolgsmaße

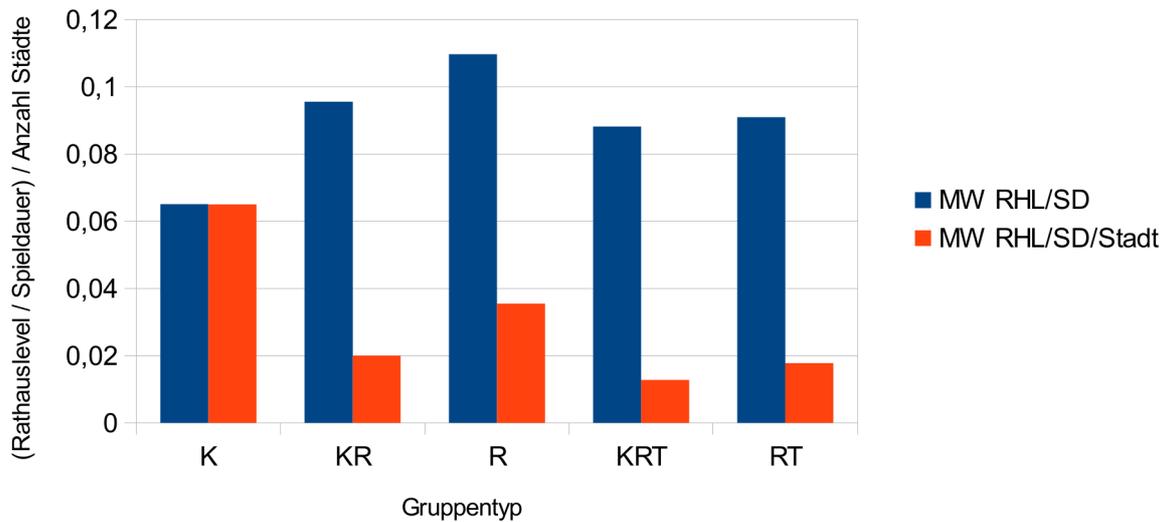


Abbildung 54. Vergleich der zwei Erfolgsmaße (RHL/SD/Stadtanzahl und RHL/SD) für den Erfolg der Gruppentypen für den letzten Snapshot

Land vor. Diese Unterschiede sind signifikant (Kruskal-Wallis-Test, $df = 4$, DE: $n = 5724$, EN: $n = 4202$, FR: $n = 6409$, GR: $n = 7544$, TR: $n = 10572$, $p < 0,001$). Im Vergleich zum Erfolgsmaß ohne Korrektur zeigen die korrigierten Post-hoc-Tests für alle paarweisen Vergleiche ein einheitliches Bild: Für Deutschland ($p < 0,001$), England ($p < 0,003$) und Frankreich ($p < 0,004$) sind sämtliche Vergleiche signifikant. Für Griechenland ($p < 0,001$) und die Türkei ($p < 0,005$) sind sie dies ebenfalls, bis auf jeweils R-K.

Diese Unterschiede sind auf die Gruppengröße zurückzuführen (durchschnittliche Anzahl der Städte: K: 2,5; RT: 11). Kleine Gruppen machen es entsprechend teurer, ein Gemeingut zu erzeugen. Das zeigt sich auch deutlich am hohen Erfolg von Kooperierenden-Gruppen nach Korrektur um die Stadtanzahl (Abbildung 54), da sie vom letzten auf den ersten Platz wechseln. Eine zusätzliche Stützung ergeben Korrelationsanalysen zwischen der Ausbaustufe des Gemeinguts und der Zahl der Städte (Tabelle 24), deren Korrelationskoeffizienten relativ hoch, und zwar zwischen 0,42 und 0,67 liegen. Viele Städte auf einer Insel korrelieren demnach mit einem hohen Ausbau der Gemeingüter:

	Ausbaustufe Sägewerk	Ausbaustufe Luxusmine	Ausbaustufe Wunder
DE-Stadtanzahl	0,65	0,67	0,57
EN-Stadtanzahl	0,53	0,57	0,56
FR-Stadtanzahl	0,62	0,65	0,57
GR-Stadtanzahl	0,54	0,56	0,53
TR-Stadtanzahl	0,54	0,60	0,42

Tabelle 24. Pearson-Korrelationskoeffizienten (einseitig) zwischen der Anzahl der Städte und den Ausbaustufen der Gemeingüter für den letzten Snapshot (13.2.2014) ohne leere Inseln; alle Zusammenhänge sind auf dem 0,01-Niveau signifikant

3.11 Diskussion H3 Ikariam

Zunächst ist festzuhalten, dass die Ergebnisse für beide Erfolgsmaße über alle Länder konsistent sind, es also nur bei Kooperierenden in Frankreich ohne Korrektur der Stadtanzahl einen Ausreißer gibt. Das Ergebnis mit Korrektur der Stadtanzahl liegt näher an den theoretischen Erwartungen. Mit Korrektur nehmen dann K-Gruppen die Spitzenposition ein, gefolgt von reinen Reziprokatoren-Gruppen (R) und KR-Inseln. Die beiden Gruppen mit Trittbrettfahrern (RT und KRT) sind am Ende zu finden. Auch wenn die erwartete Reihenfolge nicht ganz eintrifft, spiegelt zumindest das korrigierte Erfolgsmaß die theoretischen Erwartungen wider.

Von einer Bestätigung der Hypothese will ich dennoch nicht sprechen – dafür sind Ursachen und andere Bedingungen für diese Reihung zu unklar. Die Abhängigkeit des Resultats von der Anzahl der Mitspieler verdeutlicht, dass es nicht nur entscheidend ist, welche Strategien potenzielle Partner bei der Erzeugung von Gemeingütern haben, sondern auch, wie viele mobilisiert werden können. Auf den Faktor Gruppengröße als wichtiges Attribut für evolutionären Erfolg weisen auch West et al. 2011 hin.

3.12 Resultate PGG Hypothese 4: Im direkten Aufeinandertreffen von Kooperierenden und Trittbrettfahrern ist Erfolg (für T) bzw. Misserfolg (für K) am deutlichsten

Diese Hypothese greift auf, dass kooperatives Verhalten, das nicht über Verwandtschaftsbeziehungen oder Reziprozität erklärt werden kann, über seinen direkten Erfolg erklärt werden kann. Unter bestimmten Umständen – etwa einer geringen Anzahl an Trittbrettfahrern oder externen Möglichkeiten wie Reputation oder Bestrafung, Trittbrettfahrer zur Kooperation zu bewegen – sollten kooperative Strategien erfolgreicher sein, da sie effizienter sind. Diese Überlegung impliziert im Umkehrschluss, dass der Erfolg kooperativer Individuen besonders dann einbrechen sollte, wenn sie auf Trittbrettfahrer treffen.

Die folgende Abbildung 55 belegt, dass dies bei PGG tatsächlich der Fall ist. Die abhängige Variable, der Erfolg wird als Gesamtverdienst (*total profit*) gemessen, nachdem Bestrafungskosten abgezogen sind. Um die Studien vergleichbar zu machen, wird dieser Gesamtverdienst wiederum in Prozent des maximal erreichbaren Profits umgerechnet (siehe Teil C 1). Ein Vergleich dieses Erfolgs ergibt in Labor-PGG zwischen Spielertypen, die nach Gruppentypen aufgeschlüsselt sind, ein hochsignifikantes Ergebnis sowohl für Treatments mit und ohne Bestrafung (Kruskal-Wallis-Test, $df = 2$, $P: n = 5232$ für P; für NP: $df = 4$, $n = 1796$, $p < 0,001$). In den korrigierten Post-hoc-Tests erweisen sich alle paarweisen Vergleiche mindestens auf dem 0,001-Niveau als signifikant, bis auf K in KR zu T in RT im NP-Treatment.

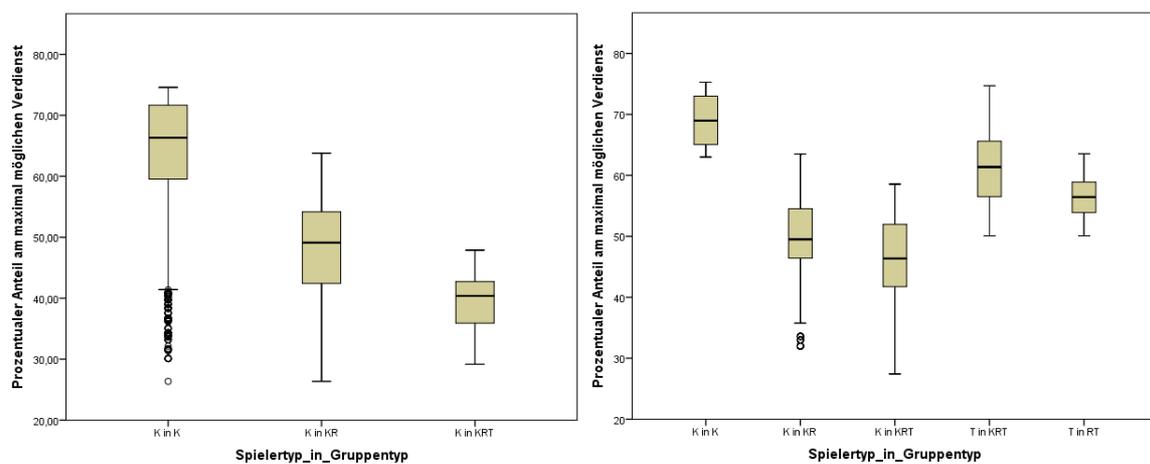


Abbildung 55: Test auf Unterschiede im Erfolg der Spielertypen in fünf Gruppensamensetzungen in vier Labor-PGG in Prozent des maximal möglichen Verdienstes (links für P, rechts für NP-Treatments)

Es zeigen sich deutliche Unterschiede, je nachdem, ob Bestrafung möglich ist oder nicht. Mit Bestrafung sind Kooperierende (K in K) mit 63% des maximal möglichen Verdienstes vor Kooperierenden in gemischten Gruppen (KRT), die 39% erreichen. Trittbrettfahrer kommen nicht vor. Ohne Bestrafung sind dagegen Trittbrettfahrer im direkten Vergleich Kooperierenden überlegen: mit 45% (K in KRT) zu 62% (T in KRT) verdienen sie 17% mehr. Gruppen, die nur aus Kooperierenden bestehen, sind in beiden Treatments die erfolgreichste Gruppe.

3.13 Diskussion H4 PGG

Damit bestätigen sich die Hypothesen vollumfänglich – in Gruppen mit Kooperierenden und Trittbrettfahrern können letztere stark profitieren. Die Unterschiede zwischen den Spielertypen in den verschiedenen sozialen Umgebungen sind beachtlich und reichen von 39% (K in

KRT mit Bestrafung) bis 62% (T in KRT ohne Bestrafung). Wenn man diese Ergebnisse mit den Häufigkeiten kombiniert, mit der diese Gruppenzusammensetzungen auftreten, ergeben sich bereits weitreichende Konsequenzen für die Fitness egoistischer und altruistischer Strategien (die in Teil E angesprochen werden).

3.14 Resultate Ikarium Hypothese 4: Im direkten Aufeinandertreffen von Kooperierenden und Trittbrettfahrern ist Erfolg (für T) bzw. Misserfolg (für K) am deutlichsten

Wie eben analysiert, ist eine wichtige erste Beobachtung, dass sich das Vorkommen von Spielertypen in Gruppentypen in der Häufigkeit auch in diesem Datensatz sehr stark unterscheidet. Die untenstehende Abbildung 56 zeigt, dass gemischte Gruppen mit Reziprokatoren (RT, KR, KRT) den Hauptanteil ausmachen.

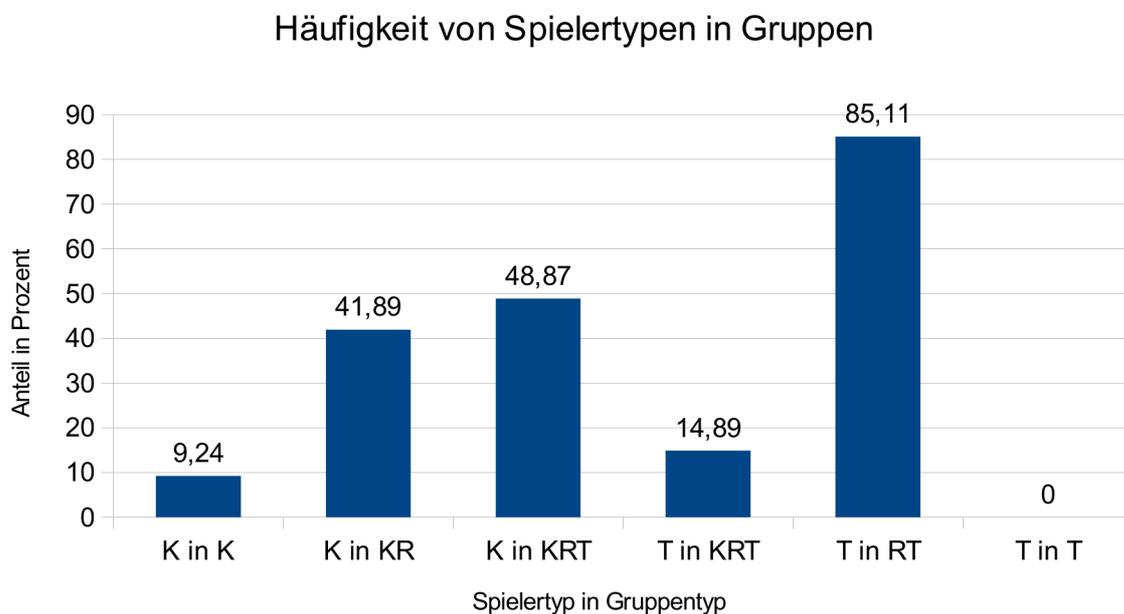


Abbildung 56. Häufigkeit von Spielertypen in Gruppen in Ikarium

Nicht-reziprokatorische Strategien (K und T) müssen sich also darauf einstellen, sehr oft in einem heterogenen Umfeld bestehen zu müssen (hier: KRT = 64% der Fälle). Ein rein kooperativ geprägtes Umfeld ist dagegen mit einer Häufigkeit von neun Prozent selten; ein Umfeld ohne Trittbrettfahrer mit 42% dagegen deutlich wahrscheinlicher. Betrachtet man die allgemeine Häufigkeit der Inseltypen (siehe Tabelle 12 in Abschnitt C 2), so ergibt sich eine Trittbrettfahrer-freie Umgebung in 29,3% der Fälle, während gemischte Umgebungen (KR und KRT) noch 21,0% ausmachen. Reine Kooperierenden-Umgebungen (1,3) sind dagegen sehr selten, KT existieren überhaupt nicht.

Der Erfolg für Ikariam wird im Folgenden wiederum als Rathauslevel pro Spieldauer pro Stadt für Inseln mit mehr als einem Spieler gemessen. Da der Gesamterfolg interessiert, wird das Ergebnis nach zehn Monaten Spielzeit verwendet, d.h. die Daten des Snapshots vom 13.2.2014.

Trittbrettfahrer und Kooperierende treffen in den Konstellationen KRT und KT aufeinander. KT-Gruppen existieren allerdings nicht. Für KRT-Gruppen ist das Ergebnis eines Tests auf Unterschiede im Erfolg zwischen den Spielertypen K und T nur für Griechenland signifikant (Mann-Whitney-U-Test, $n = 582$, $p < 0,024$; DE ($n = 187$), EN ($n = 262$), FR ($n = 249$), TR ($n = 444$) bleiben n.s.). Abbildung 57 zeigt auch die anderen Gruppen in einer Übersicht für Griechenland). Den höheren Rang im Test, also den durchschnittlich höheren Erfolg, erreichen in Deutschland, England und Frankreich allerdings Kooperierende, nicht Trittbrettfahrer. In Griechenland und in der Türkei sind Trittbrettfahrer erfolgreicher.

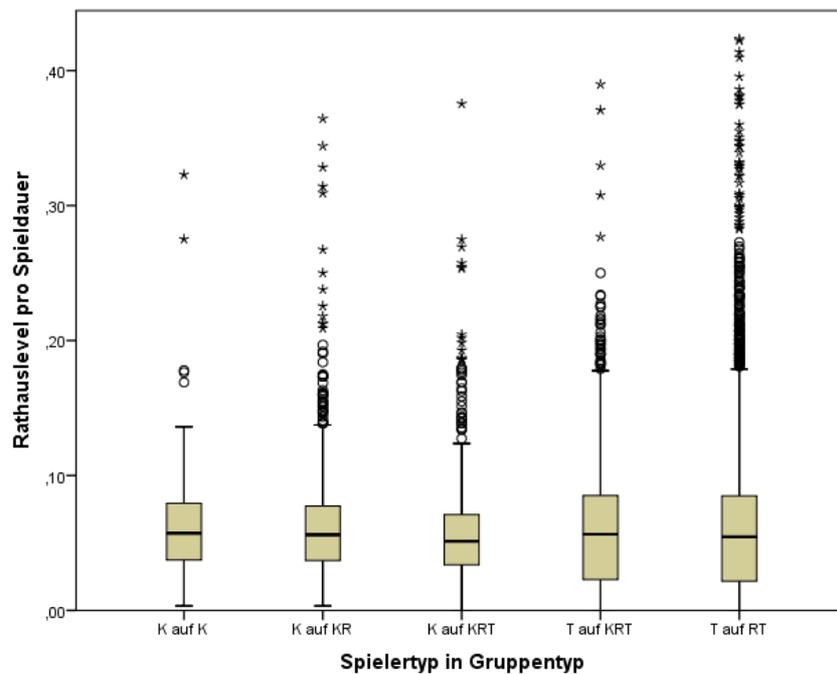


Abbildung 57. Verteilung des Erfolges zwischen Spielertypen für K- und T-Gruppen in Ikariam (für Griechenland und den 13.02)

3.15 Diskussion H4 Ikariam

Dieses Ergebnis entspricht den Erwartungen in keiner Weise. Die Messung des Erfolgs über das zweite verwendete Erfolgsmaß (RHL/SD, nicht korrigiert um die Stadtanzahl), ergibt kein wesentlich anderes Bild. Es ist schwer zu verstehen, warum Kooperierende im direkten Aufeinandertreffen Trittbrettfahrern in manchen Fällen im Erfolg überlegen sind. Möglicher-

weise sind die geringen Fallzahlen (n = 187-582) ein Grund dafür, dass sich kein klareres Ergebnis zeigt.

3.16 Resultate PGG Hypothese 5: Alle Spielertypen interagieren länger mit Kooperierenden als mit Trittbrettfahrern

Ein wichtiger Faktor für Kooperation ist die Dauer der Interaktion, da sowohl Vertrauen in den Partner als auch Gewinne mit längerer Dauer ansteigen können. Diese Überlegung wird durch die Logik bezüglich der möglichen Kombinationen von Strategien gestützt: Es ist für *jede* Strategie (K, R, T) vorteilhaft, mit Kooperierenden zu interagieren, da diese die meisten oder alle Kosten für Gemeingüter tragen. Das sollte dazu führen, dass die Wechselbereitschaft weg von Trittbrettfahrern (T) generell am größten ist, von Reziprokatoren-Strategien (R) etwas weniger stark ausgeprägt und bei Kooperierenden (K) kaum vorhanden sein sollte. Diese Hypothese lässt sich allerdings mit den Labor-PGG-Daten nicht überprüfen, da in allen Studien die Spieldauer vorgegeben und damit konstant für alle Spieler ist. Ich wende mich daher den Ikariam-Daten zu.

