

# ȘCOALA NAȚIONALĂ DE STUDII POLITICE ȘI ADMINISTRATIVE

## REZUMAT

Rezonabilitate, reciprocitate și cooperare – O reinterpretare a teoriei rațiunii publice prin  
prisma teoriei evoluționiste a jocurilor

Coordonator științific:

**Prof. Dr. Adrian Miroiu**

Doctorand:

**Oana-Alexandra Dervis**

**2021**

## TABLE OF CONTENTS

|                                                                                                                                                                         |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Chapter I. Liberal democracies, reasonability, reciprocity, and cooperation.....                                                                                        | 8  |
| 1. Political liberalism and reasonable pluralism .....                                                                                                                  | 9  |
| 2. Individual interest, cooperation, and social norms.....                                                                                                              | 15 |
| 3. Some basic notions of game theory.....                                                                                                                               | 18 |
| 4. Evolutionary game theory and the emergence of cooperation.....                                                                                                       | 26 |
| 4.1 The evolution of contingent cooperation.....                                                                                                                        | 27 |
| 4.2 Group Selection.....                                                                                                                                                | 34 |
| 5. Conclusions .....                                                                                                                                                    | 41 |
| Chapter II. The evolutionary account of public reason.....                                                                                                              | 43 |
| 1. The ideal of public reason.....                                                                                                                                      | 44 |
| 2. Theories and practice: social morality.....                                                                                                                          | 47 |
| 3. The authority relation at the heart of social morality.....                                                                                                          | 49 |
| 4. Moral emotions – the presuppositions of social life.....                                                                                                             | 56 |
| 5. Actual moral reasoners .....                                                                                                                                         | 60 |
| 6. Actual public deliberation .....                                                                                                                                     | 63 |
| 7. Coordinating on a morality.....                                                                                                                                      | 65 |
| 8. Reasonable persons as rule-following punishers.....                                                                                                                  | 67 |
| 9. Punishment as a stabilization tool.....                                                                                                                              | 73 |
| Chapter III. The normative content of evolutionary public reason.....                                                                                                   | 78 |
| 1. Context and debates.....                                                                                                                                             | 79 |
| 2. The normative content of public reason.....                                                                                                                          | 83 |
| 2.1. What is at stake? Converging on a social rule and the normative content of public reason liberalism.....                                                           | 83 |
| 2.2. Does taking an evolutionary perspective on the emergence of social norms make us reconsider why reasonability and democratic interactions might be desirable?..... | 85 |

|                                                                                              |     |
|----------------------------------------------------------------------------------------------|-----|
| 3. Conclusions.....                                                                          | 113 |
| Chapter IV. The evolutionary interpretation of public reason: institutional design .....     | 114 |
| 1. “Even those who aim to change the world had better first learn how to describe it.” ..... | 115 |
| 2. A systematic look at the problems.....                                                    | 119 |
| 3. Empirical evidence in support of the implausibility charge.....                           | 123 |
| 4. A useful framework for understanding the conditions necessary for democratic change .     | 127 |
| 5. Norm change and deliberation.....                                                         | 134 |
| Concluding remarks.....                                                                      | 146 |
| References.....                                                                              | 151 |

## REZUMAT

În ultimii ani am asistat la o re-emergență a interesului pentru explicațiile care fac apel la diverse forme ale teoriei evoluționiste pentru a elucida fenomene sociale și politice. Doua exemple relevante sunt tradiția neo-instituționalista și cea a rațiunii publice, ambele beneficiind de dezvoltări teoretice din perspectiva teoriei evoluționiste a jocurilor. Deși nu este prima dată când argumente de acest fel sunt invocate în acest sens, anii '80 fiind de asemenea marcați de ceea ce în prezent sunt cunoscute ca dezbaterile sociobiologice (en. „the sociobiology wars”), noile modele propuse sunt mult mai sofisticate și propun soluționarea sau reinterpretarea unor probleme centrale cel puțin pentru anumite tradiții teoretice din științele politice.

