Durham University

Social Complexity

Computational Social Science

Lecture 1

Welcome to Computational Social Science

Practicalities: Explain how the module works

- Introductions at end

Social science deals with complex problems

- Social systems are complex
- Problems of interest are increasingly complex

Computational social science provides a set of methods that are suited to complex problems

- Contrast with statistical methods

Need to combine data and theory for social science

Durham University

Structure of Module

SOCI44115_2022

Computational Social Science (22/23)

Content | Calendar | Discussions | Gradebook | Messages | Analytics

Course Staff

Jennifer Badham
INSTRUCTOR

Vikki Boliver
INSTRUCTOR

Show more

Details & Actions

Class register
View everyone on your course

Course Groups
View sets & groups

Progress Tracking
Turn on

Course Image
Edit display settings

Course is private
Students can't access this course

Blackboard Collaborate
Join session

Announcements
Create announcement

Books & Tools
View course & institution tools

Question Banks
Manage banks

Microsoft Teams
Enable Microsoft Teams

Add course schedule

Skip

Course Content

Module Information - Start Here!
Visible to students
Key information about the module: what it is about; who is teaching it; and how you will learn. The Module Overview includes a topic list of the lectures and workshops for each week.

Assessment Instructions and Extension Request
Visible to students
Details about assessment for the module, including assessment instructions, Turnitin submission points and extension request form.

Week 1: Social Complexity
Visible to students
What does it mean when we say society is complex? What are the implications for understanding and explaining social issues?

Week 2+3: Case-Based Complexity
Visible to students
Clustering and classification methods, finding similarity and differences between cases

Week 4+5: Qualitative Comparative Analysis
Visible to students
Deriving explanation from similarity and differences between cases

Week 6+7: Network Analysis
Visible to students
Focussing on the relationships between entities rather than the entities as individuals

Week 8+9: Simulation
Visible to students
Agent-based modelling and other simulation methods that represent processes or changes in a social system

Week 10: Computational Social Science
Visible to students
Bringing it all together: how do computational social science methods deal with issues discussed in the first week?

Reading List (Full)
Visible to students
The reading list is organised by topic and the link within the topic folder shows just the readings for that topic. If you want to see the whole reading list, you can use this link.

Teaching Staff (yes, just me)

Who am I?

Dr Jennifer Badham

- Teach and convene module
- First contact with questions or for support

Talking to me, or about me

- Jen (or Jennifer)
- Dr Badham
- She/her

Contact details

Email: jennifer.badham@durham.ac.uk

- Generally respond by end of following business day

Office hours:

By appointment

- For in-person, online or Teams chat.
- Remember, I may already be talking with another student

Room B3, 32 Old Elvet

Durham University

What will you learn about?

First and last lectures discuss key ideas from complexity and methods

Four methods with paired sessions

- Lecture ‘about’
- Workshop ‘to do’ application

Week 1: Social Complexity

Weeks 2 and 3: Big Data, Data Mining and Clustering

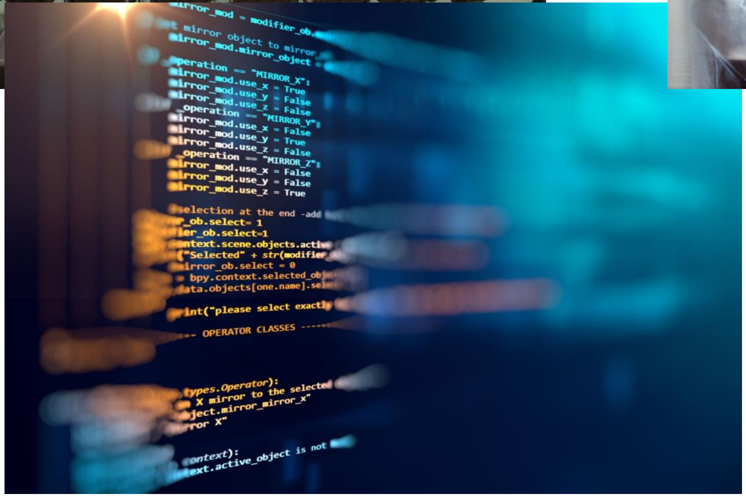
Weeks 4 and 5: Qualitative Comparative Analysis

