



Altruïsme, coöperatie en reciprociteit

Jack Vromen

Erasmus Universiteit Rotterdam

Altruïsme, coöperatie en reciprociteit

- Soms worden altruïsme, coöperatie en reciprociteit aan elkaar gelijk gesteld (als zouden zij hetzelfde betekenen)

Mogelijke redenen:

1. Er is, zeker vanuit evolutietheorie bezien, overlap tussen deze begrippen
2. Vanuit speltheoretisch perspectief ligt het voor de hand om de begrippen aan elkaar te koppelen



Opbouw presentatie

1. Altruïsme, coöperatie en reciprociteit: een eerste terreinverkenning
2. Altruïsme bezien vanuit evolutionair perspectief
3. *Ultimate and proximate causes*
4. Speltheorie
5. Het Prisoners' Dilemma (PD)
6. Hoe zou altruïsme toch kunnen evolueren in een PD?
7. Conclusie



1. Een eerste begripsverkenning

Altruïsme:



onzelfzuchtig gedrag (tegendeel van egoïsme)

- Het belang van anderen boven dat van jezelf stellen? (Motief voor handelen)
- Hulp verlenen zonder verwachting van tegenbeloning (Handelen zelf)



1. Een eerste begripsverkenning

Coöperatie (samenwerking)



Er kunnen vele redenen zijn voor samenwerking:

- je denkt er zelf beter van te worden (bijv. je kunt iets niet in je eentje voor elkaar krijgen)
- je wilt anderen helpen
- je vindt samenwerken op zich prettig
- ...



1. Een eerste begripsverkenning

- *Reciprociteit* (wederkerigheid)



Balans/evenwicht tussen dienst en wederdienst

- Voor wat hoort wat
- *Quid pro quo*
- *I'll scratch your back if you scratch mine, Of: I won't scratch your back if you won't scratch mine*

Kortom: je helpt anderen niet, tenzij je een tegenprestatie verwacht



1. Een eerste begripsverkenning



(Alan Fiske: wederkerigheid kan verschillende vormen aannemen in verschillende soorten sociale relaties; bijv. vriendschap *versus* markttransacties)

1. Een eerste begripsverkenning

- Altruïsme, coöperatie en reciprociteit lijken dus niet hetzelfde te betekenen

Hun betekenissen overlappen elkaar hooguit gedeeltelijk

- Merk op dat ze alle drie begrepen worden in termen van motieven en intenties (redenen/oorzaken voor gedrag)



2. Altruïsme bezien vanuit evolutionair perspectief

- Maar in evolutionaire theorieën worden altruïsme, coöperatie en reciprociteit (meestal) anders, nl. in termen van feitelijk gedrag en feitelijke gevolgen daarvan begrepen
- *Evolutionair* altruïsme (niet *psychologisch* altruïsme, als hierboven):
Gedrag dat *fitness* van anderen vergroot ten koste van eigen *fitness*

Hier wordt niets gezegd over evt. motieven en intenties voor gedrag



2. Altruïsme bezien vanuit evolutionair perspectief

- Evolutionair probleem/puzzle van altruïsme:

Uit Darwins evolutieleer volgt dat individuen met relatief lage *fitness* reproductief niet succesvol zijn (d.w.z.: verdwijnen in populatie), hoe kan evolutionair altruïsme zich dan vestigen en handhaven in populatie?

- Oplossing moet eruit bestaan dat *fitness*-nadeel van altruïstisch gedrag *op de een of andere manier* gecompenseerd wordt



2. Altruïsme bezien vanuit evolutionair perspectief

- Voorbeeld: *kin selection* (Hamilton 1964); Hulp aan genetisch verwanten kan *inclusive fitness* vergroten



3. *Ultimate and proximate causes* (E. Mayr 1961)

- Waarom migreren trekvogels in het najaar naar het zuiden?