Una dintre abordările recente care face uz de o explicație de natura evoluționistă este cea propusă de Gerald Gaus în cartea sa “The Order of Public Reason”, și supranumită interpretarea evoluționistă a raționării publice (en. „evolutionary public reason”). Gaus încearcă să dea o nouă interpretare justificării publice care să rezolve problemele asociate cu noțiunea de rezonabilitate aflată în centrul teoriei. Aceasta interpretare presupune prezentarea unui mecanism evoluționist de selecție a normelor sociale ca fiind responsabil pentru emergența cooperării la scară largă și propune o serie de argumente în favoarea acceptării rezultatului acestei selecții ca legitim dintr-o perspectivă liberală.

Scopul acestei lucrări a fost explorarea consecințelor reinterpretării teoriei rațiunii publice dintr-o perspectivă evoluționistă și investigarea modalităților prin care teoria evoluționistă a jocurilor ne permite să înțelegem interacțiunile din spațiul public liberal-democratic și poate chiar sa construim instituții mai democratice. În acest scop, am considerat ca este important răspunsul la următoarele patru întrebări: Care sunt provocările la care răspunde proiectul rațiunii publice? Cum poate fi folosită teoria evoluționistă a jocurilor pentru a rezolva o parte dintre problemele întâmpinate de acesta? Care sunt consecințele teoretice ale reinterpretării teoriei rațiunii publice din perspectivă evoluționistă? Există alte căi prin care rezultatele obținute de teoria evoluționistă a jocurilor pot fi folosite pentru a sprijini înțelegerea contemporană a proceselor democratice?

Pentru a înțelege relevanța acestei cercetări, cititorul trebuie să fie familiarizat cu problemele asociate noțiunii de justificare publică, și în particular proiectului liberal al rațiunii publice propus de John Rawls (1993). Acesta introduce noțiunea de rezonabilitate pentru a oferi o

soluție dificultăților de coordonare specifice democrațiilor pluraliste moderne. Rezonabilitatea este prezentată ca o caracteristică a indivizilor care le permite să rezolve probleme de cooperare pe care raționalitatea nu le-ar putea rezolva de sine stătătoare (deoarece raționalitatea este orientată doar către interesele individuale și poate opera doar pornind de la concepția individului cu privire la ceea ce este bun/dezirabil). Rawls propune ca soluție la problemele de coordonare din societățile moderne, probleme cauzate de pluralismul concepțiilor particulare despre bine (un fapt social care nu poate și nu trebuie schimbat), un dialog rezonabil despre principiile de bază ale societății, detașat de orice concepție particulară despre bine.

Există două critici majore la adresa acestei propuneri vizate de această lucrare. În primul rând, rezonabilitatea presupune posibilitatea de a sacrifica interesele individuale pe termen scurt în favoarea intereselor colective pe termen lung. Însă explicația pentru emergența și pentru viabilitatea unei astfel de trăsături în populațiile umane a fost mult timp o provocare, atât pentru teoreticienii politicului, cât și pentru biologi/antropologi. În al doilea rând, comportamentul rezonabil implică un grad ridicat de sofisticare psihologică, una dintre consecințele teoriei rawlsiene fiind excluderea unei majorități dintre cetățenii reali de la procesul de raționare publică. Însă aceste două probleme s-ar putea să devină mai ușor de abordat folosind câteva dezvoltări recente din teoria evoluționistă a jocurilor.