3.17 Resultate Ikariam Hypothese 5: Alle Spielertypen interagieren länger mit Kooperierenden als mit Trittbrettfahrern

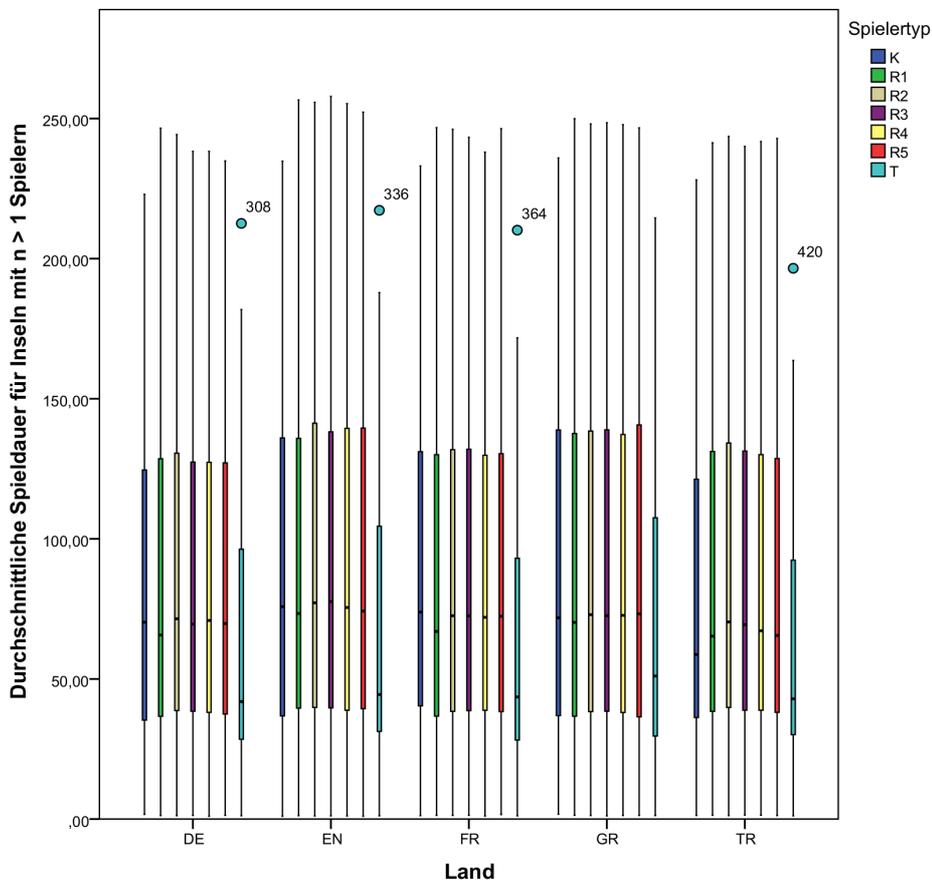
Um diese Hypothese zu überprüfen, wird zunächst die Spieldauer aufgeschlüsselt nach Spielertypen berechnet, um zu sehen, ob hier Unterschiede bestehen. Die folgende Tabelle 25 listet die statistischen Kennwerte für die Verteilung der Spieldauer über die Spielertypen auf.

Spielertyp	Mittelwert	Median	Standardabweichung	95% Konfidenzintervall des Mittelwertes
K	229,92	238,00	57,57	227,05 – 232,79
R1	246,43	263,00	53,36	245,34 – 247,13
R2	245,24	257,00	50,42	244,06 – 246,42
R3	241,91	250,00	49,54	240,21 – 243,61
R4	240,56	248,00	50,58	230,11 – 243,01
R5	241,98	248,00	49,04	238,76 – 245,21
T	207,07	211,00	68,07	205,41 – 208,72

Tabelle 25: Deskriptive Statistik für Spieldauer pro Spielertyp

Die folgende Abbildung 58 zeigt analog, dass es zwischen den Spielertypen nur zu Trittbrettfahrern Unterschiede gibt. Zwischen den Ländern ist kaum ein Unterschied zu verzeichnen, auch wenn die Spielzeiten von sehr kurz bis sehr lang reichen, die Spannweiten also extrem hoch sind.

Abbildung 58: Durchschnittliche Spieldauer pro Spielertyp in Tagen im Ländervergleich für Inseln mit mehr als einem Spieler



Für jedes Land sind alle paarweisen Vergleiche (nach Anzahl der Tests korrigiert) von Trittbrettfahrern (T) zu anderen Spielertypen für den letzten Snapshot signifikant mit negativen mittleren Rängen für Trittbrettfahrer (Kruskal-Wallis-Test, $df = 6$, DE: $n = 5724$; EN: $n = 4202$; FR: $n = 6409$; GR: $n = 7544$; TR: $n = 10572$; $p < 0,001$; einzige Ausnahmen sind T zu R5 in DE und T zu K in EN: n.s.). Praktisch alle Vergleiche zwischen Kooperierenden (K) und Reziprokatoren (R1 bis R5) bleiben dagegen nicht signifikant. Damit lässt sich feststellen, dass alle Spielertypen etwa gleich lang spielen, bis auf Trittbrettfahrer, die kürzer am Spiel teilnehmen als andere Spielertypen.

Dies bestätigt sich teilweise, wenn man genauer zwischen den Gruppentypen differenziert. Für die Inselzusammensetzung KRT sind bis auf Deutschland und England die Unterschiede zwischen Trittbrettfahrern und Kooperierenden jeweils signifikant (Mann-Whitney-U-Test, DE: $n = 187$, n.s.; EN: $n = 262$, n.s.; FR: $n = 249$, $p < 0,014$; GR: $n = 582$, $p < 0,001$; TR: $n = 444$, $p < 0,001$).

Um die Hypothese – längere Interaktion mit Kooperierenden – bestätigen oder verwerfen zu können, wird auf Unterschiede desselben Spielertyps zwischen Gruppen mit unterschiedlicher Zusammensetzung geprüft (R mit K bzw. R mit T). So unterscheiden sich R-Spieler in der Türkei nicht signifikant in ihrer Spieldauer zwischen KR und RT-Inseln (Mann-Whitney-U-Test, $n = 4601$, n.s.), tun dies aber in Frankreich (Mann-Whitney-U-Test, $n = 2650$, $p < 0,001$) mit einem höheren mittleren Rang für Reziprokatoren auf Kooperierenden-Inseln (KR). Für Deutschland, England und Griechenland ist das Gegenteil der Fall (Mann-Whitney-U-Test, DE: $n = 2365$, EN: $n = 1560$, GR: $n = 2995$, $p < 0,001$). Das heißt, auf RT-Inseln spielen Reziprokatoren signifikant länger mit Trittbrettfahrern als mit Kooperierenden zusammen. Für alle Vergleiche zwischen KT-Gruppen sind keine Berechnungen möglich, da keine KT-Inseln vorhanden sind.

Alle weiteren Mann-Whitney-U-Tests auf Unterschiede (K auf KRT gegen K auf KR, sind für Deutschland ($n = 190$), England ($n = 236$) und Frankreich ($n = 226$) nicht signifikant. In Griechenland ($n = 430$) und in der Türkei ($n = 322$) ist die Spieldauer allerdings jeweils signifikant länger für Kooperierende auf gemischten Inseln (KRT) als mit Reziprokatoren (KR) (Mann-Whitney-U, $p < 0,001$). Für Trittbrettfahrer sind für alle Länder alle Tests auf Unterschiede zwischen der Spieldauer auf RT und KRT-Inseln nicht signifikant (Mann-Whitney-U, DE: $n = 967$, EN: $n = 741$, FR: $n = 1016$, GR: $n = 1316$, TR: $n = 2459$, n.s.).

3.18 Diskussion H5 Ikariam

Da die mittleren Ränge weder für Kooperierende noch für Trittbrettfahrer ein klares Bild ergeben, da teils KRT-Inseln, teils KR-Inseln bzw. RT-Inseln im Mittel eine höhere Spieldauer aufweisen, kann nicht bestätigt werden, dass mit Kooperierenden länger als mit Trittbrettfahrern interagiert wird. Allerdings ist die Tatsache, dass es keine signifikanten Unterschiede gibt, in der Natur des Spieles begründet. Zum einen spielt jeder Teilnehmer auf vielen (bis zu zwölf Inseln) gleichzeitig. Eine komplette, auf vielen oder allen Inseln durchgeführte Assoziierung mit Kooperierenden (*assortment*) ist spieltechnisch nicht machbar, weil zu teuer bzw. zeitaufwändig. Das aber wäre nötig, um signifikante Unterschiede feststellen zu können. Zum anderen führen egoistischere Gruppenzusammensetzungen eher zum Verlagern von Ressourcen auf andere Inseln als zum Rückzug, was über die Spieldauer aber nicht zu messen ist. Für eine präzisere Prüfung dieser Hypothese ist daher eine andere Datenstruktur nötig.

3.19 Resultate PGG Hypothese 6: Je länger die Interaktionen, desto größer die Unterschiede zwischen Kooperierenden und Trittbrettfahrern

Diese Hypothese geht davon aus, dass Interaktionen auf längere Zeiträume hin ausgelegt sind.

Die Vorteile von Kooperation sollten sich gerade dann zeigen, wenn Kooperation sich nach einiger Zeit amortisieren kann. Diese Überlegung wird als Hypothese H6 formuliert: *Je länger die Interaktionen, desto größer die Unterschiede zwischen Kooperierenden und Trittbrettfahrern.*

Dahinter steckt die Überlegung, dass die Erzeugung von Gemeingütern zwar kostspielig ist, deren Aufrechterhaltung im Vergleich aber weit weniger. Beispiele sind etwa Infrastrukturprojekte, wie Straßen, Bewässerungssysteme oder ähnliches. Je länger demnach ein Gemeingut existiert, desto eher werden dessen Effizienzgewinne die Kosten amortisieren.

Wenn diese Hypothese stimmt, müssten sich die Verdienste je nach Strategie über die Zeit ändern, also eine negative Entwicklung für Trittbrettfahrer-Strategien und eine positive Entwicklung für kooperativere Strategien zu beobachten sein. Die untenstehenden Abbildungen (Abbildungen 59 und 60) zeigen die Entwicklung des Verdienstes über die Zeit wiederum als prozentualen Anteil des Verdienstes aufgeschlüsselt nach Treatments mit und ohne Bestrafung.

Mit Bestrafung erzielen Kooperierende ab der zweiten Runde den größten Erfolg (57,8% in der letzten Runde). Trittbrettfahrer treten in Treatments mit Bestrafung nicht auf. Die fünf Reziprokatoren folgen keinem erkennbaren Trend.

Erfolg Spielertypen im Zeitverlauf (nur P Treatments)

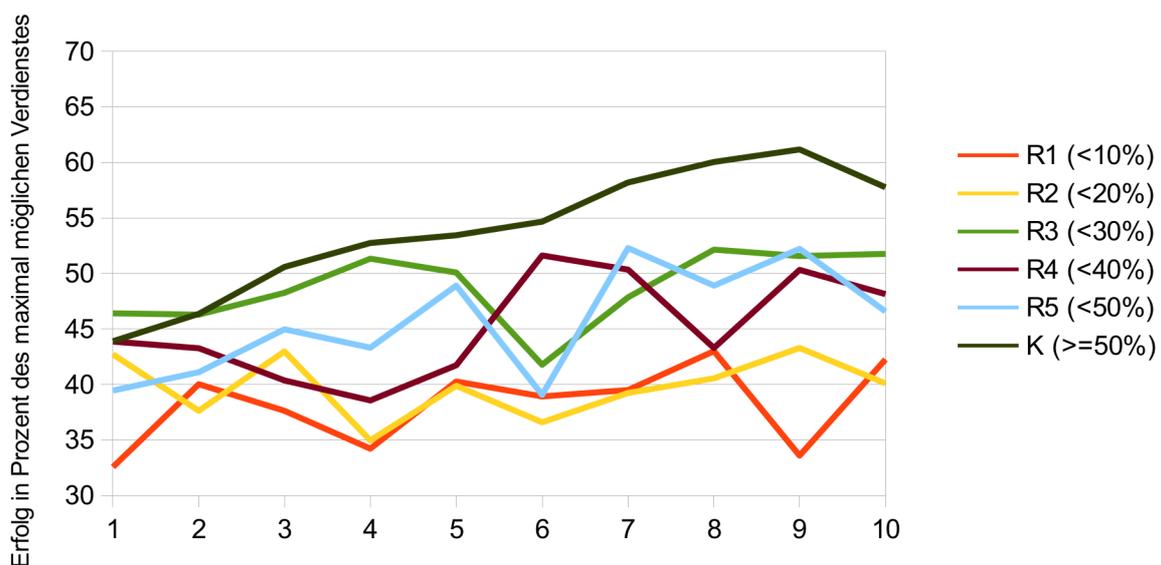


Abbildung 59: Verdienst pro Runde in Prozent des maximal möglichen Verdienstes pro Spielertyp über die Zeit in vier PGG-Studien für 10 Runden (nur P Treatments)

In der folgenden Abbildung 60 ist dagegen erkennbar, dass Trittbrettfahrer (T) zu Beginn den

höchsten und am Ende den niedrigsten Verdienst aufweisen. Kooperierende beginnen dagegen mit dem geringsten Verdienst unter allen Spielertypen, und finden sich nach dem höchsten Verdienst in der achten Runde am Ende unter den Verlierern. Der Spielertyp R2 weist in der ersten Runde den zweitletzten Verdienst, in der letzten Runde allerdings den höchsten Verdienst mit 55,2% des maximalen Verdienstes auf. Es fällt auf, dass sich die Spielertypen in ihrem Verdienst über die Runden angleichen. Liegen sie in der ersten Runde noch 14% auseinander (68,1% und 54,4%), so beträgt der Abstand in der zehnten Runde nur noch 5,7% (55,2% und 49,5%).

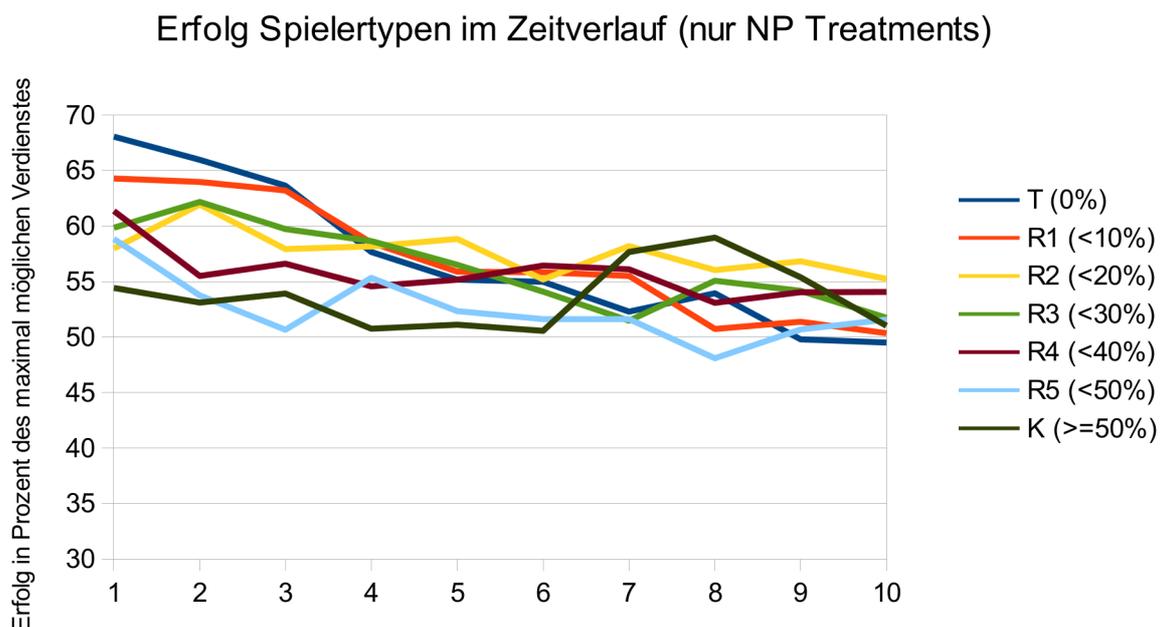


Abbildung 60: Verdienst pro Runde in Prozent des maximal möglichen Verdienstes pro Spielertyp über die Zeit in vier PGG-Studien für 10 Runden (nur NP Treatments)

Zwar bestätigt die statistische Analyse, dass die Unterschiede in der ersten Runde für Treatments ohne Sanktionen signifikant sind (Kruskal-Wallis-Test, $df = 6$, $n = 463$, $p < 0,001$), was in der zehnten Runde nicht mehr der Fall ist; es sind jedoch nur wenige paarweise Vergleiche signifikant verschieden (K zu T, R1, R2, R3 und R5 zu T).

Für Treatments mit Bestrafung ist es genau andersherum: die nicht signifikanten Unterschiede im Erfolg zu Beginn zwischen den Spielertypen sind gegen Ende des Spiels statistisch nachweisbar unterschiedlich (Kruskal-Wallis-Test, $df = 5$, $n = 304$, $p < 0,001$). Hier liegen die Kooperierenden vor den anderen Spielertypen, was die mittleren Ränge betrifft. Auch hier sind nur wenige paarweise Vergleiche signifikant verschieden (K zu R2, R4 und R5).

Da in diesem Datensatz auch eine Studie enthalten ist, die über einen längeren Zeitraum läuft

(50 Runden), wird auch hier der Erfolg (ab Runde 11 bis 50) gesondert analysiert. Die beiden folgenden Abbildungen (Abbildung 61 und 62) unterscheiden wiederum zwischen Bestrafung (PP) und Nicht-Bestrafung (PNP).