3. *Ultimate and proximate causes* (E. Mayr 1961)

1. Verklaring in termen van ultieme oorzaak: voorouders van trekvogels behaalden hierdoor evolutionair voordeel (natuurlijke selectie als ultieme oorzaak)
2. Verklaring in termen van proximale oorzaak: daling van temperatuur en/of verkorting van aantal zonen *triggers* trekgedrag in vogels (fysiologische mechanismen in vogels als proximale oorzaak)



3. *Ultimate and proximate causes* (E. Mayr 1961)

In het algemeen:

- ultieme oorzaak: evolutionair voordeel in het verleden
- proximale oorzaak: gedragsbepalende mechanismen in huidige individuen (plus omgevingsfactoren); mogelijk op verschillende niveau's van analyse (bijv. genetisch, neuraal, fysiologisch, mentaal – in termen van intenties en motieven bijv.)

Belang van onderscheid:

- Twee soorten verklaringen zijn complementair (niet rivaliserend) en beide nodig voor een compleet begrip



3. *Ultimate and proximate causes* (E. Mayr 1961)

Veel voorkomende denkfout:

- Omdat alles in evolutie is gericht op het behalen van evolutionair voordeel (ultieme oorzaak), kunnen alleen individuen die daar (intentioneel) op gericht zijn overleven (proximale oorzaak)
- Het is mogelijk dat individuen (of het soort gedrag dat ze tentoonspreiden) overleven die daar niet op gericht zijn (bijv. psychologische altruïsme)



3. *Ultimate and proximate causes* (E. Mayr 1961)

- Het evolutionaire probleem van evolutionair altruïsme is het probleem waarom en hoe de *ultimate cause* natuurlijke selectie evolutionair altruïsme niet heeft geëlimineerd

(wat zou de overlevingswaarde van evolutionair altruïsme geweest kunnen zijn?)

- Zegt in principe niets over *proximate causes* (of bijv. mensen psychologisch altruïstisch zijn)



3. *Ultimate and proximate causes* (E. Mayr 1961)

- Gangbare opvatting:
Direct gevolg van *ultimate cause* is niet gedrag, maar *proximate causes* van gedrag;
Gedrag als indirect gevolg

(Maar: hoe staat het met de verhouding tussen *ultimate* en *proximate causes* in *culturele* evolutie? Hier lijkt het onderscheid tussen beide te vervagen)

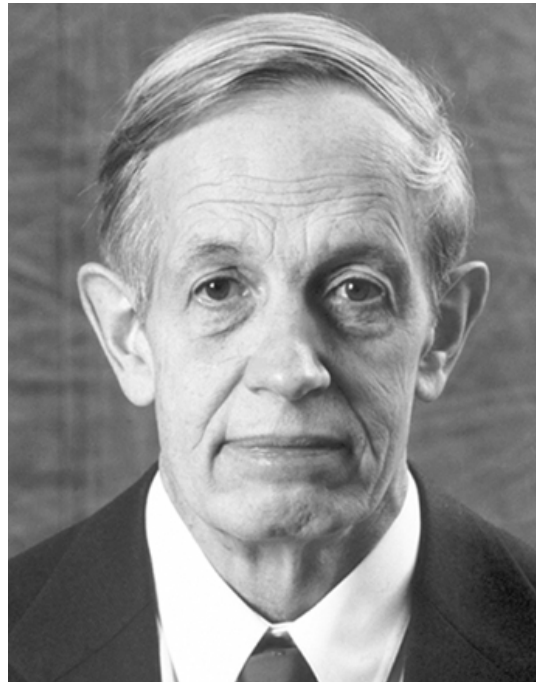


4. Speltheorie

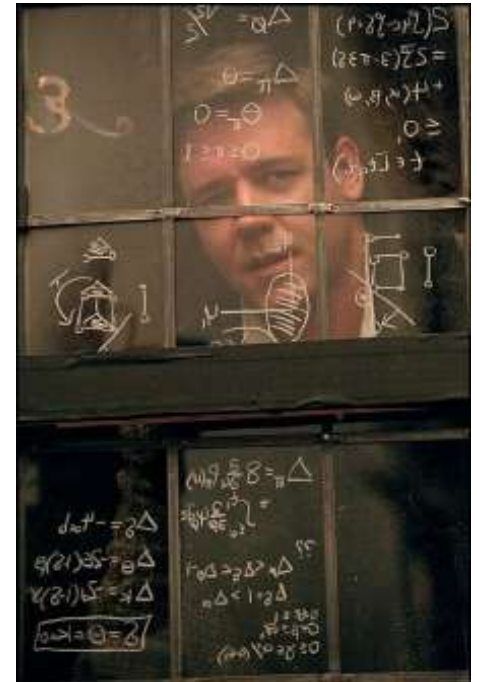
(Speltheorie is feitelijk een onderdeel van de wiskunde)