Rezultatele obținute folosind teoria evoluționistă a jocurilor de teoreticieni precum John Maynard Smith (1976), George Price (1972), Robert Trivers (1971), Robert Axelrod și William Hamilton (1981), dezvoltate în ultimii 40 de ani de mulți alți teoreticieni, oferă o soluție teoretică primei dintre problemele menționate (problemei emergenței cooperării pe termen lung în comunitățile umane). Jocul dilema prizonierului, inițial o dovadă a imposibilității apariției cooperării între indivizi raționali, a fost folosit de Axelrod și Hamilton pentru a demonstra că interacțiuni repetate în aceasta formulă pot duce, dacă se respectă anumite condiții, la emergența cooperării. De asemenea, celebra ecuație propusă de George Price a fost folosită pentru a demonstra că selecția în favoarea grupului poate, în anumite condiții, să fie mai puternică decât selecția în favoarea individului.

Concluziile generale derivate din aceste rezultate sunt că predispoziția spre cooperare în favoarea grupului (și orice altă trăsătură rentabilă pentru grup dar care defavorizează individul) poate fi încurajată de mecanismul selecției naturale cât timp: aceasta crește gradul de adaptabilitate

al grupului; grupurile se pot delimita în funcție de existența acestei predispoziții; capacitatea de a moșteni această trăsătură este corelată pozitiv cu apartenența la grup.

Implicațiile acestor descoperiri sunt relevante pentru felul în care ele vor fi folosite în reinterpretarea rațiunii publice. În primul rând, cooperarea susținută în situații de tip dilema prizonierului necesită dezvoltarea unui mecanism prin care defectorii pot fi excluși din grupul de indivizi care cooperează. De asemenea, în societățile umane, mecanismele de excludere a celor care nu se conformează normelor de cooperare sunt centrale pentru protejarea acestor norme, iar majoritatea indivizilor predispuși spre cooperare din comunitățile umane se folosesc de metode de identificare și pedepsire a defectorilor (afișând comportamente de tip „*rule-following punishment*”, cum sunt supranumite în literatura de specialitate) – altfel cooperarea la scară largă nu ar fi posibilă. O caracteristică importantă a modelelor folosite în domeniu este că ele pot fi folosite și pentru a explica evoluția culturală, nu doar pe cea biologică. În acest sens una dintre lucrările centrale din domeniu este *Culture and the Evolutionary Process* (Boyd, Richerson, 1985), care argumentează în favoarea considerării co-evoluției gene-cultură ca fiind maniera dominantă a evoluției umane.

Plecând de la aceste descoperiri, capitolele II și III ale lucrării discută consecințele teoretice ale reinterpretării teoriei rațiunii publice din perspectivă evoluționistă. După cum este menționat mai sus, în 2011, Gerald Gaus propune folosirea rezultatelor prezentate pentru a reinterpreta noțiunea de rezonabilitate în așa fel încât să ofere o rezolvare celor două probleme amintite anterior. Pentru început, o explicație a emergenței și condițiilor în care cooperarea poate fi o strategie stabilă din punct de vedere evoluționist ne poate ajuta să înțelegem care sunt condițiile în care comportamentul rezonabil este posibil (având în vedere că rezonabilitatea este tot o trăsătură a indivizilor care presupune sacrificarea/punerea între paranteze a intereselor individuale în favoarea intereselor publice). În al doilea rând, dacă este adevărat că toți indivizii care trăiesc în comunități umane au sau sunt încurajați să-și dezvolte capacitatea de a fi „*rule-following punishers*”, ar părea la prima vedere că nu mai este nevoie de raționare morală sofisticată pentru a rezolva problemele de coordonare specifice pluralismului – Gaus pretinde, bazându-se pe cercetările făcute de Boyd și Richerson (1985; 2005), că în istoria noastră evolutivă am dezvoltat așa-numite „emoții morale”, care ne pot ajuta să rezolvăm probleme ce țin de respectarea normelor. Între aceste emoții specifice capacitații de a fi „*rule-following punisher*” ar intra, printre altele, vina și indignarea.