Verdienst in % des maximalen Verdienstes für Runden 11-50 (PNP)

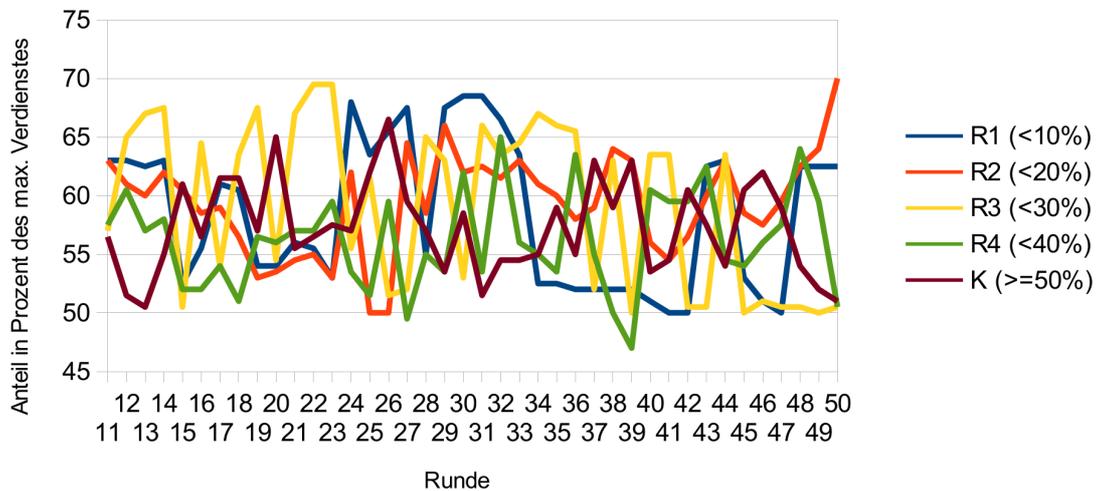


Abbildung 61: Verdienst pro Runde in Prozent des maximal möglichen Verdienstes pro Spielertyp von Runde 11 bis 50 für das Treatment PNP aus Gächter et al. 2008

Durch das andere Treatment (Bestrafung) ändern sich die auftretenden Spielertypen und der Verdienst wie folgt:

Verdienst in % des maximalen Verdienstes für Runden 11-50 (PP)

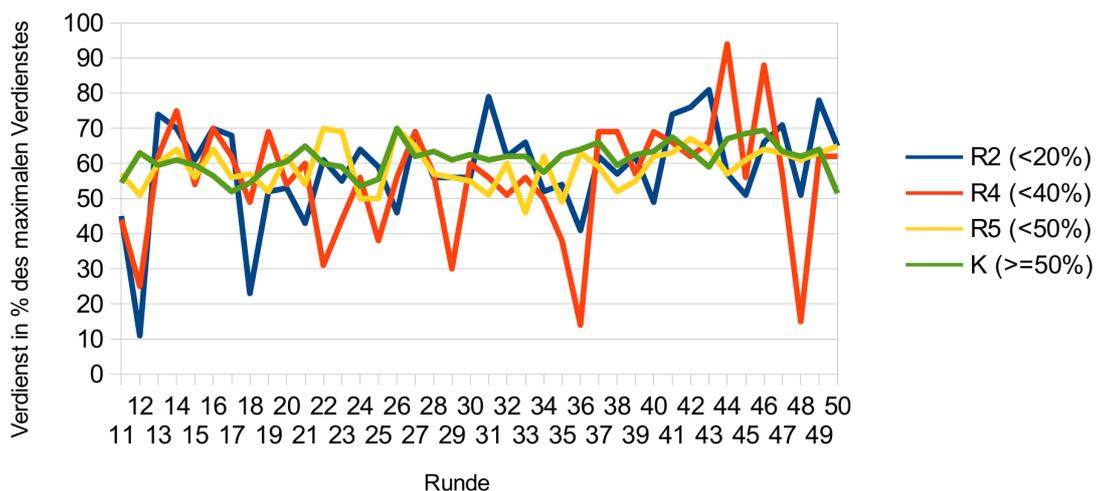


Abbildung 62: Verdienst pro Runde in Prozent des maximal möglichen Verdienstes pro Spielertyp von Runde 11 bis 50 für das Treatment PP aus Gächter et al. 2008

Die Abbildungen 61 und 62 und die Datenanalyse (nicht gezeigt) zeigen hier ein Bild ohne Tendenz.

3.20 Diskussion H6 PGG

Für die ersten zehn Runden ergibt sich ein die Hypothese bestätigendes Muster: Selbst ohne Bestrafung rutschen die Trittbrettfahrer vom ersten auf den letzten Platz, während mit Bestrafung Kooperierende sich im Verlauf des Spiels von den anderen Spielertypen nach oben absetzen können. Auch die kreuzweise wechselnden Signifikanzen der Unterschiede zu Beginn und gegen Ende sind hierfür ein Indiz. Allerdings sind die Unterschiede statistisch nicht robust nachweisbar, was vermutlich in der bereits erwähnten kurzen Spieldauer von nur zehn Runden liegt – Unterschiede können sich über diese kurze Zeitspanne offensichtlich noch nicht ausdifferenzieren.

Überraschenderweise ist für die Runden 11 bis 50 das obige Muster nicht erkennbar, auch wenn Trittbrettfahrer in beiden Treatments überhaupt nicht und R5 nicht in PNP, sowie R1 und R3 nicht in PP vorkommen. Ohne Bestrafung ergibt sich das theoretisch erwartete Bild, dass R1 und R2 einen höheren Erfolg erzielen als Strategien, die mehr in das Gemeingut investieren. Die Ausschläge im Treatment PP sind dagegen extremer, sowohl für R2 als auch R4. Diese Ausschläge sind vermutlich den Strafen durch andere Spieler geschuldet. Generell lässt sich festhalten, dass in allen drei betrachteten, doch sehr unterschiedlichen Datenzusammenstellungen alle Spielertypen bezüglich des Erfolges relativ eng zusammen liegen, und zwar zwischen 48% und 58% (für zehn Runden).

Zusammenfassend lässt sich diese Hypothese mit den Daten aus dem Labor nicht zweifelsfrei belegen. Auch wenn einige Indizien dafür sprechen, weisen andere dagegen keine Richtung auf.

3.21 Resultate Ikariam H6: Je länger die Interaktionen, desto größer die Unterschiede zwischen Kooperierenden und Trittbrettfahrern

Da die Interaktionen bei Ikariam über zehn Monate laufen, kann das Erfolgsmaß über einen relativ langen Zeitraum beobachtet werden. Die folgenden zwei Abbildungen (Abbildung 63 und 64) zeigen die auch bisher verwendeten Erfolgsmaße (Rathauslevel pro Spieldauer und Rathauslevel pro Spieldauer pro Anzahl Städte) aggregiert über alle Länder über die Zeit, d.h. von Beginn des Spieles bis zu den letzten Daten zehn Monate später.

Zu Beginn des Spieles (hier der nicht repräsentative erste Snapshot 1) wird deutlich, dass Trittbrettfahrer von einer Situation deutlich profitieren können, in der sich viele Teilnehmer erst noch strategisch orientieren müssen. Zwei Wochen später sind Kooperierende bereits erfolgreicher, was sich aber stetig angleicht. Dies dauert bis zum Snapshot 10. Ab diesem

Zeitpunkt sind Trittbrettfahrer wieder erfolgreicher. Danach (dem 26.9.2014), also genau der Hälfte greift wieder die (geringe) Überlegenheit der Kooperierenden.

Median Erfolg (RHL/SD) im Zeitverlauf

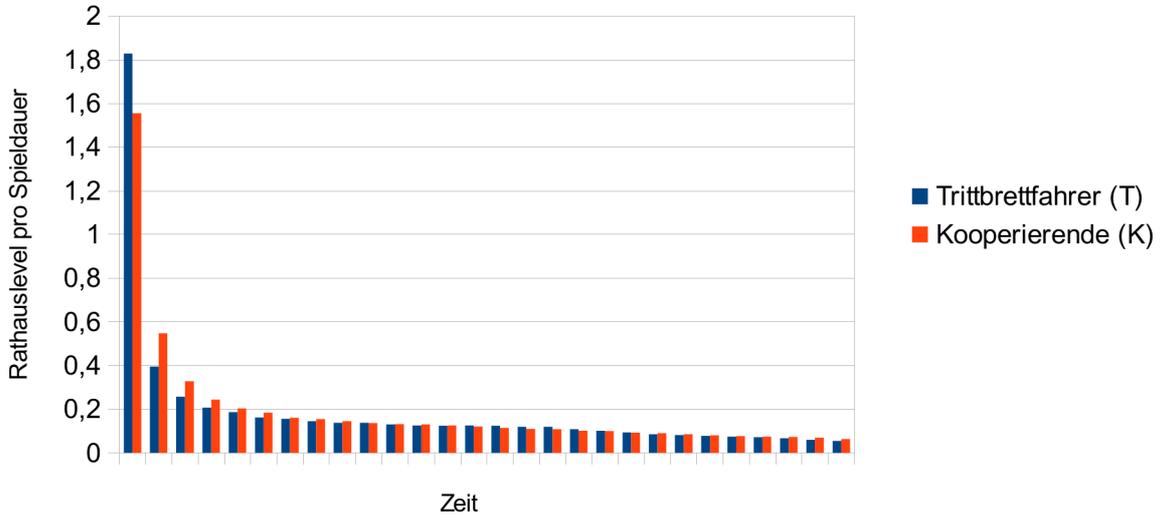


Abbildung 63. Median des Erfolges (RHL/SD) über alle Länder aggregiert im Zeitverlauf mit dem ersten Snapshot

Die untenstehende Abbildung, die das Erfolgsmaß um die Anzahl der Städte korrigiert (Abbildung 64) bestätigt diesen Befund.

Median Erfolg (RHL/SD) pro Stadt im Zeitverlauf

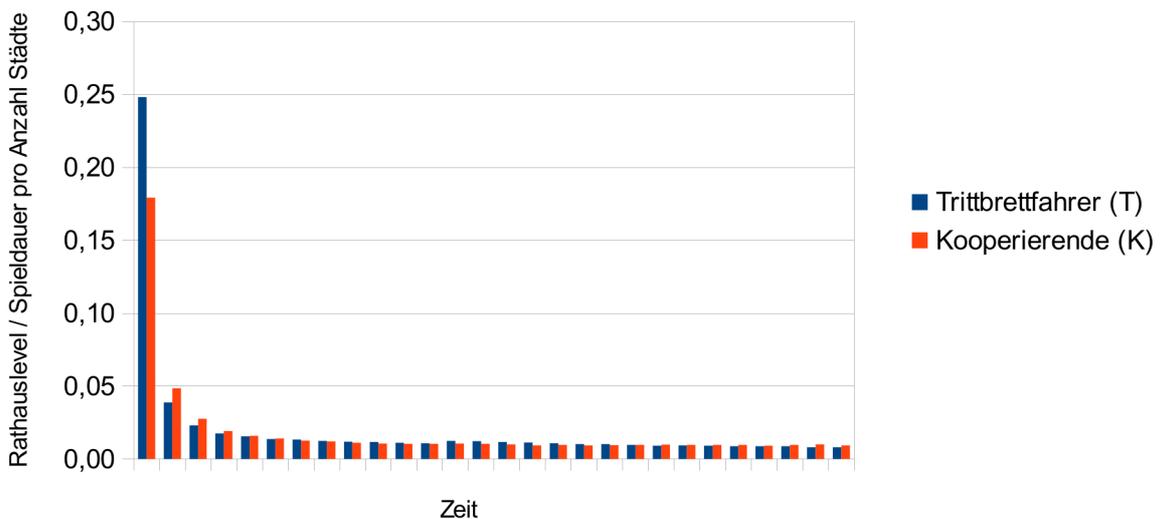


Abbildung 64. Median des Erfolges (RHL/SD) pro Anzahl der Städte über alle Länder aggregiert im Zeitverlauf mit dem ersten Snapshot

Diese Zusammenstellungen wären nicht vollständig ohne die anderen Spielertypen. Im Folgenden als Beispiel der Erfolgstrend für alle Spielertypen über die Zeit für alle Länder (Abbildung 65). Hier ergibt sich für alle Strategien ein extrem stabiles Bild in der Reihenfolge – R1 vor R2, R3, R4, R5. Etwas darunter liegen Kooperierende und Trittbrettfahrer.

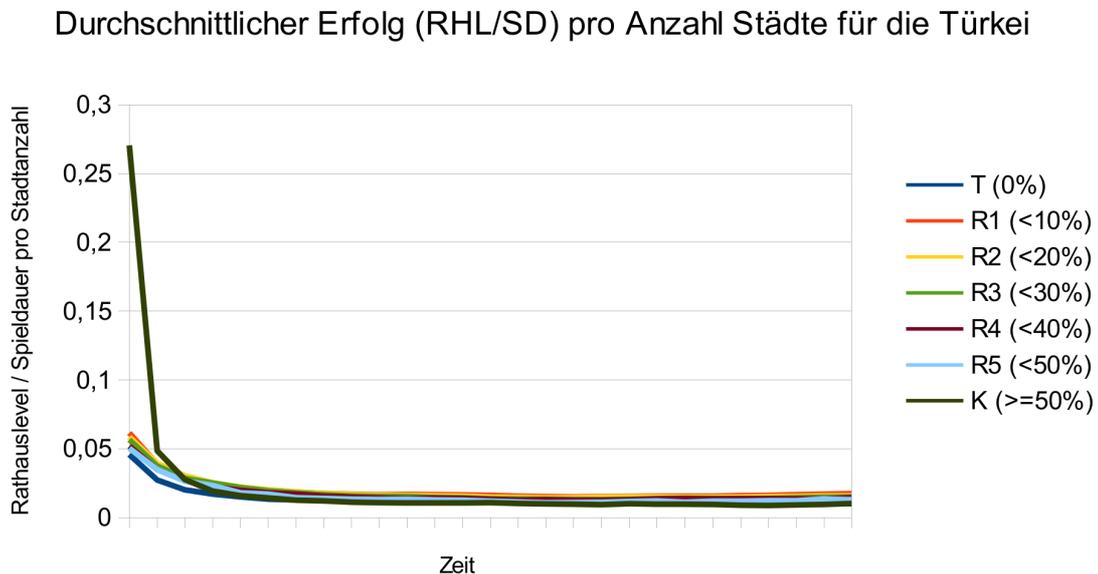


Abbildung 65: Durchschnittlicher Erfolg (RHL/SD) pro Stadtanzahl im Zeitverlauf für die Türkei ohne den ersten Snapshot

Ein Test auf Unterschiede des Erfolgsmaßes zwischen den Snapshots ist signifikant für alle Länder (Friedman-Rang-Test für verbundene Stichproben, $df = 27$, DE: $n = 92$, EN: $n = 92$, FR: $n = 135$, GR: $n = 137$, TR: $n = 167$, $p < 0,001$). In den korrigierten Post-hoc-Tests erweisen sich von den 351 paarweisen Vergleichen pro Land praktisch alle als signifikant.

3.22 Diskussion H6 Ikariam

Die Hypothese wird durch den Datensatz der Labor-PGG für die ersten zehn Runden gut bestätigt. Die beiden Datenreihen, die strategisches Verhalten über einen längeren Zeitraum beobachten (Gächter et al. 2008 und Ikariam) zeigen jedoch keine eindeutige Bestätigung. Zwar geht bei Ikariam der Trend möglicherweise langfristig zu den Kooperierenden, möglicherweise dreht er sich jedoch auch für den Zeitraum nach der Datensammlung, für den keine Daten vorliegen. Für die Türkei ist sogar der gegenläufige Trend zu beobachten. Insgesamt liegt demnach keine eindeutige Bestätigung vor. Es zeigt sich nicht die erwartete, zu Ende stark steigende Überlegenheit der Kooperierenden. Die statistisch signifikanten Unterschiede zwischen den Zeitpunkten für Ikariam erklären sich durch die zu Beginn des Spieles möglichen starken Wachstumsraten, die sich im Verlauf durch die stark steigenden Kosten sehr abschwächen.

3.23 Resultate PGG Hypothese 7: Die R1-Strategie mit Trade-off zwischen Investitionen in das öffentliche Gut und persönlichem Vorteil sollte sehr guten Erfolg zeigen

Die letzte Hypothese H7 wendet sich ab von den bisher behandelten Spielertypen mit extremen Strategien – Trittbrettfahrern (T) und Kooperierenden (K) – und zielt darauf, die Erfolgchancen einer Trade-off-Strategie (R1) genauer zu analysieren. Die Hypothese lautet: *Strategien mit Trade-off zwischen Investitionen in das öffentliche Gut und persönlichem Vorteil (R1) sollten sehr guten Erfolg zeigen*. Der Abgleich erfolgt hier zwischen der Anpassung an die Investitionen anderer – deshalb wird von R1-Spielern überhaupt in das Gemeingut eingezahlt – und der dominanten Strategie, nichts einzuzahlen.

Drei Überlegungen stecken hinter dieser Hypothese. Erstens sollten R1-Strategien durch ihre geringeren Investitionen in das Gemeingut mehr profitieren als R2 bis R5 sowie K. Zweitens könnten ihre Investitionen, auch wenn sie gering sind, die prinzipielle Kooperationsbereitschaft von R2-R5 Spielertypen besser aufrechterhalten als es eine Trittbrettfahrer-Strategie vermag. Damit ergibt sich, drittens, möglicherweise eine Situation, in der andere Spielertypen die Hauptlast für die Erzeugung von Gemeingütern tragen. Durch die geringeren Investitionen von R1-Spielern und infolgedessen der anderen wird allerdings nicht das soziale Optimum erreicht – im Vergleich zu Situationen, die von Trittbrettfahrerverhalten geprägt sind, sollten diese Strategien allerdings besser abschneiden.

Ein Test über alle Treatments unabhängig von den Gruppentypen zeigt signifikante Unterschiede zwischen den Spielertypen bezüglich des Erfolgs (Kruskal-Wallis-Test, $n = 12182$, $p < 0,001$; siehe auch Abbildung 41, S. 124 und dortige Post-hoc-Tests). Die Trade-off-Strategie R1 erreicht aber nur den vierten Rang im Erfolg, deutlich hinter T und K und knapp hinter R2. Separiert man nach Treatment, so ist die R1-Strategie in beiden Bestrafung-treatments (partner und stranger) jeweils letzte. Ist keine Bestrafung möglich, so erreicht sie im Stranger-Treatment den zweiten nach Trittbrettfahrern, im Partner-Treatment den fünften Platz (siehe auch Abbildung 42, S. 125).

Analysiert man den Erfolg von R1 je nach Gruppentyp, so fährt diese Strategie am besten in KR-Gruppen, gefolgt von KRT-Gruppen und reinen Reziprokatoren-Gruppen (R). Am schlechtesten schneidet sie in Gruppen mit Trittbrettfahrern (RT) ab (Kruskal-Wallis-Test, $df = 3$, $n = 722$, $p < 0,001$; siehe Abbildung 66). In den paarweise korrigierten Post-hoc-Vergleichen fallen nur RT zu R und KRT zu KR nicht signifikant aus. Das bedeutet, dass R1 von Kooperierenden profitieren kann und von Trittbrettfahrern geschädigt wird. Damit kann diese

Hypothese nicht bestätigt werden. Stattdessen ergibt sich das Bild aus Hypothese 1, bei der Trittbrettfahrer am erfolgreichen sind.

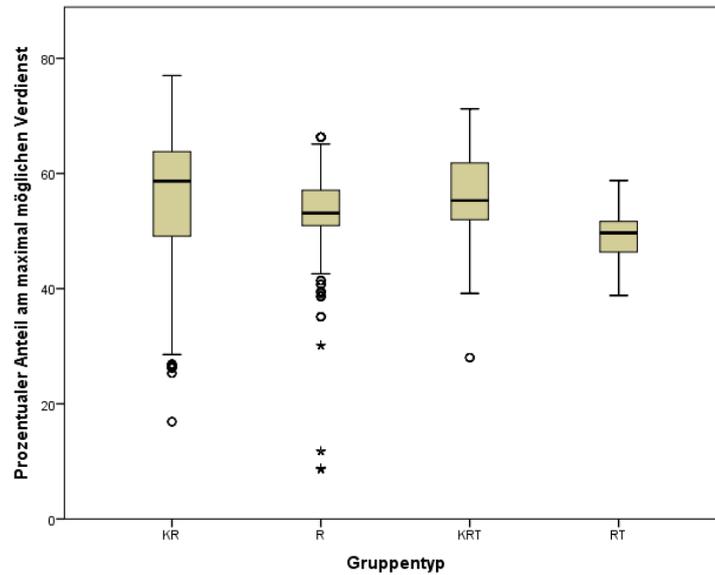


Abbildung 66. Test auf Unterschiede im Erfolg von Reziprokatoren (R1-R5) pro Gruppentyp (Verdienst in % des maximal möglichen Verdienstes für vier PGG-Studien)

3.24 Resultate Ikariam Hypothese 7: Die R1-Strategie mit Trade-off zwischen Investitionen in das öffentliche Gut und persönlichem Vorteil sollte sehr guten Erfolg zeigen

Die Ergebnisse der Erfolgsvariablen wurden bereits bei Hypothese 1 ausführlich diskutiert (siehe auch Abbildung 44). Zur Vervollständigung wird die zweite Erfolgsvariable (RHL/SD korrigiert um die Stadtanzahl) gezeigt. Es ist zu sehen, dass R1 in allen Ländern den höchsten oder zweithöchsten Erfolg erzielen kann. Beide Extremstrategien (K und T) sind weit abgeschlagen.