John von Neumann



John Nash



Russell Crowe als Nash

4. Speltheorie

- Evolutionair probleem kan treffend worden weergegeven in speltheorie, met name (maar niet alleen!) aan de hand van het *Prisoners' Dilemma* (PD)
- Eerst wat wetenswaardigheden over speltheorie



4. Speltheorie

- Grofweg zijn er twee soorten speltheorie:
 1. Klassieke 'rationalistische' speltheorie:
 - *payoffs* in termen van nut (omvat alles waar mensen waarde aan hechten)
 - individuen kiezen optie die hoogste nut oplevert
 - individuen worden geacht superrationeel en volledig geïnformeerd te zijn (*common knowledge*)



4. Speltheorie

2. Evolutionaire speltheorie

- *payoffs* in termen van *fitness*
- individuen zijn 'geprogrammeerd' om vaste strategie te spelen (populatie bestaat uit verschillende strategische types; wat evolueert is samenstelling van strategieën in populatie)
- individuen zijn hooguit beperkt rationeel en hebben hooguit beperkte informatie



4. Speltheorie

- Een interessant inzicht is wel dat (in veel gevallen; er zijn uitzonderingen!) evolutionaire processen (zoals gemodelleerd in evolutionaire speltheorie) convergeren naar uitkomsten die de superrationele individuen (uit de klassieke 'rationalistische' speltheorie) ook zouden kiezen: nl. naar een Nash-evenwicht



4. Speltheorie

- Nash evenwicht: niemand kan voordelen behalen door eenzijdig af te wijken
- Evolutionair stabiele strategie (ESS; Maynard Smith 1982): mutant-strategie kan populatie niet binnen dringen
(ESS is een verfijning van het Nash evenwicht)



4. Speltheorie

- De dynamische variant van de evolutionaire speltheorie:

Replicator dynamics: hoe de *fitness* van een strategie zich verhoudt tot de gemiddelde *fitness* van populatie bepaalt de groeivoet van frequentie van de strategie in de populatie

- Veronderstelling: dit is Darwiniaanse evolutie door natuurlijke selectie in zijn meest simpele, geïdealiseerde ('uitgeklede') vorm



4. Speltheorie

- *Evolutionary dynamics* relateren feitelijke gevolgen (relatieve *performance*) aan groeivoet van strategie (in populatie)

Zeggen dus (in principe) niets over redenen/beweegredenen/oorzaken (proximate causes) van gedrag

Zeggen alleen iets over *ultimate causes*



4. Speltheorie

Terug nu naar interessant inzicht:

- Er zijn twee mogelijke manieren waarop Nash-evenwicht kan worden bereikt :
 1. Rationeel redeneren en kiezen door individuen (klassieke speltheorie)
 2. Evolutionair proces (evolutionaire speltheorie)
- Maar... Nash evenwicht hoeft niet samen te vallen met optimale uitkomst!



5. Het Prisoners' Dilemma

- PD is het bekendste (maar zeker niet het enige!) spel in de speltheorie
- Het PD is vaak gebruikt om de mogelijkheden voor de evolutie van altruïsme te onderzoeken

Maar is het PD een *toy game*?



5. Het Prisoners' Dilemma

	C	D
C	(2,2)	(0,3)
D	(3,0)	(1,1)

In *one-shot* versie, ontkom je niet aan conclusie dat niet-samenwerken de dominante strategie is (het maakt niet uit wat anderen doen, niet-samenwerken is altijd de beste keuze); enige Nash evenwicht is (D,D)