La o privire mai atentă însă, substituirea rezonabilității cu un comportament de tip „*rule-following punishing*” are o serie de caracteristici mai puțin dezirabile din perspectivă normativă. Ca ele să devină aparente, trebuie să revenim la principiul liberal al legitimității, cel care stipulează exact care ar fi cerințele unui mecanism de identificare normelor morale și politice legitime la nivelul unei comunități:

*“Our exercise of political power is fully proper only when it is exercised in accordance with a constitution the essentials of which all citizens as **free and equal** may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason.”* (Rawls, 1993, p.137).

Gaus (2011) face trimitere la o serie de exemple și studii din psihologie pentru a demonstra că traiul în comun ne-a făcut să ne dezvoltăm anumite predispoziții emoționale morale compatibile cu tendința de a ne trata concetățenii conform cu principiul liberal al legitimității: felul în care tratăm persoanele care nu pot recunoaște validitatea justificărilor morale; felul în care tratăm copiii; felul în care ne schimbăm comportamentul când înțelegem că o regulă a fost încălcată din ignoranță, și nu din rea-voință. În toate aceste cazuri nu reacționăm emoțional sau violent, nu suntem indignați, ci înțelegători cu persoanele din exemple. Dar aceste exemple sunt, în cel mai rău caz, anecdotice, iar în cel mai bun caz, limitate doar la o fracțiune din interacțiunile din societate și, în orice caz, implică excluderea acestor persoane de la posibilitatea de a fi cetățeni egali. Se poate arăta că în istoria apropiată a umanității există nenumărate comportamente care implică excluderea unui individ sau a unei comunități doar pentru că nu recunoștea validitatea acelorași reguli ca majoritatea. De asemenea, revenind la implicațiile rezultatelor din teoria evoluționistă a jocurilor, aceleași predispoziții morale de natură emoțională care protejează stabilitatea unui sistem valoric pot fi folosite pentru a proteja o serie de norme mai mult decât chestionabile, deoarece mecanismele principale evoluționiste care le-au dat naștere sunt protejarea conformismului și excluderea indivizilor nonconformiști. Există, de asemenea, și inegalități structurale presupuse de unele dintre modelele evoluționiste care explică emergența cooperării. De regula, dacă admitem interacțiune non-atomistă în interiorul grupului și acceptăm că indivizii, datorită poziționării sociale, tind să interacționeze mai mult cu anumiți vecini decât cu alții, distribuția strategiilor aflate în echilibru în orice moment într-o societate va depinde de pozițiile pe care diverși indivizi le ocupă în societatea respectivă (Alexander, 2007). Într-o astfel de situație, cei aflați în poziții cheie beneficiază de mai multă libertate decât ceilalți membri ai societății. Din

aceste motive, chiar dacă concluziile despre emergența și stabilitatea cooperării obținute folosind teoria evoluționistă a jocurilor pot descrie care sunt mecanismele care încurajează cooperarea și respectarea normelor în societățile contemporane, ele nu trebuie să fie folosite pentru a justifica aranjamente care din punct de vedere normativ nu sunt în conformitate cu cerințele unei societăți liberale.

Însă această concluzie nu este finalul investigației, deoarece trebuie să ne întrebăm dacă nu cumva există alte căi prin care rezultatele obținute de teoria evoluționistă a jocurilor pot fi folosite pentru a sprijini înțelegerea contemporană a proceselor democratice. Aceasta este investigația căreia îi este dedicat ultimul capitol al lucrării.

Printre altele, teoria evoluționistă a jocurilor a fost folosită și pentru a explica fenomenul ignoranței pluraliste (en. „pluralistic ignorance”; Bicchieri 2006), un fenomen care depinde de inegalitățile create de structurile sociale menționate mai sus. Ignoranța pluralistă descrie situațiile în care majoritatea membrilor unui grup au o preferință în favoarea abolirii unei anumite norme, însă se simt obligați să se conformeze cu aceasta și să-i pedepsească pe cei care nu se conformează deoarece consideră că sunt singurii din grup care au această preferință - și se tem la rândul lor de consecințe. Exemplele din această categorie includ anumite norme religioase din comunități foarte conservatoare, unele norme din grupurile adolescenților, sau unele norme de gen din societățile contemporane. Acest comportament poate fi cauzat de lipsa accesului la informație sau de proeminența unei minorități cu o atitudine favorabilă pentru norma respectivă în poziții cheie din interiorul grupului.