Erfolg von Spielertypen (RHL/SD) pro Stadtanzahl und Median

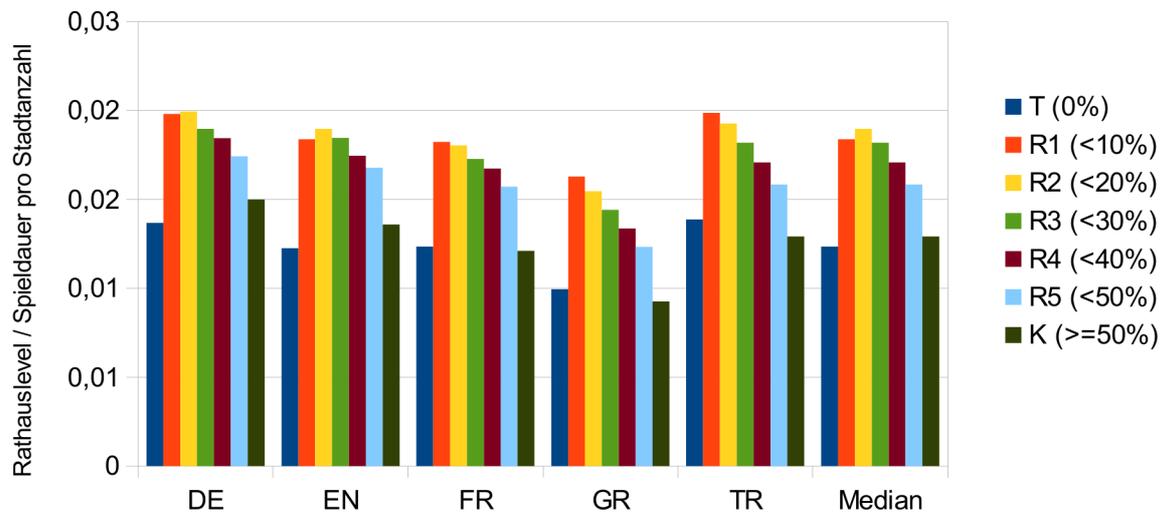


Abbildung 67. Erfolg von Spielertypen (RHL durch SD pro Stadtanzahl) und Median im Ländervergleich (alle Snapshots ohne 25.04)

Wie sieht es nun zu dem letzten verfügbaren Zeitpunkt, nach zehn Monaten Spiel aus? Hier ergibt sich folgendes Bild:

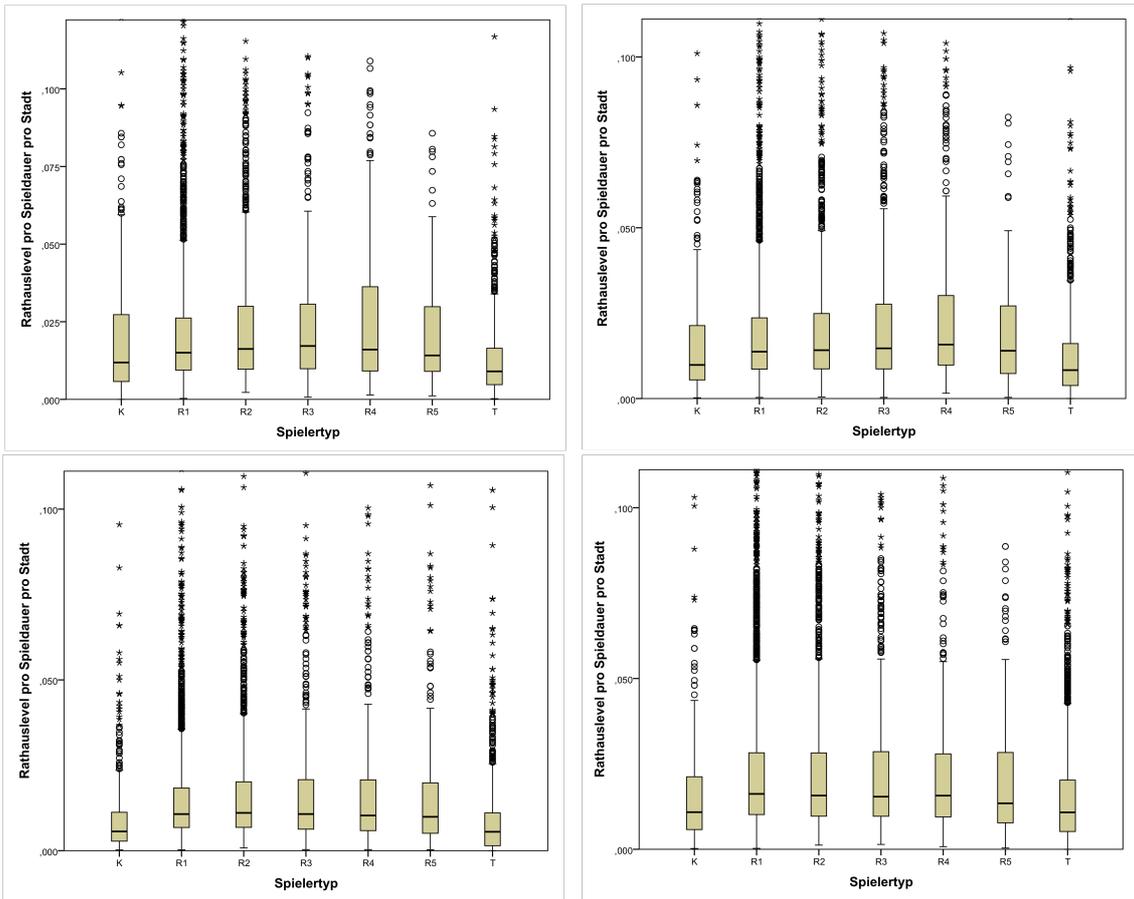


Abbildung 68: Erfolg (Rathauslevel pro Spieldauer und Stadtanzahl) pro Spielertyp und Land (im Uhrzeigersinn: DE, FR, TR, GR) nach 10 Monaten

Damit ergeben sich über die Länder hinweg die folgenden statistischen Kennwerte:

Spielertyp	Median	Standardabweichung
T	0,0082	0,0161
R1	0,0136	0,0238
R2	0,0142	0,0268
R3	0,0149	0,0295
R4	0,0147	0,0274
R5	0,0131	0,0293
K	0,0084	0,0241

Tabelle 26: Deskriptive Statistik für den Erfolg (RHL/SD) pro Stadtanzahl pro Spielertyp nach 10 Monaten

Die Strategie R1 kann also nach einer Spielzeit von zehn Monaten nur die Extremstrategien Kooperierende, Trittbrettfahrer und R5 schlagen. Dagegen sind alle stärker in das Gemeingut investierenden Reziprokatoren (R2-R4) erfolgreicher als R1. Pro Land sind diese Unterschiede signifikant (Kruskal-Wallis-Test, $df = 6$, DE: $n = 5724$, EN: $n = 4202$, FR: $n = 6409$,

GR: $n = 7544$, TR: $n = 10572$, $p < 0,001$), wobei R1 immer besser als Kooperierende und Trittbrettfahrer abschneidet, und in der Türkei die beste Strategie stellt, in allen übrigen Ländern jedoch von höher investierenden Reziprokatoren übertroffen wird (DE: R2, R3, R4; EN: R2, R3, R4, R5; FR: R3, R4; GR: R2, R3). In den korrigierten Post-hoc-Tests erweisen sich alle paarweisen Vergleiche mindestens auf dem 0,04-Niveau als signifikant, bis auf R1-R2, R3-R4, R4-R5 und K-T.

3.25 Diskussion H7 Ikariam

Wie in den PGG-Daten kann die R1-Strategie auch bei längerfristigen Interaktionen nicht als Sieger gelten, was den Gesamterfolg betrifft. Zwar liegt sie im Mittel über die Zeit vorne, nach zehn Monaten scheint das Kosten-Nutzen-Optimum der Gemeingüter allerdings eher bei einer Investition von etwa 30-40% der eigenen Ressourcen in die Gemeingüter zu liegen. Diese Erklärung zielt auf die Amortisation der Ressourcen durch höhere Erträge über längere Zeiträume (zehn Monate) ab. Möglicherweise ziehen R3 und R4-Spieler auch weniger Strafen auf sich.

E Gesamtfazit und allgemeine Diskussion

1 Zusammenhang der untersuchten drei Kooperationsstudien

Bevor diese Arbeit zu einem Gesamtfazit kommt, wird der thematische Zusammenhang der behandelten drei größeren Analyseblöcke aufgezeigt. Die erste Klammer besteht im gemeinsamen Ausgangspunkt, der Frage nach Einflussfaktoren auf menschliches Kooperationsverhalten. Unter dieser Leitfrage adressieren alle drei Analysen zwei zentrale Fragestellungen der Kooperationsforschung – individuelle Strategien sowie Abwehr von Trittbrettfahrern beispielsweise durch Bestrafung. Hinter den ersten beiden Analysen (Differenzierung von Spielertypen bzw. Effizienz von Bestrafung) steht zudem der Versuch, Befunde auf der Ebene der Methodik auf eine tragende Grundlage zu bringen. Im ersten Fall prüft diese Arbeit bisherige Befunde auf ihre Stabilität, indem das Paradigma PGG in einem veränderten Kontext (in einem Online-Spiel) aufgenommen wird, um es mit einer größeren Anzahl an Personen in mehreren Ländern zu überprüfen. Im zweiten Fall bedeutet die vorliegende Analyse eine kritische Reinterpretation und Differenzierung vorhandener Ergebnisse.

Eine zweite Klammer betrifft die Ergebnisse, die jeweils zu einem besseren institutionellen Design führen sollen. Da die zweite Studie Strafen als Institution *rehabilitieren* kann, indem sie deren Effizienz belegt, ergibt sich in Verbindung mit anderen Resultaten, etwa der Wirk-

samkeit von nicht-monetären (sozialen) Strafen (Masclet et al. 2003) und der wirksamen Kombination von Bestrafungs- und Reputationsmechanismen (Rockenbach & Milinski 2006), eine unmittelbare Anwendung auf Institutionen im realen Leben. Mit den gewonnenen Erkenntnissen lassen sich einige Institutionen zudem möglicherweise effizienter gestalten.

Dies gilt ebenso für Spielertypen: Hat man erst einmal erkannt, dass Menschen verschiedene Kooperationsstrategien bevorzugen, liegt es auf der Hand, zu treffende Maßnahmen auf diese Typen zuzuschneiden. So zeigen Burlando & Guala 2005, dass Gruppen, die nur aus konditional Kooperierenden bestehen, fast perfekte Kooperation erreichen. Dies beobachtet man vermutlich, weil keine Trittbrettfahrer die Einzahlungen nach unten ziehen (vgl. auch Page et al. 2005, der ähnliche Ergebnisse für die Sortierung von Gruppen nach Einzahlungshöhe berichtet). Um höhere Kooperationsraten zu erzielen, wären demnach Gruppen aus konditionalen Kooperierenden mit jeweils einem Kooperierenden optimal – Trittbrettfahrer müssten aus diesen produktiven Gruppen aus- und untereinander zusammengeschlossen werden.

Und schließlich lässt sich, als dritte Klammer, auf den Ergebnissen wechselseitig aufbauen. Eine Differenzierung nach Spielertypen kann nämlich nicht nur bezüglich der Gruppenzusammensetzung genutzt werden, sondern gerade auch in Hinblick auf Strafen. Es ist anzunehmen, dass konditionale Kooperierende deutlich besser als Trittbrettfahrer auf Strafen wie Reputationsverlust oder Missbilligung ohne monetäre Konsequenzen ansprechen werden. Bestrafung existiert in vielen Variationen – z. B. zentral-dezentral; monetär-nicht-monetär; kostenloses-kostspieliges Strafen. Diese Variationen könnten gerade in Bezug auf die Differenzierung von Spielertypen auf die jeweilige Wirksamkeit geprüft werden, wie ich es bereits für einige Fragestellungen in Studie 2 demonstriert habe.

2 Neue Erkenntnisse

Der besseren Übersichtlichkeit halber werden zunächst die Untersuchungsergebnisse zu den Arbeitshypothesen in einer Tabelle (Tabelle 27) zusammengefasst.

Studie und Nummer der Hypothese im Text	Hypothese in Stichworten	Bestätigung durch Ikariam-Daten	Bestätigung durch PGG-Daten
S1 – H1	Spielertypen sind auch außerhalb von Laborbedingungen auffindbar	Ja	Hier nicht relevant
S1 – H2	Spielertypen sind unter verschiedenen Bedingungen und Kontexten stabil	Ja	Teilweise ja
S2 – H1a	Kooperierende strafen häufiger als Trittbrettfahrer	Nicht möglich	Tendenziell ja, mit unklarer Befundlage
S2 – H1b	Trittbrettfahrer werden häufiger bestraft als Kooperierende	Nicht möglich	Ja
S2 – H2	Spielertyp und Straftyp treten in bestimmten Kombinationen auf	Nicht möglich	Nein
S2 – H3	Gruppenzusammensetzung ändert das Strafverhalten	Nicht möglich	Ja
S2 – H4	Unabhängig vom Spielertyp wird anfangs mehr als im weiteren Verlauf bestraft	Nicht möglich	Ja
S2 – H5	Hohe Einzahlungen und Bestrafung führen zur Überlegenheit von Kooperierenden	Nicht möglich	Ja
S3 – H1	Trittbrettfahrer sind ohne zusätzliche Randbedingungen Kooperierenden überlegen	Nein	Ja
S3 – H2	Kooperierende unter sich sind Trittbrettfahrern überlegen	Ja	Ja
S3 – H3	Gruppenerfolg hängt von Gruppenzusammensetzung ab (Erwartung: K, KR, R, KRT, KT, RT, T)	Teilweise ja	Tendenziell ja, aber mit wenigen Daten
S3 – H4	Direktes Aufeinandertreffen von Kooperierenden und Trittbrettfahrern favorisiert Trittbrettfahrer	Nein	Ja
S3 – H5	Interaktionsdauer ist mit Kooperierenden länger als mit Trittbrettfahrern	Nein	Nicht möglich, weil Dauer konstant
S3 – H6	Unterschiede zwischen Kooperierenden und Trittbrettfahrern werden mit der Zeit größer zu Gunsten von K	Nein	Teilweise ja
S3 – H7	Trade-off-Strategien zwischen Kooperation und Egoismus können Erfolg maximieren	Nein	Nein

Tabelle 27. Zusammenfassung der Ergebnisse zu den Arbeitshypothesen der drei Studien

Im folgenden Abschnitt werden die Befunde zu den Hypothesen in Abschnitt D aufgenommen,

es wird zusammengefasst, was die Arbeit an neuen Erkenntnissen gebracht hat und dies wird zu einer Gesamtschau auf das Phänomen Kooperation verbunden. Dies geschieht vor allem im Hinblick auf die zu Grunde liegenden evolutionären Mechanismen. Den Abschluss bilden die allgemeinen Schlussfolgerungen, die diese Arbeit für die Kooperationsforschung ableitet.

Was hat die vorliegende Arbeit an neuen Erkenntnissen gebracht?

1. Die in der ökonomischen Experimentalliteratur beschriebenen Spielertypen (Fischbacher et al. 2001) konnten auch außerhalb des Laborkontextes in einem Online-Spiel bestätigt werden.

Diese Erkenntnis ist gerade im Hinblick auf die oft kritisierte Künstlichkeit der Laborumgebung von Bedeutung. Diese Arbeit kann relativ stabil bleibende Kooperationsniveaus von Spielertypen auch in einem anderen Kontext, der nicht durch Geld motiviert ist, nachweisen.

2. Die extremen Variationen in der Häufigkeit dieser Spielertypen konnten auf die Verwendung verschiedener Definitionen zurückgeführt werden. Durch die Berechnung der Häufigkeiten für alle gängigen Definitionen an einem Datensatz (Ikariam) konnten die bisherigen Literaturangaben vergleichbar gemacht werden.

Ein Ziel dieser Arbeit war es, ein Maß für die Robustheit der bisherigen und der vorliegenden Ergebnisse in Bezug auf die Häufigkeiten von Spielertypen zu gewinnen. Dies geschah nicht nur durch hohe Spielerzahlen und einen Ländervergleich in einem anderen Kontext, sondern auch durch die Gegenüberstellung bisheriger Resultate und Definitionen. Ein aussagekräftiger Vergleich wurde allerdings nur durch die Anwendung *verschiedener* Definitionen an *einem* Datensatz möglich.

3. Der oft vernachlässigte Spielertyp Altruist / stark Kooperierender wurde erstmals robust gemessen. Sein Anteil liegt bei etwa 5 % der Individuen in Gemeingutsituationen (Ikariam).

Nur wenige Studien führen den Spielertypus Kooperierende überhaupt auf. Dementsprechend schwanken die Schätzungen bezüglich der Häufigkeit seines Auftretens. Konsistent mit den theoretischen Erwartungen findet diese Arbeit einen kleinen, aber stabilen Anteil. Wichtig ist, dass diese Individuen unabhängig von den Investitionen der anderen agieren – sie investieren einen Großteil der eigenen Ressourcen in das Gemeingut, auch wenn Trittbrettfahrer vorhanden sind.

4. Durch die erstmals mögliche Unterscheidung zwischen pessimistischen konditional Kooperierenden und Trittbrettfahrern (Kocher et al. 2008) konnte diese Arbeit den durchschnittlichen Anteil an tatsächlichen Trittbrettfahrer präziser bestimmen. Er liegt bei einem Computerspiel im Durchschnitt von fünf Ländern und 18 000 Versuchspersonen bei etwa 25 % der Individuen in Gemeingutsituationen.

Trittbrettfahrer ist derjenige Spielertyp, der am leichtesten zu definieren und von anderen abzugrenzen ist. Konsistent mit bisherigen Ergebnissen diverser Studien kann diese Arbeit den Anteil von Trittbrettfahrern bei Gemeingutproblemen bei etwa 25 % verorten.

5. Diese Arbeit bietet für zwei bislang ungeklärte Effekte aus der Experimentalliteratur – den *endgame*- und den *partner/stranger*-Effekt – eine evolutionäre Erklärung an.

Eine evolutionäre Interpretation kann auch einige bislang ungeklärte Effekte aus der ökonomischen Experimentalliteratur erhellen (vgl. Frey & Rusch 2012). So sollte etwa aus ökonomischer Sicht kein Unterschied zwischen Gruppen, die ihre Zusammensetzung nicht ändern (Partner) und Gruppen, in denen dies der Fall ist (Stranger), feststellbar sein, weil dies an der nutzenmaximierenden Strategie nichts ändert. Man beobachtet allerdings durchgängig höhere Einzahlungen in das öffentliche Gut bei ersteren (Keser & van Winden 2000; Chaudhuri 2011). Aus soziobiologischer Sicht ist eine Trennung in *in-* und *outgroup* gut belegt und erklärt (Voland 2013). Meine Erklärung für Bestrafung, die eine *Kooperationsnorm* in einer Gruppe etabliert, fügt sich dem nahtlos an.