Weeks 6 and 7: Social Network Analysis

Weeks 8 and 9: Simulation, particularly Agent-Based Modelling

Week 10: Synthesis

How will you learn? Four modes work together



Durham
University

Lectures

Lectures cover key ideas

- Information is organised
- Provide scaffolding

Recorded, available within day

Materials uploaded in advance

- Do NOT read in advance
- Take notes during lecture

Durham University

Workshops

Workshop for each of four methods

- Practice doing the method

Three in R with specialist packages

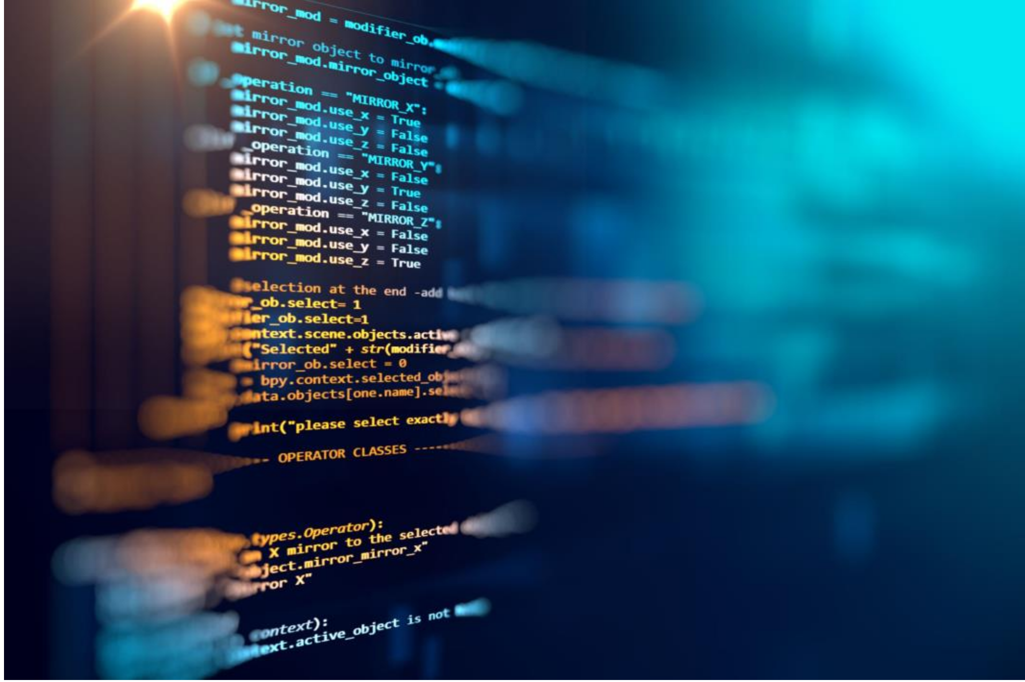
- Download provided script
- Run in sections and discuss

ABM workshop in NetLogo

Not recorded

Material uploaded after workshop

- Some will complete yourself



Discussion

Active participation in lectures and workshops

- Ask questions if unclear
- Consider questions posed
  - Better to think than simply listen to someone else's answer
- Help with workshop errors

Some specific breaks for discussion

Durham University

Reading

Extensive reading list

Required reading

- Insufficient lecture time for everything you should know

Preparatory reading

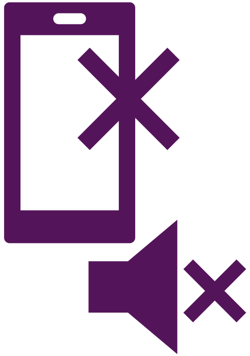
- Advance of lecture, needed to participate

Further reading

- For interest and for assessment

Logistics

Every Thursday (10 weeks)
16:05 to 17:55
2 hours, so break
Note the room changes



One assessment (22 April), plus optional formative (1 March)

Two options, ONE of the following (3000 words)

1/ Compare two (from four in class) methods

- Responding to different aspects of a topic you choose
- Why that method for that aspect, both strengths and limitations
- HINT: Think about methods that could complement each other

2/ Use one of the methods (from four in class)

- Choose a question with a relevant dataset / model
- Provide context about the question and choice of method
- Undertake partial analysis – see assessment description for detail
- HINT: Be guided by what is possible with your dataset or model

Lots of information in the assessment description

Extensions and adjustments have forms

Blackboard Ultra run-through

SOCI44115_2022

Computational Social Science (22/23)

Content | Calendar | Discussions | Gradebook | Messages | Analytics

Course Staff

Jennifer Badham
INSTRUCTOR

Vikki Boliver
INSTRUCTOR

Show more

Details & Actions

- Class register — View everyone on your course
- Course Groups — View sets & groups
- Progress Tracking — Turn on
- Course Image — Edit display settings
- Course is private — Students can't access this course
- Blackboard Collaborate — Join session
- Announcements — Create announcement
- Books & Tools — View course & institution tools
- Question Banks — Manage banks
- Microsoft Teams — Enable Microsoft Teams

Add course schedule

Skip

Course Content

Module Information - Start Here!
Visible to students
Key information about the module: what it is about; who is teaching it; and how you will learn. The Module Overview includes a topic list of the lectures and workshops for each week.