Idealul rațiunii publice presupune crearea unui cadru în care astfel de fenomene nu influențează indivizii implicați în selectarea normelor unei societăți, prin urmare un pas către crearea unor instituții mai democratice ar presupune identificarea unor proceduri care minimizează efectul lor. În acest sens, încerc să investighez propunerea conform căreia deliberarea, în anumite condiții, poate duce la minimizarea efectului ignoranței pluraliste, aducând un argument în favoarea interpretării deliberative a rațiunii publice.

În urma studiilor de caz prezentate, se poate argumenta că deliberarea, dacă are loc într-un mediu propice, poate minimiza asimetria informațională din interiorul grupului sau poate disipa credințele false cu privire la motivațiile din spatele comportamentului altor indivizi. De asemenea, deliberarea poate oferi un cadru în care atitudinile se schimbă colectiv, prin urmare niciun individ nu trebuie să suporte un cost disproporționat ocupând singur poziția costisitoare de „trendsetter”.

Nu în ultimul rând, un cadru deliberativ creează premisele pentru interacțiune socială independentă de structurile încetățenite de interacțiune, în măsura în care fiecare persoană poate participa la dezbateri și argumentele sale sunt ascultate de toți cei prezenți – prin urmare, membri privilegiați din punct de vedere structural ai grupului nu mai pot exercita aceeași influență.

Concluziile cercetării sunt prezentate la sfârșitul fiecărui capitol, lucrarea fiind organizată sub forma mai multor articole de sine stătătoare. Însă în final ele pot fi rezumate de următoarele trei idei generale. În primul rând, deși teoria evoluționistă a jocurilor poate fi folosită cu succes pentru a explica emergența cooperării și pentru a investiga condițiile în care aceasta este posibilă, implicațiile pe care le are folosirea acestui cadru teoretic într-o teorie cu caracter normativ sunt multifacțate. În al doilea rând, încercarea lui Gerald Gaus de a aduce împreună cele două cadre teoretice nu trebuie subestimată, însă cercetările viitoare în domeniu ar trebui să acorde mai multă atenție incompatibilităților metodologice dintre cele două abordări și implicațiile descoperirilor pentru filosofia morală și politică. Nu în ultimul rând, deși poate nu are consecințele majore presupuse de G. Gaus, includerea modelelor evoluționiste de selecție a normelor în cercetarea instituțiilor democratice poate duce la descoperiri utile din punct de vedere al design-ului instituțional.

### ***Bibliografie***

- Alexander, J. M. 2007. *The Structural Evolution of Morality*. Cambridge University Press
- Axelrod, R., & W. D. Hamilton. 1981. "The evolution of cooperation". *Science* 211:1390– 1396.
- Bicchieri, C. 2006. *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press.
- Boyd, R. & P. J. Richerson. 1985. *Culture and the evolutionary process*. Chicago: University of Chicago Press.
- Boyd, R. & P. J. Richerson. 2005. *The Origin and Evolution of Cultures (Evolution and Cognition)*, Oxford University Press
- Gaus, G. 2011. *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*. Cambridge University Press
- Maynard Smith, J. 1976. "Group selection". *Quarterly Review of Biology* 51:277–283.
- Price, G. R. 1972. "Extension of selection covariance mathematics". *Annals of Human Genetics* 39: 455–458.
- Rawls, J. 1996[1993]. *Political Liberalism*. New York: Columbia University Press
- Trivers, R. L. 1971. "The evolution of reciprocal altruism". *Quarterly Review of Biology* 46:35–57.