Ein weiterer ungeklärter Befund ist der *Endgame*-Effekt. In der letzten Runde fallen typischerweise die Einzahlungen rapide. Damit steigt gleichzeitig die Bestrafung steil an. Da Strafen ja auf Verhaltensänderung abzielen, ist dies aus einer nutzenmaximierenden Perspektive sinnlos und nicht erklärbar. Warum aber wird dennoch bestraft? Das wäre etwa dann rational, wenn unsere Heuristiken Kooperationsinteraktionen als primär zwischen *Personen* und nicht Situationen (pro PGG) auffassen. Dann wiederum ist es sinnvoll, eine Person, mit der man vermutlich wieder (in einer anderen Situation) zusammenarbeiten muss (was sich aus der Tatsache der Kleingruppen bei Jägern und Sammlern ergibt), am Ende einer Interaktion zu warnen, dass weiteres Fehlverhalten in einer neuen Interaktion nicht geduldet werden wird. Bei dieser Betrachtung ist es selbstverständlich rational, am Ende von Situation A zu bestrafen, da die Interaktion zwischen den beteiligten Personen ja meist *weitergeht*.

6. Diese Arbeit konnte – gegen die vorherrschende Forschungsmeinung – zeigen, dass Sanktionen nicht die Wohlfahrt schmälern, sondern im Gegenteil vergrößern. Dieses

Muster konnte diese Arbeit selbst in den *für die gegenteiligen Behauptungen herangezogenen* Experimentaldaten nachweisen (etwa bei Fehr & Gächter 2000 und Fehr & Gächter 2002).

Strafkosten nehmen schnell ab und sind bei längerfristigen Interaktionen effizienter als Umgebungen bzw. Institutionen ohne Sanktionsmöglichkeit. Trittbrettfahrer werden anfangs hart durch die Gruppe bestraft und verlieren einen Großteil ihrer Gewinne, wenn sie ihr Verhalten nicht kooperativer gestalten. Da viele Trittbrettfahrer auf Kooperation wechseln, ergibt sich eine höhere Effizienz. Diese zeitliche Dynamik ist weitgehend unabhängig von Spielertypen.

7. Diese Arbeit fasst Bestrafung als ein ubiquitär und universal verwendetes Mittel wehrhafter Kooperierender auf. Kleine Strafen zu Beginn, die graduell gesteigert werden können, etablieren eine Gruppennorm. Die bloße Existenz reicht dann in der Folge als plausible Drohung aus, um Kooperation zu gewährleisten.

Die zeitliche Dynamik von Sanktionen zeigt deutlich, dass *zu Beginn* festgelegt wird, wie die sozialen Interaktionen zukünftig ablaufen werden, d.h. welches Kooperationsniveau etabliert werden kann. Die erstaunlich hohe Anzahl an Individuen, die bereit sind zu strafen, ist aus evolutionärer Sicht ein Beleg dafür, dass die Kosten nicht so hoch sein können, wie oft angenommen wird. Gemäß dem *folk theorem* können sie auch egoistischen Motiven entspringen und den individuellen Nutzen erhöhen.

Damit haben sich vor allem zwei Mechanismen als wichtig für funktionierende menschliche Kooperation erwiesen: effiziente Bestrafung und die aktive Assoziierung mit Kooperierenden (etwa über Reputation). Kooperative Strategien müssen allerdings mit einer großen Häufigkeit heterogener Gruppenzusammensetzungen rechnen und damit zurechtkommen können.

8. Diese Arbeit qualifiziert die bisherige Forschung dahingehend, dass die besonderen Stärken von Sanktionen bei konstant zusammengesetzten, homogenen Gruppen mit längerfristigen Interaktionen genauer verortet werden. Ich weise dieses Muster in zwei realen Kontexten nach: sowohl in sozial-ökologischen-Systemen als auch bei Jägern und Sammlern (vgl. Frey & Rusch 2012). Man findet jeweils geringfügige, billige und angepasste Strafen, die nur selten bis zum Extrem eskaliert werden (Guala 2012; Ostrom 1990). Damit lässt sich Bestrafung als notwendiger Erfolgsfaktor für Kooperation bei Menschen etablieren.

Die verschiedenen Kontexte belegen, dass Sanktionen ein ausgezeichnetes Mittel sind, um

Kooperation zu etablieren. Allerdings müssen bestimmte Charakteristika in der Situation erfüllt sein – dazu gehören der Wiederholungscharakter sozialer Interaktionen und die Möglichkeit, Kosten dadurch zu senken, dass bestimmte, allen Gruppenmitgliedern bekannte Regeln nur von Zeit zu Zeit durchgesetzt werden müssen.

9. Diese Arbeit analysiert zum ersten Mal den Erfolg von Spielertypen im Vergleich.

Bislang schien dieser Zusammenhang so offensichtlich, dass sich meines Wissens keine Studie damit befasst. Trittbrettfahrer sollten immer besser fahren als Individuen, die in Gemeingüter einzahlen. Diese Arbeit kann diese Vermutung bestätigen (Studie 3, Hypothese 1), qualifiziert sie allerdings dahingehend, dass Kooperierende unter bestimmten Randbedingungen die erfolgreichste Strategie darstellen (Studie 3, Hypothese 2).

10. Diese Arbeit verknüpft zwei Forschungsgebiete, indem der Erfolg verschiedener Kooperationsstrategien in *public goods games* mit einer Diskussion der zu Grunde liegenden evolutionären Mechanismen verbunden wird.

Der Grad des kooperativen Verhaltens in sozialen Interaktionen ist aus biologischer Sicht ein Kernelement für die Fitness menschlicher Individuen. In ökonomischen Experimenten kommen diese Erklärungsansätze jedoch kaum vor. Die vorliegende Arbeit liefert einige Schnittstellen, an denen diese beiden Gebiete nutzbringend verbunden werden können.

11. Diese Arbeit erklärt über die evolutionäre Funktionslogik den hohen Anteil konditional kooperierender Strategien als Reaktion auf Trittbrettfahrer.

Reziprokerische Strategien stellen in vielen (und auch der vorliegenden) Studie den Hauptanteil. Aus evolutionärer Sicht gibt es mehrere Gründe dafür: erstens können die Umgebungsparameter oft wechseln; als flexible Strategie sind Reziprokatoren hierfür gut angepasst. Das gilt insbesondere für die Anpassung an andere Strategien – es ist gezeigt worden, dass starre Mechanismen, die unabhängig von anderen Strategien arbeiten, hier schnell ausselektiert werden (Axelrod 1984/2000).

Die zwei voll bestätigten Hypothesen (Studie 3, Hypothesen 1 und 2, siehe Tabelle 27) machen deutlich, dass der Erfolg der eigenen Strategie stark von den Strategien abhängig ist, mit denen man interagieren muss. Wie häufig eine bestimmte Gruppenzusammensetzung vorkommt, wie leicht Migration möglich ist und ob Trittbrettfahrer abgewehrt werden können, sind dafür entscheidende Parameter. Diese Abhängigkeit von anderen Strategien impliziert, dass vermutlich ein hoher Selektionsdruck darauf herrscht, kooperatives Verhalten flexibel

umzustellen und an die Erfordernisse anzupassen. Dies wird durch die hohen Anteile konditional Kooperierender in Experimenten (Fischbacher et al. 2001; Rustagi et al. 2010; ich finde 19%-67%, je nach Metrik) unterstützt.

Ist dieses Umfeld unkooperativ, so haben wir gesehen, dass es entweder erfolgreich sein kann, auf Defektion zu wechseln (in PGG), oder aber sich in kleine kooperative Gruppen zurückziehen (in Ikariam, Gruppentyp K), um von den Effizienzvorteilen eines kleinen Gemeinguts zu profitieren – durch eigene Investitionen und durch die des kleinen Anteils unkonkonditional Kooperierender.

Ist das Umfeld dagegen kooperativ und die Migration von Trittbrettfahrern eingeschränkt, greifen die Vorteile von Kooperation in vollem Umfang. Eine Abschottung ist beispielsweise möglich durch aktive Wahl der Interaktionspartner (Page et al. 2005), Abwehrmaßnahmen, räumliche Trennung von Gruppen durch Territorialität, geografische Bedingungen oder rein stochastische Prozesse.

12. Diese Arbeit macht plausibel, dass Investitionen in ein Gemeingut nicht notwendigerweise altruistisch oder mutualistisch motiviert sein müssen, sondern auch aus egoistischen Motiven entstehen können. Damit liegt eine proximate Erklärung vor, wie die hohe menschliche Kooperationsbereitschaft in solchen Situationen ohne Rückgriff auf Verwandtschaft oder teure Signale zustande kommen könnte.

Ein Hauptergebnis dieser Arbeit ist die Erkenntnis, dass die Investition in ein Gemeingut auch egoistisch motiviert sein kann, da ein einmal erzeugtes Gemeingut auf lange Sicht eben gewisse Vorteile bietet – im Gegensatz zu nicht-kooperierenden Gruppen, die keines erzeugen. Das stellt auf die Einsicht ab, dass auch viele Gemeingüter lokal begrenzt sind. Dabei kommt es entscheidend auf die Kosten, Effizienzvorteile und Ausschlussmöglichkeiten an, die das Gemeingut mit sich bringt.

Die Bestätigung der dafür entscheidenden Hypothesen (Studie 3, Hypothesen 1 und 2) ist daher in Einklang mit dem oben entworfenen Erklärungsansatz (B 1 und D 3): Investition in ein Gemeingut muss nicht immer altruistisch oder mutualistisch, sondern könnte sogar egoistisch motiviert sein. Unter gewissen Randbedingungen kann es sich lohnen, in Gemeingüter zu investieren. Ich habe ebenso gezeigt, dass diese Strategie nicht zum Erfolg führt, wenn Trittbrettfahrer nicht abgewehrt werden können, als auch, dass sie erfolgreich ist, wenn das gelingt. Wichtig für die erstmalige Entstehung ist dabei, dass sich solche Umweltbedingungen auch schon durch *zufällige* Gruppenaufteilungen ergeben können. Dem steht nicht entgegen,

dass ein solches Verhalten über inklusive Fitnessvorteile bei Verwandten entstanden sein kann.

Damit komme ich zu den allgemeinen Forschungsheuristiken, die sich aus meiner Sicht von den durchgeführten spezielleren Analysen ableiten lassen.

Erstens: Selbst sehr „robuste“ Ergebnisse aus Laborexperimenten, beispielsweise die schlechte Effizienz von Strafen, können sich nach kleinen Modifikationen als nicht haltbar erweisen, wie es etwa die Erweiterung von 10 auf 50 Runden in PGG zeigt (Gächter et al. 2008). Diese Arbeit schließt sich daher im Großen und Ganzen der oft geäußerten Kritik an Laborexperimenten als zu artifiziell an. Es gibt zwar viele Resultate, die durch externe Überprüfungen bestätigt werden (siehe oben, Spielertypen), das entscheidende Problem liegt jedoch darin, vorher nicht wissen zu können, welche der Resultate denn nun Artefakte sind und welche nicht. Um das entscheiden zu können, müssen andere Experimental-Paradigmen (z.B. Online-Spiele oder realweltliche Experimentalsituationen) stets ergänzend hinzugezogen werden.

Diese Warnung vor Verallgemeinerungen zeigt sich auch in dieser Arbeit. Ich habe immer vorsichtshalber mit zwei (fünf Labor-PGG und Ikariam) bzw. mehreren (Abschnitt D 2) Datensätzen oder Bestimmungsmethoden (Abschnitt D 1) gearbeitet. Trotzdem weichen Einzelergebnisse bei identischer Problemstruktur teilweise stark voneinander ab. Die Notwendigkeit, bei solchen Divergenzen eine genaue Analyse der Gründe vorzunehmen, versteht sich zwar von selbst, kommt aber erst durch ein solches redundantes Forschungsdesign in den Blick. Auch in anderen Experimenten zeigt bereits die Veränderung eines einzigen Wortes in den Instruktionen (Yamagishi et al. 2007) große Wirkung, was als Einflussfaktor für Ergebnisse extrem leicht übersehen werden kann, falls es nicht wie hier von Yamagishi et al.(2007) gezielt manipuliert wird.

Zweitens: Studien zur Kooperation müssen die komplexen Wechselwirkungen, die durch das Aufeinandertreffen verschiedener Strategien entstehen, mit in Betracht ziehen. Meines Wissens wurden bzw. werden *public goods games* im Labor, sowohl was Hypothesen, Design als auch Auswertung betrifft, standardmäßig ohne Berücksichtigung verschiedener Spielertypen durchgeführt. Selbst wenn sich alle Detailergebnisse in dieser Arbeit als falsch erweisen sollten, sollte doch deutlich geworden sein, dass die *Gruppenzusammensetzungen*, die sich wiederum durch die Verteilung *verschiedener Spielertypen* ergeben, von entscheidender Bedeutung für das Kooperationsniveau und den Erfolg verschiedener Strategien sind.

Drittens: Es ist extrem schwierig, für ein bestimmtes Verhalten einen evolutionären Mechanismus zweifelsfrei verantwortlich zu machen (vgl. Voland 2013). Das verhindern die enormen Schwierigkeiten, der Währung, in der gerechnet werden muss (lebenslange Fitnessbilanz (West et al. 2011), habhaft zu werden. In Hinblick auf meine Hypothese, dass Investitionen in Gemeingüter auch egoistisch motiviert sein könnten, schließt diese Arbeit mit dem dazu passenden Zitat von Christian Vogel: „Der wahre Egoist kooperiert!“ (Vogel 1993).

3 Offene Fragen

Schließlich komme ich noch auf offene Fragen zu sprechen. Bislang nicht erforscht ist beispielsweise die Dynamik in Gruppen *homogener* Spielertypen, also beispielsweise Gruppen ausschließlich Kooperierender. Auch wenn diese in der Realität selten vorkommen, können solche homogenen Gruppen doch durch institutionelle Anreize oder Rahmenbedingungen erzeugt werden. Was passiert in solche Gruppen? Denkbar wäre, dass Spieler ihre bisherigen Strategien weiterfahren, ohne sie in Abhängigkeit der nunmehr homogenen Gruppenzusammensetzung zu modifizieren. Denkbar ist aber auch, dass einige Spieler ihre Strategie in einem solchen kooperativen Umfeld zu Trittbrettfahrerverhalten verändern, da sich der Nutzen dieser Strategien abhängig von einem derartigen Umfeld stark erhöhen würde. Für die erste Überlegung sprechen Befunde, die ein extrem hohes Niveau an Einzahlungen bei Gruppen homogener konditionaler Kooperierenden feststellen (Burlando & Guala 2005), für die zweite der Hinweis auf allgemeine evolutionäre Dynamiken (Semmann et al. 2003). Eine weitere damit verbundene Fragestellung, die aber meines Wissens bislang noch nicht weiterverfolgt wurde, ist, ob sich reine Trittbrettfahrergruppen aus ihrem dann unproduktiven Verhalten befreien können.

Ein zweiter offener Fragekomplex betrifft die Übertragung der Befunde und Überlegungen auf reale Gemeingutprobleme. Wie Zusammenarbeit bei der Erzeugung öffentlicher Güter von extrem niedrigen Transaktionskosten durch Online-Kommunikation profitieren kann, ist etwa Ziel des gerade gestarteten DFG-Graduiertenkollegs in Düsseldorf (<http://www.fortschrittskolleg.de/>). Wie Gemeinden, Städte oder Dörfer sich organisieren können, um über kollektives Handeln erfolgreich Energie-Autarkie zu erreichen, zeigen inzwischen zahlreiche Beispiele (Morbach bei Trier, Feldheim bei Berlin; Jühnde bei Göttingen, usw.). Eine systematische, wissenschaftliche Verbindung mit der Forschung zu Gemeingutproblemen ist allerdings noch nicht in Sicht.

Eine dritte offene Frage ist, wie starr bzw. flexibel menschliches kooperatives Verhalten tatsächlich ist. Die starke Modifikation von Verhalten (Güerck et al. 2006) im Experiment, die starken Unterschiede im Anteil der Spielertypen in verschiedenen Ländern (Kocher et al. 2008), der sehr hohe Anteil an konditional Kooperierenden (Fischbacher et al. 2001), und die variierenden Ultimatum- und Diktatorspiel-Offerten in fünfzehn Kulturen (Henrich et al. 2010) sprechen eher dafür, dass diese Heuristiken sehr flexibel eingesetzt werden können. Auch aus theoretischer Sicht geht die Erwartung zur Flexibilität; da die Kosten einer langfristigen „Kooperation“ mit einem Trittbrettfahrer immens sind, sollten Individuen in solchen Situationen auf einen Wechsel eingestellt sein. *Dagegen* könnten einerseits Bauplan-*constraints* oder die Kosten flexibler kognitiver Lösungen sprechen.

F Literaturverzeichnis

- Aktipis, C. A. 2011. Is cooperation viable in mobile organisms? Simple Walk Away rule favors the evolution of cooperation in groups. *Evolution and Human Behavior* 32(4): 263-276.
- Alexander, R. D. 1987. *The Biology of Moral Systems*. Berlin: De Gruyter.
- Anderies, J. M. and Janssen, M. A. 2013. *Sustaining the Commons*. Online verfügbar unter <https://csid.asu.edu/publications/sustaining-commons>, zuletzt geprüft am 26.03.2014.
- Andreoni, J. 1988. Why free ride? Strategies and learning in public goods experiments. *Journal of Public Economics* 37(3): 291-304.
- Andreoni, J., Harbaugh, W. and Vesterlund, L. 2003. The Carrot or the Stick: Rewards, Punishments and Cooperation. *American Economic Review* 93(3): 893-902.
- Axelrod, R. 1984/2000. *Die Evolution der Kooperation*. München: Oldenbourg.
- Axelrod, R. and Hamilton, W. 1981. The evolution of cooperation. *Science* 211(4489): 1390-1396.
- Balliet, D., Mulder, L. B. and Van Lange, P. A. 2011. Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin* 137(4): 594-615.
- Barclay, P. 2008. Enhanced recognition of defectors depends on their rarity. *Cognition* 107: 817-828.
- Barkow, J. H.; Cosmides, L.; Tooby, J. (Hg.). 1992. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford: Oxford University Press.
- Bell, R., Buchner, A. and Musch, J. 2010. Enhanced old-new recognition and source memory for faces of cooperators and defectors in a social-dilemma game. *Cognition* 117(3): 261-275.