5. Het Prisoners' Dilemma

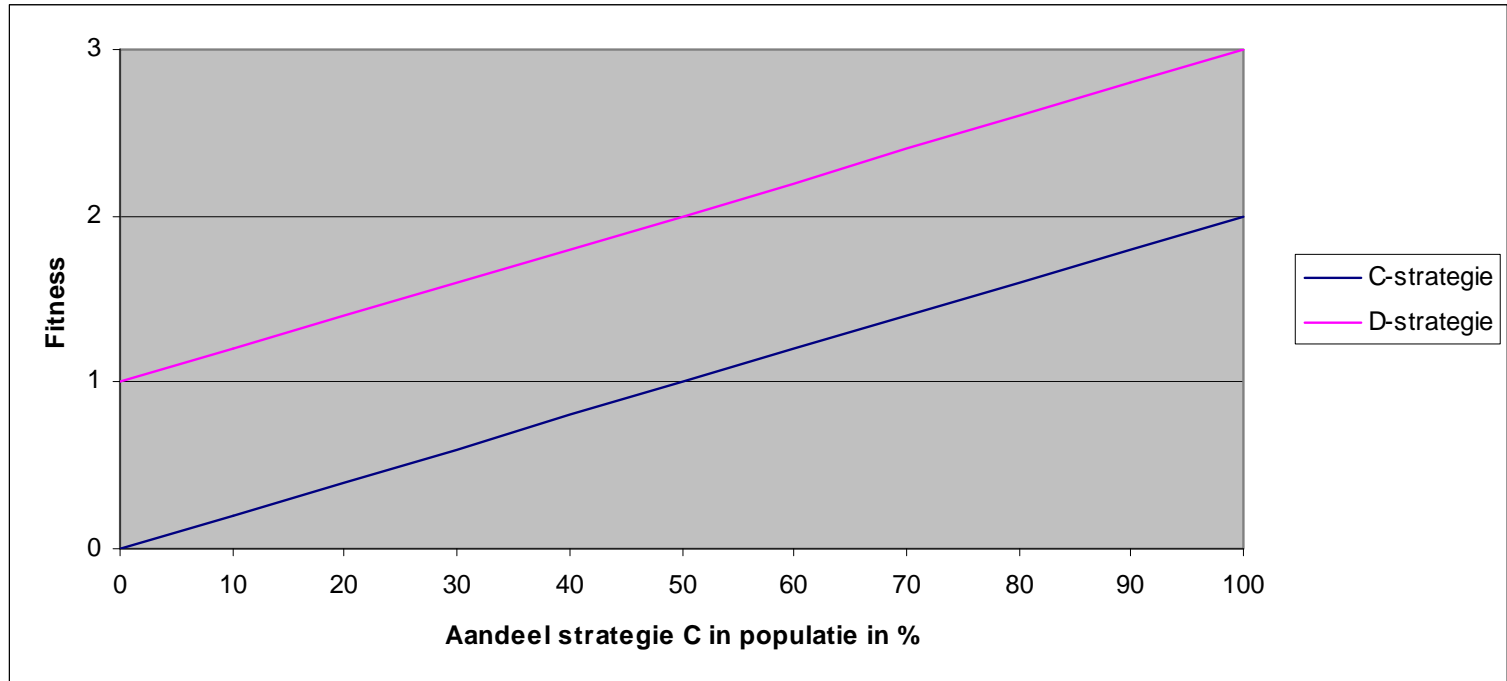
- (C,C) zou voor beide spelers beter zijn (is optimaler)

(Het spelen van C is altruïstisch in de zin dat, wat de ander ook doet, de *fitness* van de ander groter is en jouw eigen *fitness* lager is dan wanneer je D zou spelen)

- In evolutionaire setting: het maakt niet uit hoe de samenstelling van de populatie is, niet-samenwerken heeft altijd een hogere *fitness* dan samenwerken



5. Het Prisoners' Dilemma



5. Het Prisoners' Dilemma

- Vanuit speltheoretisch probleem is het evolutionair probleem van altruïsme er een van *free riding*: evt. samenwerken wordt altijd van binnenuit uitgehold door *free riders*

(Dit is ook één van de belangrijkste theoretische tegenwerpingen tegen **groepsselectie** “oude stijl”: ook al hebben individuen in groepen met veel samenwerkers een hogere *fitness* dan individuen in andere groepen, in elke groep wordt samenwerking van binnenuit uitgehold door *free riders*)