- Bergmüller, R., Johnstone, R. A., Russell, A. F. and Bshary, R. 2007. Integrating cooperative breeding into theoretical concepts of cooperation. *Behavioural Processes* 76(2): 61-72.
- Bierhoff, H.-W. and Fetchenhauer, D. 2006. How to Explain Prosocial and Solidary Behavior: A Comparison of Framing Theory with Related Meta-Theoretical Paradigms. pp. 225-242 In: Fetchenhauer, D., Flache, A., Buunk, B. and Lindenberg, S., eds. *Solidarity and Prosocial Behavior: An Integration of Sociological and Psychological Perspectives*. Springer: New York.
- Binmore, K. 2006. Why do people cooperate? *Politics, Philosophy & Economics* 5(1): 81-96.
- Bochet, O., Page, T. and Putterman, L. 2006. Communication and punishment in contribution experiments. *Journal of Economic Behavior and Organization* 60(1): 11-26.
- Boesch, C. 2001. Cooperative Hunting Roles Among Tai Chimpanzees. *Human Nature* 13(1): 27-46.
- Bolton, G. E. and Ockenfels, A. 2000. ERC: A theory of equity, reciprocity, and competition 90(1): 166-193.
- Boomsma, J. J., Nielsen, J., Sundström, L., Oldham, N. J., Tentschert, J., Petersen, H. C. and Morgan, E. D. 2003. Informational constraints on optimal sex allocation in ants. *Proceedings of the National Academy of Sciences of the United States of America* 100(15): 8799-8804.
- Bortz, J. and Schuster, C. 2010. *Statistik für Human- und Sozialwissenschaftler*. 7., vollst. überarb. u. erw. Aufl. Berlin: Springer (Springer-Lehrbuch).
- Brandts, J., Saijo, T. and Schram, A. 2004. How universal is behavior? A four country comparison of spite and cooperation in voluntary contribution mechanisms. *Public Choice* 119: 381-424.
- Brockhurst, M. A., Buckling, A., Racey, D. and Gardner, A. 2008. Resource supply and the evolution of public-goods cooperation in bacteria. *BMC Biology* 6: 20-26.
- Brockhurst, M. A., Hochberg, M. E., Bell, T. and Buckling, A. 2006. Character Displacement Promotes Cooperation in Bacterial Biofilms. *Current Biology* 16(20): 2030-2034.
- Bshary, R. and Bergmüller, R. 2008. Distinguishing four fundamental approaches to the evolution of helping. *Journal of Evolutionary Biology* 21(2): 405-420.
- Bshary, R. and Bronstein, J. L. 2004. Game Structures in Mutualistic Interactions: What Can the Evidence Tell Us About the Kind of Models We Need?. pp. 59-101 In: *Advances in the Study of Behavior*, Bd. 34. Elsevier.
- Bshary, R. and Grutter, A. S. 2002. Asymmetric cheating opportunities and partner control in a cleaner fish mutualism. *Animal Behaviour* 63(3): 547-555.

- Bshary, R. and Grutter, A. S. 2005. Punishment and partner switching cause cooperative behaviour in a cleaning mutualism. *Biology Letters* 1(4): 396-399.
- Bshary, R. and Grutter, A. S. 2006. Image scoring and cooperation in a cleaner fish mutualism. *Nature* 441: 975-978.
- Bshary, R. and Schäffer, D. 2002. Choosy reef fish select cleaner fish that provide high-quality service. *Animal Behaviour* 63(3): 557-564.
- Buchan, N. R., Croson, R. T. and Dawes, R. M. 2002. Swift Neighbors and Persistent Strangers: A Cross-Cultural Investigation of Trust and Reciprocity in Social Exchange. *American Journal of Sociology* 108(1): 168-206.
- Buchan, N. R., Grimalda, G., Wilson, R., Brewer, M., Fatas, E. and Foddy, M. 2009. Globalization and human cooperation. *Proceedings of the National Academy of Sciences* 106(11): 4138-4142.
- Burlando, R. and Hey, J. D. 1997. Do Anglo-Saxons free-ride more? *Journal of Public Economics* 64: 41-60.
- Burlando, R. M. and Guala, F. 2005. Heterogeneous Agents in Public Goods Experiments. *Experimental Economics* 8(1): 35-54.
- Burnham, T. C. and Johnson, D. D. 2005. The Biological and Evolutionary Logic of Human Cooperation. *Analyse und Kritik* 27: 113-135.
- Buss, D. M. 2004. *Evolutionary psychology: The New Science of the Mind*. Boston: Pearson.
- Cardenas, J.-C., Stranlund, J. and Willis, C. 2000. Local Environmental Control and Institutional Crowding-Out. *World Development* 28(10): 1719-1733.
- Cardenas, S. 2000. How Do Groups Solve Local Commons Dilemmas? Lessons from Experimental Economics in the Field. *Environment, Development and Sustainability* 2(3-4): 305-322.
- Chaudhuri, A. 2011. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics* 14(1): 47-83. Online verfügbar unter [10.1007/s10683-010-9257-1](https://doi.org/10.1007/s10683-010-9257-1).
- Chaudhuri, A. and Paichayontvijit, T. 2006. Conditional cooperation and voluntary contributions to a public good. *Economics Bulletin* 3(8): 1-14.
- Chesney, T., Chuah, S.-H. and Hoffmann, R. 2009. Virtual world experimentation: An exploratory study. *Journal of Economic Behavior & Organization* 72(1): 618-635.
- Chhatre, A. and Agrawal, A. 2008. Forest commons and local enforcement. *Proceedings of the National Academy of Sciences* 105(36): 13286-13291.
- Cinner, J. E., McClanahan, T. R., MacNeil, M. A., Graham, N. A., Daw, T. M., Mukminin, A.

- et al. 2012. Comanagement of coral reef social-ecological systems. *Proceedings of the National Academy of Sciences* 109(14): 5219-5222.
- Clutton-Brock, T. 2002. Breeding together: kin selection and mutualism in cooperative vertebrates. *Science* 296(5565): 69-72.
- Clutton-Brock, T. H. and Parker, G. A. 1995. Punishment in animal societies. *Nature* 373(6511): 209-216.
- Cornwallis, C. K., West, S. A., Davis, K. E. and Griffin, A. S. 2010. Promiscuity and the evolutionary transition to complex societies. *Nature* 466(7309): 969-972.
- Cosmides, L. 1989. The Logic of Social Exchange: Has Natural Selection Shaped How Humans Reason? Studies with the Wason Selection Test. *Cognition* 31(3): 187-276.
- Cosmides, L. and Tooby, J. 1992. Cognitive Adaptations for Social Exchange. pp. 163-228 In: Barkow, J. H., Cosmides, L. and Tooby, J., eds. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford University Press: Oxford.
- Creel, S. and Creel, N. M. 1995. Communal hunting and pack size in African wild dogs, *Lycaon pictus*. *Animal Behaviour* 50: 1325-1339.
- Croson, R. T. and Marks, M. B. 2000. Step Returns in Threshold Public Goods: A Meta- and Experimental Analysis. *Experimental Economics* 2(3): 239-259.
- Dawkins, R. 1976. *The Selfish Gene*. Oxford: Oxford University Press.
- Dawkins, R. 1982/2008. *The Extended Phenotype: The Long Reach of the Gene*. Oxford: Oxford University Press.
- de Waal, F. B. 1989. Food sharing and reciprocal obligations among chimpanzees. *Journal of Human Evolution* 18(5): 433-459.
- Donlan, R. and Costerton, J. W. 2002. Biofilms: Survival Mechanisms of Clinically Relevant Microorganisms. *Clinical Microbiological Reviews* 15(2): 167-193.
- Drea, C. and Carter, A. N. 2009. Cooperative problem solving in a social carnivore. *Animal Behaviour* 78(4): 967-977.
- Ducheneaut, N., Yee, N., Nickell, E. and Moore, R. J. 2006. "Alone together?": Exploring the Social Dynamics of Massively Multiplayer Online Games In: Olson, G. and Jeffries, R., eds. *the SIGCHI conference*. Montreal, Canada. pp. 407-416.
- Dunbar, R. I. 1998. The Social Brain Hypothesis. *Evolutionary Anthropology* 6(5): 178-190.
- Dunbar, R. I. 2003. The Social Brain: Mind, Language, and Society in Evolutionary Perspective. *Annual Review of Anthropology* 32: 163-181.
- Egas, M. and Riedl, A. 2008. The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society B: Biological Sciences* 275: 871-878.

- Fehr, E. and Fischbacher, U. 2003. The nature of human altruism. *Nature* 425(6960): 785-791.
- Fehr, E., Fischbacher, U. and Gächter, S. 2002. Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature* 13(1): 1-25.
- Fehr, E. and Gächter, S. 2000. Cooperation and Punishment in Public Goods Experiments. *American Economic Review* 90(4): 980-994.
- Fehr, E. and Gächter, S. 2002. Altruistic punishment in humans. *Nature* 415(6868): 137-140.
- Fehr, E. and Schmidt, K. M. 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114(3): 817-868.
- Fetchenhauer, D. and Bierhoff, H.-W. 2004. Altruismus aus evolutionstheoretischer Perspektive. *Zeitschrift für Sozialpsychologie* 35(3): 131-141.
- Fetchenhauer, D. and Dunning, D. 2009. Do people trust too much or too little? *Journal of Economic Psychology* 30(3): 263-276.
- Fischbacher, U. and Gächter, S. 2010. Social preferences, beliefs, and the dynamics of free riding in public good experiments. *American Economic Review*(100): 541-556.
- Fischbacher, U., Gächter, S. and Fehr, E. 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters* 71(3): 397-404.
- Fischbacher, U., Gächter, S. and Quercia, S. 2012. The behavioral validity of the strategy method in public good experiments. *Journal of Economic Psychology* 33(4): 897-913.
- Fischer, J. 2010. Nothing to talk about: On the linguistic abilities of nonhuman primates (and some other animal species) In: Frey, U. J., Störmer, C. and Willführ, K. P., eds. *Homo Novus: A Human Without Illusions*. Springer: Heidelberg.
- Frank, S. A. 2003. Repression of competition and the evolution of cooperation. *Evolution* 57(4): 693-705.
- Frey, B. S. and Meier, S. 2004. Pro-social behavior in a natural setting. *Journal of Economic Behavior and Organization* 54(1): 65-88.
- Frey, U. J. 2007. *Der blinde Fleck - Kognitive Fehler in der Wissenschaft und ihre evolutionsbiologischen Grundlagen*. Heusenstamm: Ontos.
- Frey, U. J. 2010. Modern illusions of humankind. pp. 263-288 In: Frey, U. J., Störmer, C. and Willführ, K. P., eds. *Homo Novus: A Human Without Illusions*. Springer: Heidelberg.
- Frey, U. J. and Rusch, H. 2012. An evolutionary perspective on the long-term efficiency of costly punishment. *Biology & Philosophy* 27(6): 811-831.
- Frey, U. J. and Rusch, H. 2013. Introducing Artificial Neural Networks to the Analysis of Social-Ecological Systems. *Ecology and Society* 18(2): 40. Online verfügbar unter <http://www.ecologyandsociety.org/vol18/iss2/art40/>.

- Frey, U. J. and Rusch, H. 2014. Modelling ecological success of common pool resource systems using large datasets. *World Development* 59: 93-103.
- Gächter, S. 2007. Conditional Cooperation: Behavioral Regularities from the Lab and the Field and Their Policy Implications. pp. 19-50 In: Frey, B. S. and Stutzer, A., eds. *Economics and Psychology: A Promising New Cross-Disciplinary Field*. MIT Press: Cambridge, Massachusetts.
- Gächter, S. and Herrmann, B. 2009. Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1518): 791-806.
- Gächter, S., Herrmann, B. and Thöni, C. 2004. Trust, voluntary cooperation, and socio-economic background: survey and experimental evidence. *Journal of Economic Behavior and Organization* 55: 505-531.
- Gächter, S., Herrmann, B. and Thöni, C. 2010. Culture and cooperation. *Philosophical Transactions of The Royal Society of London: Biological Sciences* 365: 2651-2661.
- Gächter, S., Renner, E. and Sefton, M. 2008. The Long-Run Benefits of Punishment. *Science* 322: 1510.
- Gardner, A. and West, S. A. 2010. Greenbeards. *Evolution* 64(1): 25-38.
- Gibson, C. C., Williams, J. T. and Ostrom, E. 2005. Local Enforcement and Better Forests. *World Development* 33(2): 273-284.
- Gintis, H., Bowles, S., Boyd, R. and Fehr, E. 2003. Explaining altruistic behavior in humans. *Evolution and Human Behavior* 24(3): 153-172.
- Gomes, T. and Carvalho, A. A. 2011. The Pedagogical Potential of MMOG: An Exploratory Study Including Four Games and their Players. pp. 103-121 In: Cruz-Cunha, M. M., Varvalho, V. H. and Tavares, P., eds. *Computer Games as Educational and Management Tools*. IGI Global.
- Grafen, A. 1990. Do animals really recognize kin? *Animal Behaviour* 39(1): 42-54.
- Griggs, R. A. and Cox, J. R. 1982. The elusive thematic-materials effect in Wason's selection Task. *British Journal of Psychology* 73(3): 407-420.
- Gruber, E. 2008. Key Principles of Community-Based Natural Resource Management: A Synthesis and Interpretation of Identified Effective Approaches for Managing the Commons. *Environmental Management* 45: 52-66.
- Gutter, A. S. 1995. Relationship between cleaning rates and ectoparasite loads in coral reef fishes. *Marine Ecology Progress Series* 118: 51-58.
- Gutter, A. S. 1999. Cleaner fish really do clean. *Nature* 398: 672-673.

- Guala, F. 2012. Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences* 35(1): 1-15.
- Gürerk, Ö., Irlenbusch, B. and Rockenbach, B. 2006. The Competitive Advantage of Sanctioning Institutions. *Science* 312(5770): 108-111.
- Gurven, M. 2004. Reciprocal altruism and food sharing decisions among Hiwi and Ache hunter-gatherers. *Behavioral Ecology and Sociobiology* 56(4): 366-380. Online verfügbar unter 10.1007/s00265-004-0793-6.
- Gutiérrez, N. L., Hilborn, R. and Defeo, O. 2011. Leadership, social capital and incentives promote successful fisheries. *Nature* 470(7334): 386-389.
- Hagen, E. H. and Hammerstein, P. 2006. Game theory and human evolution: a critique of some recent interpretations of experimental games. *Theoretical population biology* 69(3): 339-348.
- Hamilton, W. D. 1964a. The Genetical Evolution of Social Behaviour. I. *Journal of Theoretical Biology* 7: 1-16.
- Hamilton, W. D. 1964b. The Genetical Evolution of Social Behaviour. II. *Journal of Theoretical Biology* 7: 17-52.
- Hatchwell, B. and Komdeur, J. 2000. Ecological constraints, life history traits and the evolution of cooperative breeding. *Animal Behaviour* 59(6): 1079-1086.
- Hawkes, K. 1993. Why Hunter-Gatherers Work: An Ancient Version of the Problem of Public Goods. *Current Anthropology* 34(4): 341-361.
- Hawkes, K. and Bird, R. B. 2002. Showing Off, Handicap Signaling, and the Evolution of Men's Work. *Evolutionary Anthropology* 11(2): 58-67.
- Heinsohn, R. and Packer, C. 1995. Complex cooperative strategies in group-territorial African lions. *Science* 269(5228): 1260-1262.
- Helanterä, H. and Sundström, L. 2007. Worker policing and nest mate recognition in the ant *Formica fusca*. *Behavioral Ecology and Sociobiology* 61(8): 1143-1149.
- Helbing, D. and Yu, W. 2009. The outbreak of cooperation among success-driven individuals under noisy conditions. *Proceedings of the National Academy of Sciences* 106(10): 3680-3685.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C. F., Fehr, E., Gintis, H. and McElreath, R. 2001. In Search of Homo economicus: Behavioral Experiments in 15 Small-Scale Societies. *American Economic Review* 91(2): 73-78.
- Henrich, J., Heine, S. J. and Norenzayan, A. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33: 61-135.

- Herrmann, B. and Thöni, C. 2009. Measuring conditional cooperation: a replication study in Russia. *Experimental Economics* 12(1): 87-92.
- Herrmann, B., Thöni, C. and Gächter, S. 2008. Antisocial Punishment Across Societies. *Science* 319: 1362-1367.
- Hess, C. and Ostrom, E. 2003. Ideas, Artifacts, and Facilities: Information as a Common-Pool Resource. *Law and Contemporary Problems* 66: 111-146.
- Hill, K. R. 2001. Altruistic cooperation during foraging by the Ache, and the evolved human predisposition to cooperate. *Human Nature* 13(1): 105-128.
- Hill, K. R. and Hurtado, M. A. 2009. Cooperative breeding in South American hunter-gatherers. *Proceedings of the Royal Society B: Biological Sciences* 276(1674): 3863-3870. Online verfügbar unter doi:10.1098/rspb.2009.1061.
- Hill, K. R., Walker, R. S., Bozicevic, M., Eder, J., Headland, T., Hewlett, B. et al. 2011. Co-Residence Patterns in Hunter-Gatherer Societies Show Unique Human Social Structure. *Science* 331(6022): 1286-1289.
- Horton, J. J., Rand, D. G. and Zeckhauser, R. J. 2011. The online laboratory: conducting experiments in a real labor market. *Experimental Economics* 14(3): 399-425.
- Isaac, R. M., McCue, K. F. and Plott, C. R. 1985. Public Goods Provision in an Experimental Environment. *Journal of Public Economics* 26: 51-74.
- Isaac, R. M. and Walker, J. M. 1988. Group size effects in public goods provision: The voluntary contributions mechanism. *Quarterly Journal of Economics* 103: 179-199.
- Isaac, R. M., Walker, J. M. and Williams, A. W. 1994. Group size and the voluntary provision of public goods: Experimental evidence utilizing large groups. *Journal of Public Economics* 54(1): 1-36.
- Jandér, K. C. and Herre, E. A. 2010. Host sanctions and pollinator cheating in the fig tree-fig wasp mutualism. *Proceedings. Biological sciences / The Royal Society* 277(1687): 1481-1488.
- Janssen, M. A., Holahan, R., Lee, A. and Ostrom, E. 2010. Lab Experiments for the Study of Social-Ecological Systems. *Science* 328(5978): 613-617.
- Johnstone, R. A. 1995. Sexual Selection, Honest Advertisement and the Handcap Principle: Reviewing the Evidence. *Biological Review* 70: 1-65.
- Keller, L. 1997. Indiscriminate altruism: unduly nice parents and siblings. *Trends in Ecology & Evolution* 12(3): 99-103.
- Kerr, B., Neuhauser, C., Bohannan, B. J. and Dean, A. M. 2006. Local migration promotes competitive restraint in a host-pathogen 'tragedy of the commons'. *Nature* 442: 75-78.