6. Hoe zou altruïsme toch kunnen evolueren in een PD?

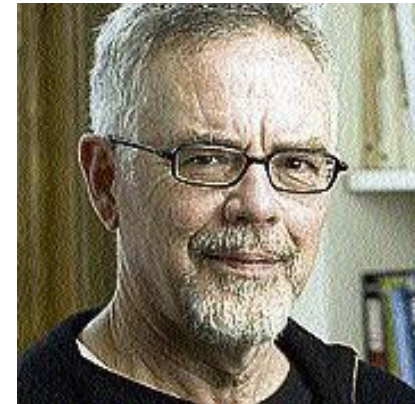
- Hoe is aan de conclusie te ontsnappen dat altruïsme niet kan evolueren?
- Grofweg twee oplossingen:
 1. *Herhalingen* van spel (plus aanvullende veronderstellingen als toekomstige opbrengsten moeten voldoende gewaardeerd worden, individuen moeten kunnen onthouden wat anderen gedaan hebben, ...)
 2. *Selectieve interactie*: samenwerkers herkennen elkaar en kunnen interactie met niet-samenwerkers weigeren



6. Hoe zou altruïsme toch kunnen evolueren in een PD?

Ad 1. Herhalingen

- Wederkerig altruïsme (R. Trivers)



- Tit For Tat (R. Axelrod; eigenlijk Rapoport): 'bestrafen' van niet-samenwerken en 'belonen' van samenwerken is ingebouwd



6. Hoe zou altruïsme toch kunnen evolueren in een PD?

- Strikt genomen zeggen deze vormen van (wederkerig) altruïsme niets over motieven en intenties voor gedrag, maar de neiging om motieven en intenties hierin te lezen is natuurlijk en misschien onbedwingbaar



6. Hoe zou altruïsme toch kunnen evolueren in een PD?

- Zie hier ook weer het verband dat vaak wordt gelegd tussen altruïsme, coöperatie en reciprociteit
- *Folk theorem* in speltheorie (Binmore): als spelen een onbekend aantal keren herhaald worden, komen ook coöperatieve oplossingen in het vizier



6. Hoe zou altruïsme toch kunnen evolueren in een PD?

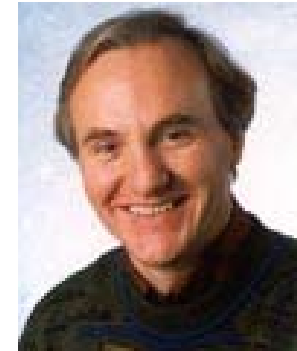
Ad 2. Selectieve interactie

- correlatie (Skyrms)
- groepsselectie à la Sober and Wilson (1998)
- emoties als *commitment devices* (Frank 1987)
- ...



6. Hoe zou altruïsme toch kunnen evolueren in een PD?

- Recente ontwikkeling (controversieel)



Ernst
Fehr

Strong reciprocity (Trivers' *reciprocal altruism* en Axelrods *Tit For Tat* worden als vormen van *weak reciprocity* gezien):

Combinatie van *altruistic punishment* en *altruistic rewarding* (Fehr, Bowles, Gintis, Henrich, ...)



6. Hoe zou altruïsme toch kunnen evolueren in een PD?

- *Strong reciprocity* wordt wel degelijk geïnterpreteerd in termen van *proximate causes* (“Geeft zonder hoop op beloning”): experimenten lijken uit te wijzen dat mensen bereid zijn om kosten te maken om zelf niet-samenwerkers te bestraffen en samenwerkers te belonen zonder dat ze verwachten daar zelf beter van te worden (het zijn vooral anderen die profiteren)



6. Hoe zou altruïsme toch kunnen evolueren in een PD?

- *Proximate cause:*

“wraak smaakt zoet” (straffen geeft psychische bevrediging); bevestigd in “neuro-experimenten”?

- *Ultimate cause:*

Culturele groepsselectie (theoretisch kader: Boyd and Richersons (1985) *gene-culture co-evolution*)



7. Conclusie

- Vanuit de evolutionaire speltheorie bezien is het begrijpelijk dat altruïsme, coöperatie en reciprociteit vaak aan elkaar worden gekoppeld
- Het is voor een goed begrip van evolutietheoretische verklaringen van altruïsme van het grootste belang om ultieme en proximale oorzaken te onderscheiden;

Tegelijkertijd blijkt dat (ook voor evolutietheoretici zelf) uitermate moeilijk te zijn