- Keser, C. and van Winden, F. 2000. Conditional Cooperation and Voluntary Contributions to Public Goods. *Scandinavian Journal of Economics* 102(1): 23-39.
- Kiers, E. T., Rousseau, R. A., West, S. A. and Denison, R. F. 2003. Host sanctions and the legume-rhizobium mutualism. *Nature* 425: 78-81.
- Kimmich, C. 2012. *Networks of Coordination and Conflict: Governing Electricity Transactions for Irrigation in South India*.
- Kocher, M. G., Cherry, T., Kroll, S., Netzer, R. J. and Sutter, M. 2008. Conditional cooperation on three continents. *Economics Letters* 101: 175-178.
- Komdeur, J. and Hatchwell, B. J. 1999. Kin recognition: function and mechanism in avian societies. *Trends in Ecology and Evolution* 14(6): 237-241.
- Koopmans, R. and Rebers, S. 2009. Collective action in culturally similar and dissimilar groups: an experiment on parochialism, conditional cooperation, and their linkages. *Evolution and Human Behavior* 30(3): 201-211.
- Kreft, J.-U. 2004. Biofilms promote altruism. *Microbiology* 150: 2751-2760.
- Kühberger, A. 1998. The Influence of Framing on Risky Decisions: A Meta-analysis. *Organizational Behaviour and Human Decision Processes* 75: 23-55.
- Kurzban, R. and Houser, D. 2001. Individual differences in cooperation in a circular public goods game. *European Journal of Personality* 15: S37-S52.
- Kurzban, R. and Houser, D. 2005. Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences* 102(5): 1803-1807.
- Ledyard, J. O. 1995. Public Goods: A Survey of Experimental Research. pp. 111-194 In: Kagel, J. H. and Roth, A. E., eds. *The Handbook of Experimental Economics*. Princeton University Press: Princeton.
- Lehmann, L. and Keller, L. 2006. The evolution of cooperation and altruism – a general framework and a classification of models. *Journal of Evolutionary Biology* 19: 1365-1376.
- Leigh, J. E. G. 2010. The evolution of mutualism. *Journal of Evolutionary Biology* 23(12): 2507-2528.
- Leimar, O. and Hammerstein, P. 2001. Evolution of cooperation through indirect reciprocity. *Proceedings of the Royal Society B: Biological Sciences* 268: 745-753.
- MacLean, R. C. and Gudelj, I. 2006. Resource competition and social conflict in experimental populations of yeast. *Nature* 441(7092): 498-501.
- Mares, R., Young, A. and Clutton-Brock, T. H. 2012. Individual contributions to territory

- defence in a cooperative breeder: weighing up the benefits and costs. *Proceedings of the Royal Society B: Biological Sciences* 279(1744): 3989-3995.
- Marlowe, F. W., Berbesque, J. C., Barrett, C., Bolyanatz, A., Gurven, M. and Tracer, D. 2011. The 'spiteful' origins of human cooperation. *Proceedings of the Royal Society B: Biological Sciences* 278(1715): 2159-2164.
- Masclet, D., Noussair, C. N., Tucker, S. and Villeval, M.-C. 2003. Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism. *American Economic Review* 93(1): 366-380.
- Mateo, J. M. 2002. Kin-recognition abilities and nepotism as a function of sociality. *Proceedings of the Royal Society B: Biological Sciences* 269(1492): 721-727.
- Melis, A. P., Hare, B. and Tomasello, M. 2006. Engineering cooperation in chimpanzees: tolerance constraints on cooperation. *Animal Behaviour* 72(2): 275-286.
- Milinski, M., Semmann, D. and Krambeck, H.-J. 2002. Reputation helps solve the 'tragedy of the commons'. *Nature* 415(6870): 424-426.
- Milinski, M., Semmann, D., Krambeck, H.-J. and Marotzke, J. 2006. Stabilizing the Earth's climate is not a losing game: Supporting evidence from public goods experiments. *Proceedings of the National Academy of Sciences* 103(11): 3994-3998.
- Mitchell, G. 2012. Revisiting Truth or Triviality: The External Validity of Research in the Psychological Laboratory. *Perspectives on Psychological Science* 7(2): 109-117.
- Mulder, R. A. and Langmore, N. E. 1993. Dominant males punish helpers for temporary defection in superb fairy-wrens. *Animal Behaviour* 45: 830-833.
- Neuhauser, C. and Fargione, J. E. 2004. A mutualism–parasitism continuum model and its application to plant–mycorrhizae interactions. *Ecological Modelling* 177(3-4): 337-352.
- Nikiforakis, N. 2008. Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics* 92(1): 91-112.
- Nikiforakis, N. and Normann, H.-T. 2008. A comparative statics analysis of punishment in public-good experiments. *Experimental Economics* 11(4): 358-369.
- Noë, R. 2006. Cooperation experiments: coordination through communication versus acting apart together. *Animal Behaviour* 71(1): 1-18.
- Noë, R., van Hooff, J. A. and Hammerstein, P. 2001. *Economics in nature: Social dilemmas and biological markets*. New York: Cambridge University Press.
- Nowak, M. A. 2006. Five Rules for the Evolution of Cooperation. *Science* 314: 1560-1563.
- Nowak, M. A. and Sigmund, K. 1998. Evolution of indirect reciprocity by image scoring. *Nature* 393: 573-577.

- Ockenfels, A. and Weimann, J. 1999. Types and patterns: an experimental East-West-German comparison of cooperation and solidarity. *Journal of Public Economics* 71: 275-287.
- Oda, R. 1997. Biased face recognition in the Prisoner's Dilemma Game. *Evolution and Human Behavior* 18(5): 309-315.
- Ostrom, E. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.
- Ostrom, E., Gardner, R. and Walker, J. M. 1994. *Rules, Games, and Common-Pool Resources*. Ann Arbor: University of Michigan Press.
- Page, T., Putterman, L. and Unel, B. 2005. Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry and Efficiency. *Economic Journal* 115(506): 1032-1053.
- Poteete, A. R. and Ostrom, E. 2008. Fifteen Years of Empirical Research on Collective Action in Natural Resource Management: Struggling to Build Large-N Databases Based on Qualitative Research. *World Development* 36(1): 176-195.
- Raihani, N. J., Thornton, A. and Bshary, R. 2012. Punishment and cooperation in nature. *Trends in Ecology & Evolution* 27(5): 288-295.
- Rainey, P. B. and Rainey, K. 2003. Evolution of cooperation and conflict in experimental bacterial populations. *Nature* 425(6953): 72-74.
- Rankin, D. J., Bargum, K. and Kokko, H. 2007. The tragedy of the commons in evolutionary biology. *Trends in Ecology and Evolution* 22(12): 643-651.
- Ratnieks, F. L., Foster, K. R. and Wenseleers, T. 2006. Conflict resolution in insect societies. *Annual review of entomology* 51: 581-608.
- Rockenbach, B. and Milinski, M. 2006. The efficient interaction of indirect reciprocity and costly punishment. *Nature* 444: 718-723.
- Russell, A. F. and Hatchwell, B. J. 2001. Experimental evidence for kin-biased helping in a cooperatively breeding vertebrate. *Proceedings of the Royal Society. Biological sciences* 268(1481): 2169-2174.
- Russell, Y. I., Call, J. and Dunbar, R. I. 2008. Image scoring in great apes. *Behavioural Processes* 78: 108-111.
- Rustagi, D., Engel, S. and Kosfeld, M. 2010. Conditional Cooperation and Costly Monitoring Explain Success in Forest Commons Management. *Science* 330(6006): 961-965.
- Sachs, J. L. and Simms, E. L. 2006. Pathways to mutualism breakdown. *Trends in Ecology & Evolution* 21(10): 585-592.
- Schino, G. 2007. Grooming and agonistic support: a meta-analysis of primate reciprocal altruism. *Behavioral Ecology* 18: 115-120.

- Schultheiss, D. 2007. *Long-term motivations to play MMOGs: A longitudinal study on motivations, experience and behavior*. Proceedings of DiGRA 2007 Conference. Online verfügbar unter http://danielschultheiss.de/pdf/Schultheiss_Long-term_motivations_to_play_MMOGs.pdf, zuletzt geprüft am 10.07.2014.
- Seed, A. M., Clayton, N. S. and Emery, N. J. 2008. Cooperative problem solving in rooks (*Corvus frugilegus*). *Proceedings. Biological sciences / The Royal Society* 275(1641): 1421-1429.
- Sefton, M., Shupp, R. and Walker, J. M. 2007. The Effect of Rewards and Sanctions in Provision of Public Goods. *Economic Inquiry* 45(4): 671-690.
- Semmann, D., Krambeck, H.-J. and Milinski, M. 2003. Volunteering leads to rock–paper–scissors dynamics in a public goods game. *Nature* 425(6956): 390-392.
- Seyfarth, R. M. and Cheney, D. L. 2003. Signalers and receivers in animal communication. *Annual Review of Psychology* 54: 145-173.
- Smith, E. A. and Bliege Bird, R. 2005. Costly Signaling and Cooperative Behavior In: Gintis, H., Bowles, S., Boyd, R. and Fehr, E., eds. *Moral sentiments and material interests: The foundations of cooperation in economic life*. MIT Press: Cambridge, Mass. (Economic learning and social evolution, 6).
- Soares, M. C., Bshary, R. and Côté, I. M. 2008. Does cleanerfish service quality depend on client value or choice options? *Animal Behaviour* 76(1): 123-130.
- Sutter, M., Haigner, S. and Kocher, M. G. 2008. *Choosing the carrot or the stick? – Endogenous institutional choice in social dilemma situations*. Online verfügbar unter http://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=res2006&paper_id=395, zuletzt geprüft am 08.07.2012.
- Tang, S. Y. 1989. *Institutions and Collective Action in Irrigation Systems*. Online verfügbar unter <http://dlc.dlib.indiana.edu/dlc/handle/10535/3596>, zuletzt geprüft am 30.09.2011.
- Tebbich, S., Bshary, R. and Grutter, A. S. 2002. Cleaner fish *Labroides dimidiatus* recognise familiar clients. *Animal Cognition* 5(3): 139-145.
- Tinbergen, N. 1963. On aims and methods of ethology. *Zeitschrift für Tierpsychologie* 20: 410-433.
- Trivers, R. L. 1971. The Evolution of Reciprocal Altruism. *Quarterly Review of Biology* 46(1): 35-57.
- Trivers, R. L. 1974. Parent-Offspring Conflict. *American Zoologist* 14(1): 249-264.
- Uhl, M. and Voland, E. 2002. *Angeber haben mehr vom Leben*. Heidelberg: Spektrum, Akademischer Verlag.

- van Laerhoven, F. 2010. Governing community forests and the challenge of solving two-level collective action dilemmas—A large-N perspective. *Global Environmental Change* 20(3): 539-546.
- Vogel, C. 1993. Der wahre Egoist kooperiert. Ethische Probleme im Bereich der Evolutionsbiologie, Verhaltensforschung und Sozialbiologie. *Universitas* 48(1): 25-38.
- Voland, E. 2003. Eigennutz und Solidarität – Das konstruktive Potenzial biologisch evolvierter Kooperationsstrategien im Globalisierungsprozess. *Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik* 26(4): 15-20.
- Voland, E. 2007. Wir erkennen uns als den anderen ähnlich: Die biologische Evolution der Freiheitsintuition. *Deutsche Zeitschrift für Philosophie* 55(5): 739-749.
- Voland, E. 2013. *Soziobiologie – Die Evolution von Kooperation und Konkurrenz*. Heidelberg: Springer.
- Wason, P. C. 1968. Reasoning about a Rule. *Quarterly Journal of Experimental Psychology* 20: 273-281.
- Wedekind, C. and Milinski, M. 2000. Cooperation Through Image Scoring in Humans. *Science* 288(5467): 850-852.
- Wenseleers, T., Helanterä, A., Hart, A. and Ratnieks, F. L. 2004. Worker reproduction and policing in insect societies: an ESS analysis. *Journal of Evolutionary Biology* 17: 1035-1047.
- West, S. A., El Mouden, C. and Gardner, A. 2011. Sixteen common misconceptions about the evolution of cooperation in humans. *Evolution and Human Behavior* 32(4): 231-262, zuletzt geprüft am 21.07.2011.
- West, S. A., Griffin, A. S. and Gardner, A. 2007. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology* 20(2): 415-432.
- Wiessner, P. 2009. Experimental Games and Games of Life among the Ju/'hoan Bushmen. *Current Anthropology* 50(1): 133-138.
- Willems, E. P., Hellriegel, B. and van Schaik, C. P. 2013. The collective action problem in primate territory economics. *Proceedings of the Royal Society B: Biological Sciences* 280(1759): 20130081.
- Wilson, E. O. 2013. *The social conquest of Earth*. New York, London: Liveright Publishing Corp.
- Yamagishi, T., Terai, S., Kiyonari, T., Mifune, N. and Kanazawa, S. 2007. The Social Exchange Heuristic: Managing Errors in Social Exchange. *Rationality and Society*

19(3): 259-291.

Young, A. J., Carlson, A. A., Monfort, S. L., Russell, A. F., Bennett, N. C. and Clutton-Brock, T. 2006. Stress and the suppression of subordinate reproduction in cooperatively breeding meerkats. *Proceedings of the National Academy of Sciences of the United States of America* 103(32): 12005-12010.

Zahavi, A. and Zahavi, A. 1997. *The Handicap Principle: A Missing Piece of Darwin's Puzzle*. Oxford: Oxford University Press.

Zefferman, M. R. 2014. Direct reciprocity under uncertainty does not explain one-shot cooperation, but demonstrates the benefits of a norm psychology. *Evolution and Human Behavior* 35(5): 358-367.

Zelmer, J. 2003. Linear Public Goods Experiments: A Meta-Analysis. *Experimental Economics* 6(3): 299-310.

G Anhang

1 SQL-Skripte

Um von den Rohdaten bei Ikariam zu verwendbaren Daten zu kommen, waren umfangreiche Umkodierungen in SQL (<http://dev.mysql.com/doc/refman/5.1/en/>) erforderlich. Diese können hier nicht aufgelistet werden, da ihr Umfang mehrere tausend Seiten weit übersteigt. Sie sind auf Nachfrage erhältlich. Um trotzdem zwei zentrale Analysen im Ansatz sichtbar und nachvollziehbar zu machen, sind im Folgenden zwei Beispielberechnungen angegeben, einmal für das Erfolgsmaß für Spielertypen (Skript 1) und die Berechnung des Erfolges der Spielertypen pro Gruppentyp (Skript 2).

Es ist jeweils nur das Skript für *einen* Snapshot für *ein* Land angegeben – tatsächlich muss jedes unten angegebene Skript für jedes Land 28-mal (für jedes Datum) durchgeführt werden. Die Speicherung der Werte findet in den mit @ gekennzeichneten Variablen statt, die mit einem weiteren Skript (hier nicht gezeigt) ausgelesen werden.

Softwareseitig wird die Sprache MySQL (<http://dev.mysql.com/>), als Datenbank MySQL-Server 5.6 und als Oberfläche HeidiSQL 8.3.0 (<http://www.heidisql.com/>) verwendet.

1.1 Skript 1

```
USE de_2504;SELECT @Alt_de_2504 := (SELECT  
(AVG(c_Rathauslev)/AVG(a_Spieldauer)) FROM avi_ds WHERE  
(d_Proz_Don_von_Res_Ges >= 50) AND (i_Anz_Staedte_pro_Insel > 1)); SELECT
```

```

@Free_de_2504 := (SELECT (AVG(c_Rathauslev)/AVG(a_Spieldauer)) FROM avi_ds
WHERE (d_Proz_Don_von_Res_Ges = 0) AND (i_Anz_Staedte_pro_Insel > 1)); SELECT
@Normal_de_2504_N1 := (SELECT (AVG(c_Rathauslev)/AVG(a_Spieldauer)) FROM avi_ds
WHERE (d_Proz_Don_von_Res_Ges < 10 AND d_Proz_Don_von_Res_Ges > 0) AND
(i_Anz_Staedte_pro_Insel > 1)); SELECT @Normal_de_2504_N2 := (SELECT
(AVG(c_Rathauslev)/AVG(a_Spieldauer)) FROM avi_ds WHERE
(d_Proz_Don_von_Res_Ges < 20 AND d_Proz_Don_von_Res_Ges > 9) AND
(i_Anz_Staedte_pro_Insel > 1)); SELECT @Normal_de_2504_N3 := (SELECT
(AVG(c_Rathauslev)/AVG(a_Spieldauer)) FROM avi_ds WHERE
(d_Proz_Don_von_Res_Ges < 30 AND d_Proz_Don_von_Res_Ges > 19) AND
(i_Anz_Staedte_pro_Insel > 1)); SELECT @Normal_de_2504_N4 := (SELECT
(AVG(c_Rathauslev)/AVG(a_Spieldauer)) FROM avi_ds WHERE
(d_Proz_Don_von_Res_Ges < 40 AND d_Proz_Don_von_Res_Ges > 29) AND
(i_Anz_Staedte_pro_Insel > 1)); SELECT @Normal_de_2504_N5 := (SELECT
(AVG(c_Rathauslev)/AVG(a_Spieldauer)) FROM avi_ds WHERE
(d_Proz_Don_von_Res_Ges < 50 AND d_Proz_Don_von_Res_Ges > 39) AND
(i_Anz_Staedte_pro_Insel > 1));

```

1.2 Skript 2

```

USE tr_1302; SELECT (c_Rathauslev / a_Spieldauer), ((c_Rathauslev / a_Spieldauer) /
i_Anz_staedte_pro_insel) FROM avi_ds WHERE d_Proz_Don_von_Res_Ges >= 50 AND
ANF_type_of_island = "1" AND i_Anz_staedte_pro_insel > 1 INTO OUTFILE
'D:\\Forschung\\Dissertation\\Ikariam\\200_Ikariam_Datenauswertung_A_RHL\\tr_1302__Sty
p_6_Gtyp_1.csv' FIELDS TERMINATED BY ',' OPTIONALLY ENCLOSED BY '"';

```

2 Internetspeicherorte der Rohdaten der verwendeten Laborexperimente

Fehr & Gächter 2000:

www.econ.uzh.ch/faculty/fehr/publications/AER2000ExpData.xls

Fehr & Gächter 2002:

www.econ.uzh.ch/faculty/fehr/publications/Fehr-Gaechter_Data_Nature2002.xls

Gächter et al. 2008:

Dieser Datensatz wurde mir auf Anfrage geschickt und ist nicht öffentlich zugänglich.

Nikiforakis 2008 und Nikiforakis & Normann 2008:

Die Daten sind inzwischen nicht mehr öffentlich verfügbar; der Download erfolgte von der Webseite von Nikos Nikiforakis im Dezember 2010.

3 Ikariam-Umfrage

Die durchgeführte Umfrage fand online statt. Ich verwende als Umfrage-Software Limesurvey (<http://www.limesurvey.org>), die Opensource ist. Die Umfrage wurde auf dem deutschen Ikariam-Portal beworben (<http://www.ikariam.de>). Sie lief vom 14. bis zum 30. März 2014 in deutscher Sprache. An der Umfrage beteiligten sich 429 Teilnehmer. 239 der 429 Teilnehmer (56%) füllten die Mehrzahl der 73 Fragen aus. Sie wurden für die weitere Datenanalyse ausgewählt. Zwei weitere Teilnehmer wurden ausgeschlossen, weil sie offensichtlich falsche Werte eintrugen.

Die Fragen beinhalten demographische Daten wie Geschlecht, Alter, Bildungsgrad, die Motivation für das Spielen von Ikariam, Kennzahlen für das Spiel wie Häufigkeit der Logins und Dauer des Spiels), den Kauf von Vorteilen im Spiel und Fragen über die Gemeingüter, wie etwa das theoretische Verständnis, nach welchen Kriterien sie in das Gemeinwohl investieren, den Umgang mit Trittbrettfahrern, wovon Investitionen abhängig gemacht werden, usw.

Beispiele:

1. Die Frage “Wie oft loggen sie sich ein, um Ikariam zu spielen?” wurde 207-mal beantwortet. Der Mittelwert lag bei 3,8.
2. Eine Textanalyse der freien Textfragen bezüglich des Verständnisses der Gemeingüter basiert vor allem auf den beiden folgenden Fragen:
 - a) “Warum investieren Sie in das Sägewerk / die Luxusminen / das Wunder?”
 - b) “ Bitte beschreiben Sie, wie Sie feststellen, ob jemand zu wenig in das Sägewerk / die Luxusminen / das Wunder investiert?”

Alle weiteren Fragen wurden nur zur weiteren Abklärung der analysierten Sachverhalte genutzt. Da sie nicht für eigene Analysen verwendet wurden, sind sie hier nicht angegeben.

4 Berechnung des Step return

Da es sich bei den Gemeingütern in Ikariam um ein *threshold public goods game* handelt, also

Gemeingüter, die dann erzeugt werden, wenn eine gewisse Grenze an Investitionen erreicht wird, kann eine wichtige Kennzahl für öffentliche Güter nicht berechnet werden (Croson & Marks 2000), der *marginal per capita return* (MPCR). Diese Größe drückt aus, wie stark es sich lohnt, in das Gemeingut zu investieren, das heißt wie viel pro eingesetzter Einheit wieder an den Investor zurückfließt. Ein analoges Konzept, das sich *step return* nennt, kann allerdings berechnet werden. Es ist der aggregierte Verdienst der gesamten Gruppe aus dem Gemeingut geteilt durch den Schwellenwert aller Einzahlungen (*total contribution threshold*, Croson & Marks 2000).

Im vorliegenden Fall ist der Verdienst abhängig von der Zeit, denn je früher der Schwellenwert erreicht wird, desto länger wird im Spielverlauf das öffentliche Gut bereitgestellt, so dass Ressourcen effizienter genutzt werden können. Daher erreicht der *step return* nach einigen Tagen bis Wochen einen Wert größer 1, nachdem die nächste Ausbaustufe erreicht wurde. Die Zeitdauer hängt von verschiedenen Faktoren ab, vor allem von der Ausbaustufe des Gemeinguts (höher = schwieriger), aber auch von verschiedenen Boni. Analog zu dem MPCR geht es bei dem *step return* um das Abgleichproblem, wie viele private Ressourcen in das Gemeingut transferiert werden sollen. Bis sich die Investition der Zahlung durch höheres stündliches Einkommen gelohnt hat, vergehen unterschiedliche Zeitspannen. Die unten stehende Tabelle 28 stellt diese Berechnungen im Detail an. Obwohl es jeweils eine Schwelle gibt, sind Investitionen in das Gemeingut keine Alles- oder-Nichts-Situation, in der es für Trittbrettfahrer rational sein könnte, den Schwellenwert durch eigene Investitionen zu erreichen. Investitionen gehen nicht verloren und akkumulieren sich über die Zeit, bis der Schwellenwert erreicht ist. Es ist also immer sinnvoll zu warten, bis andere die erforderlichen Einzahlungen bereitstellen.

Durch Variation der Spielerzahl, der Boni und der Stufen des Gemeinguts sind beliebig viele Kombinationen an *step return*-Berechnungen möglich. Für die Beispielrechnung nehme ich daher den Mittelwert der Spielerzahl auf einer Insel (8) und einen durchschnittlichen Bonus für die Effizienzverbesserung der Ernte (20%). Weitere Berechnungen, die hier nicht abgebildet sind, zeigen deutliche Unterschiede, je nach Spielerzahl und Boni in der Zeitspanne, die benötigt wird, damit sich Investitionen amortisieren.

Stufe des Gemeinguts	Gruppenverdienst	Schwellenwert für die Gesamtinvestitionen	Step Return	Tage, bis Gewinnschwelle erreicht ist
1	0	0	0	0
2	1843,2	394	4,678172589	0,21
3	2764,8	992	2,787096774	0,36
4	3225,6	1732	1,862355658	0,54
5	3686,4	2788	1,322238164	0,76
6	3686,4	3783	0,974464711	1,03
7	4147,2	5632	0,736363636	1,36
8	4608	8139	0,566162919	1,77
9	4608	10452	0,44087256	2,27
10	4608	13298	0,346518273	2,89
11	5068,8	18478	0,274315402	3,65
12	5068,8	23213	0,218360401	4,58
13	5068,8	29038	0,174557476	5,73
14	5529,6	39494	0,140011141	7,14
15	5529,6	49107	0,112603091	8,88
16	5990,4	66010	0,090749886	11,02
17	5990,4	81766	0,073262725	13,65
18	5990,4	101146	0,059225278	16,88
19	6451,2	134598	0,04792939	20,86
20	5990,4	154304	0,038822066	25,76
21	6451,2	205012	0,031467426	31,78
22	6912	270839	0,025520697	39,18
23	6451,2	311541	0,020707387	48,29
24	6912	411229	0,016808153	59,49
25	6912	506475	0,013647268	73,27
26	7372,8	665201	0,011083567	90,22
27	6912	767723	0,009003247	111,07
28	7372,8	1007959	0,007314583	136,71
29	7372,8	1240496	0,005943429	168,25
30	7372,8	1526516	0,004829822	207,05
31	7833,6	1995717	0,003925206	254,76
32	7372,8	2311042	0,003190249	313,46
33	7833,6	3020994	0,002593054	385,65
34	8294,4	3935195	0,002107748	474,44
35	7833,6	4572136	0,001713335	583,66
36	7833,6	5624478	0,001392769	717,99
37	8294,4	7325850	0,00113221	883,23
38	8294,4	9011590	0,000920415	1086,47
39	8294,4	11085051	0,000748251	1336,45
40	8294,4	13635408	0,000608299	1643,93
41	8755,2	17704143	0,000494528	2022,13
42	8294,4	20630781	0,00040204	2487,31
43	8755,2	26786470	0,000326852	3059,49
44	8755,2	32948197	0,000265726	3763,27
45	8755,2	40527121	0,000216033	4628,92
46	9216	52472840	0,000175634	5693,67

47	8755,2	61315353	0,00014279	7003,31
48	9216	79388129	0,000116088	8614,16
49	9216	97648282	0,000094380	10595,52
50	9216	120108270	0,000076731	13032,58

Tabelle 28. Beispielberechnung des step return für acht Spieler bei einem durchschnittlichen Bonus von 20%

5 Beschreibung der Variablen bei Ikariam

Die folgenden vier Tabellen (für Spieler (29), Städte (30), Gemeingüter (31) und Inseln (32)) geben die Datenstruktur der Rohdaten von Gameforge an.

Variablenname	Beschreibung
id	Player-ID
registration_time	Time of registration
gold	Amount of overall gold
formOfGovernment	Form of government
gender	Gender

Tabelle 29. Information über die Spieler (5 Variablen)

Variablenname	Beschreibung
id	City-ID
capital	Is this city the capital or not?
island_id	Island-ID
owner_id	Player-ID
citizens	Number of citizens
resource_workers	Number of workers in public goods 1
tradegood_workers	Number of workers in public goods 2
scientists	Number of scientists
level	Level of town hall
tradegood1	Amount of wine stored in this city
tradegood2	Amount of marble stored in this city
tradegood3	Amount of crystal glass stored in this city
tradegood4	Amount of sulphur stored in this city
resource	Amount of wood stored in this city
p1t	Type of building on that building plot
p1l	Level of building on that building plot
p2t	Type of building on that building plot
p2l	Level of building on that building plot
p3t	Type of building on that building plot
p3l	Level of building on that building plot
p4t	Type of building on that building plot
p4l	Level of building on that building plot
p5t	Type of building on that building plot
p5l	Level of building on that building plot
p6t	Type of building on that building plot
p6l	Level of building on that building plot
p7t	Type of building on that building plot
p7l	Level of building on that building plot
p8t	Type of building on that building plot
p8l	Level of building on that building plot
p9t	Type of building on that building plot
p9l	Level of building on that building plot
p10t	Type of building on that building plot
p10l	Level of building on that building plot
p11t	Type of building on that building plot
p11l	Level of building on that building plot
p12t	Type of building on that building plot
p12l	Level of building on that building plot
p13t	Type of building on that building plot
p13l	Level of building on that building plot
p14t	Type of building on that building plot
p14l	Level of building on that building plot
p15t	Type of building on that building plot
p15l	Level of building on that building plot
p16t	Type of building on that building plot
p16l	Level of building on that building plot
p17t	Type of building on that building plot
p17l	Level of building on that building plot
priests	Number of priests

Tabelle 30. Information über die Städte und deren Gebäude (49 Variablen)

Variablenname	Beschreibung
avatar_id	Player-ID
island_id	Island-ID
gold	Amount of gold invested in the public good
type	Type of public good (saw mill, luxury mine, miracle)

Tabelle 31. Information über Einzahlungen in die Gemeingüter (4 Variablen)

Variablenname	Beschreibung
id	Island-ID
wonder_type_id	Wonder-Type ID
tradegood	Type of luxury good
tradegood_level	Level of luxury public good
resource_level	Level of wood public good
city_count	Number of cities on this island
resource_donated	Investments so far in the wood public good
tradegood_donated	Investments so far in the luxury public good
wonder_donated	Investments so far in the miracle public good
wonder_level	Level of miracle public good
wonder_belief	Level of belief

Tabelle 32. Information über die Inseln (11 Variablen)

6 Ausbaustufen der drei Gemeingüter

Die folgenden drei Tabellen (33, 34 und 35) geben die Ausbaustufen der drei Gemeingüter (Sägewerk, Luxusminen, Wunder) an.

Stufe	Kosten der Stufe	Gesamtkosten	Maximum an einsetzbaren Arbeitern
1	0	0	30
2	394	394	38
3	992	1386	50
4	1732	3118	64
5	2788	5906	80
6	3783	9689	96
7	5632	15321	114
8	8139	23460	134
9	10452	33912	154
10	13298	47210	174
11	18478	65688	196
12	23213	88901	218
13	29038	117939	240
14	39494	157433	264
15	49107	206540	288
16	66010	272550	314
17	81766	354316	340
18	101146	455462	366
19	134598	590060	394
20	154304	744364	420
21	205012	949376	448
22	270839	1220215	478
23	311541	1531756	506
24	411229	1942985	536
25	506475	2449460	566
26	665201	3114661	598
27	767723	3882384	628
28	1007959	4890343	660
29	1240496	6130839	692
30	1526516	7657355	724
31	1995717	9653072	758
32	2311042	11964114	790
33	3020994	14985108	824
34	3935195	18920303	860
35	4572136	23492439	894
36	5624478	29116917	928
37	7325850	36442767	964
38	9011590	45454357	1000
39	11085051	56539408	1036
40	13635408	70174816	1072
41	17704143	87878959	1110
42	20630781	108509740	1146
43	26786470	135296210	1184
44	32948197	168244407	1222
45	40527121	208771528	1260
46	52472840	261244368	1300

47	61315353	322559721	1338
48	79388129	401947850	1378
49	97648282	499596132	1418
50	120108270	619704402	1458

Tabelle 33. Stufen des Gemeinguts Sägewerk

Stufe	Kosten der Stufe	Gesamtkosten	Maximum an einsetzbaren Arbeitern
1	0	0	20
2	1303	1303	32
3	2689	3992	48
4	4373	8365	66
5	7421	15786	88
6	10037	25823	110
7	13333	39156	132
8	20665	59821	158
9	26849	86670	184
10	37305	123975	212
11	47879	171854	240
12	65572	237426	270
13	89127	326553	302
14	106217	432770	332
15	152739	585509	366
16	193512	779021	400
17	244886	1023907	434
18	309618	1333525	468
19	414190	1747715	504
20	552058	2299773	542
21	660106	2959879	578
22	925396	3885275	618
23	1108885	4994160	656
24	1471979	6466139	696
25	1855942	8322081	736
26	2339735	10661816	776
27	3096779	13758595	818
28	3903252	17661847	860
29	5153666	22815513	904
30	6199765	29015278	946
31	8185063	37200341	990
32	10314552	47514893	1034
33	13588513	61103406	1080
34	17122961	78226367	1126
35	21576366	99802733	1172
36	27187657	126990390	1218
37	35747356	162737746	1266
38	45043166	207780912	1314
39	56755887	264536799	1362
40	71513915	336050714	1410
41	93863574	429914288	1460
42	118269663	548183951	1510
43	149021335	697205286	1560
44	187768443	884973729	1610
45	246053390	1131027119	1662
46	298104705	1429131824	1712

47	390638028	1819769852	1764
48	511136520	2330906372	1818
49	620180600	2951086972	1870
50	811484147	3762571119	1924

Tabelle 34. Stufen des Gemeinguts Luxusmine

Stufe	Kosten der Stufe	Gesamtkosten
1	0	0
2	8800	8800
3	31800	40600
4	73200	113800
5	135500	249300

Tabelle 35. Stufen des Gemeinguts Wunder

7 Weitere Abbildungen

Die folgenden Abbildungen dienen zur Vervollständigung der Abbildungen im Text.

Durchschnittlicher Erfolg (RHL/SD) pro Anzahl Städte für Deutschland

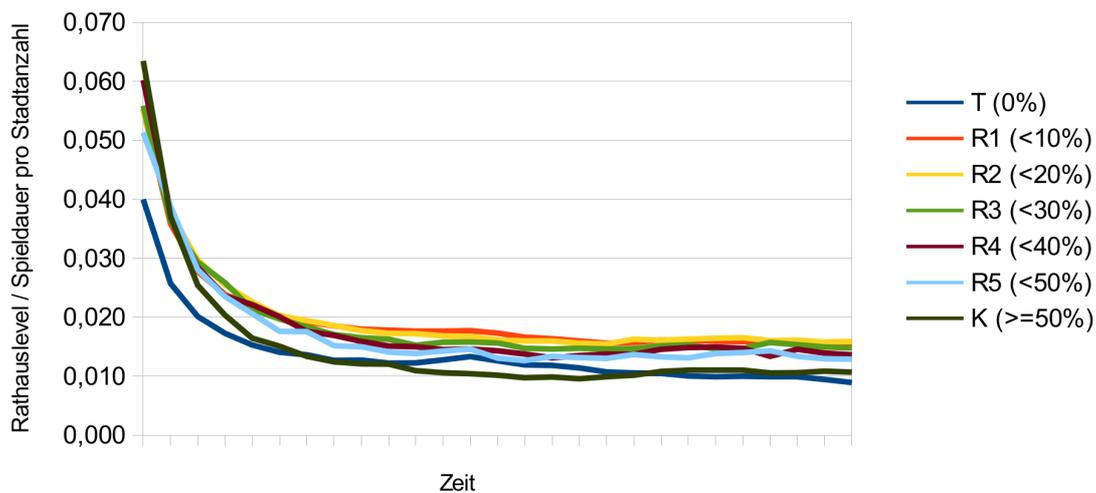


Abbildung 69: Durchschnittlicher Erfolg (RHL/SD) pro Stadtanzahl im Zeitverlauf für Deutschland ohne den ersten Snapshot

Durchschnittlicher Erfolg (RHL/SD) pro Anzahl Städte für England

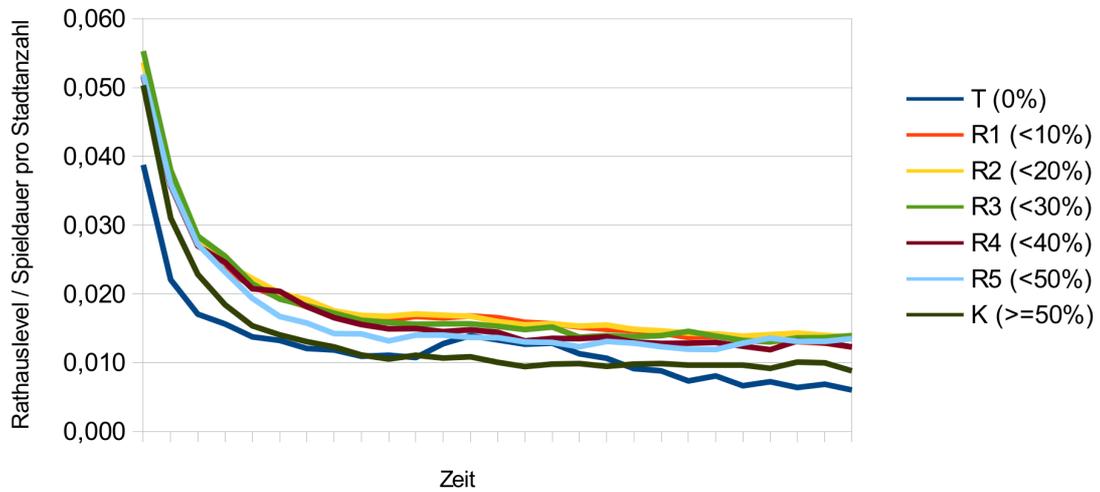


Abbildung 70: Durchschnittlicher Erfolg (RHL/SD) pro Stadtanzahl im Zeitverlauf für England ohne den ersten Snapshot

Durchschnittlicher Erfolg (RHL/SD) pro Anzahl Städte für Frankreich

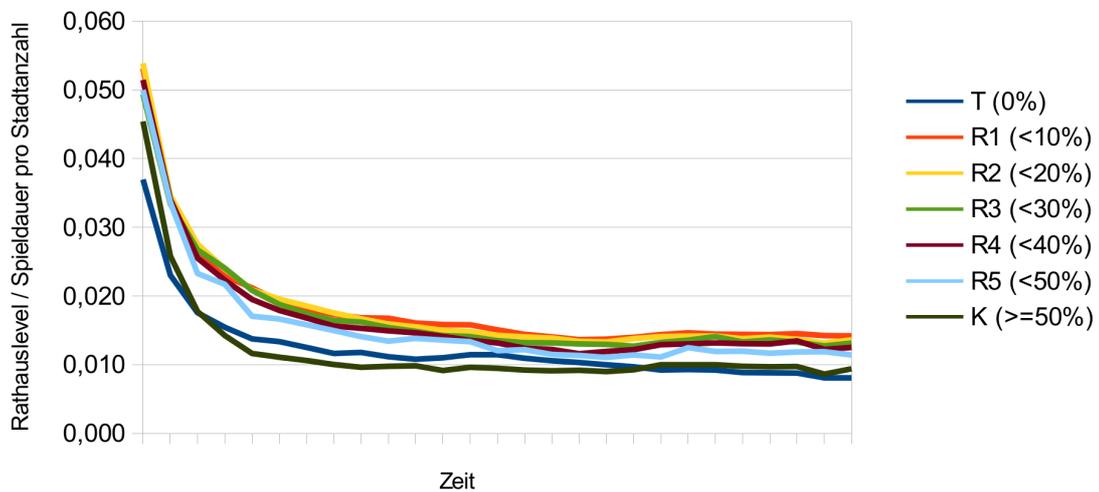


Abbildung 71: Durchschnittlicher Erfolg (RHL/SD) pro Stadtanzahl im Zeitverlauf für Frankreich ohne den ersten Snapshot

Durchschnittlicher Erfolg (RHL/SD) pro Anzahl Städte für Griechenland

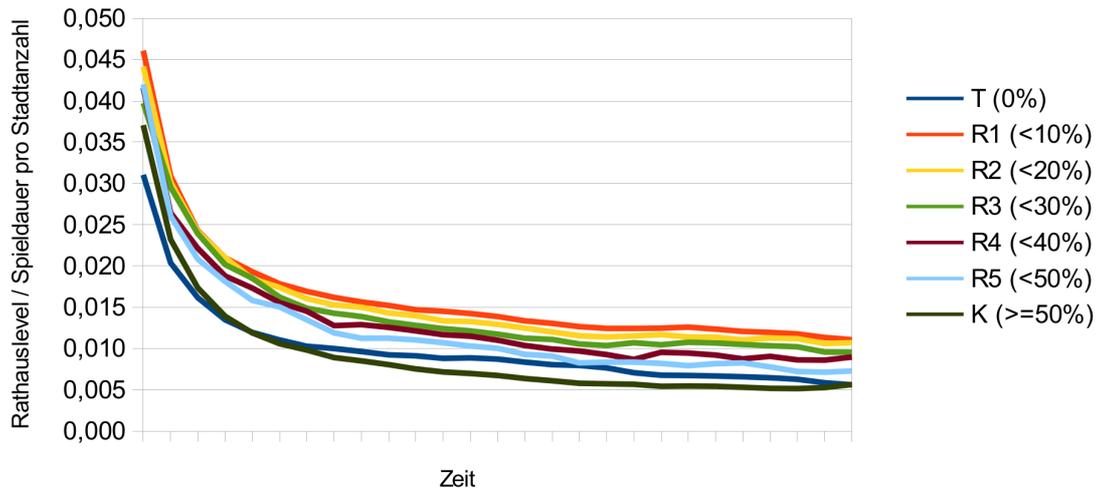


Abbildung 72: Durchschnittlicher Erfolg (RHL/SD) pro Stadtanzahl im Zeitverlauf für Griechenland ohne den ersten Snapshot

H Eigenständigkeitserklärung

Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.

Stuttgart, den 29.9.2015

(Ulrich Frey)