Assessment Instructions and Extension Request
Visible to students
Details about assessment for the module, including assessment instructions, Turnitin submission points and extension request form.

Week 1: Social Complexity
Visible to students
What does it mean when we say society is complex? What are the implications for understanding and explaining social issues?

Week 2+3: Case-Based Complexity
Visible to students
Clustering and classification methods, finding similarity and differences between cases

Week 4+5: Qualitative Comparative Analysis
Visible to students
Deriving explanation from similarity and differences between cases

Week 6+7: Network Analysis
Visible to students
Focussing on the relationships between entities rather than the entities as individuals

Week 8+9: Simulation
Visible to students
Agent-based modelling and other simulation methods that represent processes or changes in a social system

Week 10: Computational Social Science
Visible to students
Bringing it all together: how do computational social science methods deal with issues discussed in the first week?

Reading List (Full)
Visible to students
The reading list is organised by topic and the link within the topic folder shows just the readings for that topic. If you want to see the whole reading list, you can use this link.

Durham University

Why are you here?

Please introduce yourself

- Name
- Course (MARM, MDS, PhD….)

Why did you pick this module?

Durham University

Introduction

It is a nuisance, but God has chosen to give the easy problems to the physicists.

Lave & March (1993). *An Introduction to Models in the Social Sciences.* University Press of America. Lanham, MD

Weaver’s organisation of scientific problems

SCIENCE AND COMPLEXITY

By WARREN WEAVER

Rockefeller Foundation, New York City

SCIENCE has led to a multitude of results that affect men’s lives. Some of these results are embodied in mere conveniences of a relatively trivial sort. Many of them, based on science and developed through technology, are essential to the machinery of modern life. Many other results, especially those associated with the biological and medical sciences, are of unquestioned benefit and comfort. Certain aspects of science have profoundly influenced men’s ideas and even their ideals. Still other aspects of science are thoroughly awesome.

How can we get a view of the function that science should have in the developing future of man? How can we appreciate what science really is and, equally important, what science is not? It is, of course, possible to discuss the nature of science in general philosophical terms. For some purposes such a discussion is important and necessary, but for the present a more direct approach is desirable. Let us, as a very realistic politician used to say, let us look at the record. Neglecting the older history of science, we shall go back only three and a half centuries and take a broad view that tries to see the main features, and omits minor details. Let us begin with the physical sciences, rather than the biological, for the place of the life sciences in the descriptive scheme will gradually become evident.

Edited volume 1947: *Scientists Speak*

- Series of 81 talks from scientists
  - Presented during intermission of New York Philharmonic Symphony
- Organised the material with themes

Introduction adapted for the *American Scientist* article

- Three types of problems

Weaver (1948). "Science and Complexity". *American Scientist*, 36(4): 536-44

Phase 1 (to about 1900): Problems of simplicity

Physical sciences

Experimental methods

Mathematical relationships between specific variables, eg

- Force, motion and acceleration
- Motion of single billiard ball

Developed technology


Life sciences

Problems of simplicity not important

Limited quantitative theory

- Problems involve multiple variables changing in interdependent ways
- Cannot isolate two variables and hold others constant

Collect, describe, classify

- Observed correlations

Durham University

Phase 2 (current, 1948): Problems of disorganised complexity

Statistical methods are powerful

Very large number of variables that behave randomly

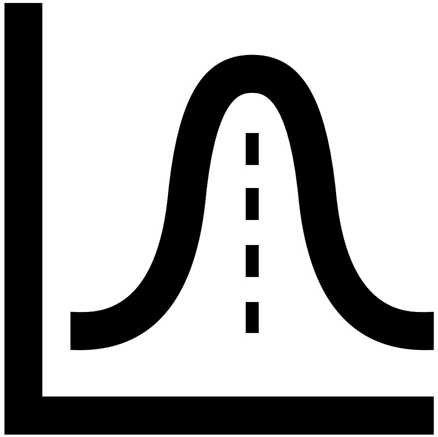
- Random is why referred to as disorganised

Average behaviour becomes predictable

- Relationship between pressure and temperature in gas
- Frequency of collisions between millions of billiard balls

Applies even if the simple version not understood

- Mortality rates predictable where individual deaths are not



Phase 3 (future): Problems of organised complexity

Consider several billiard balls that start in a row and all pushed directly across the table, rebounding to same place

- Would never collide as moving on parallel paths
- The starting conditions are organised, not random

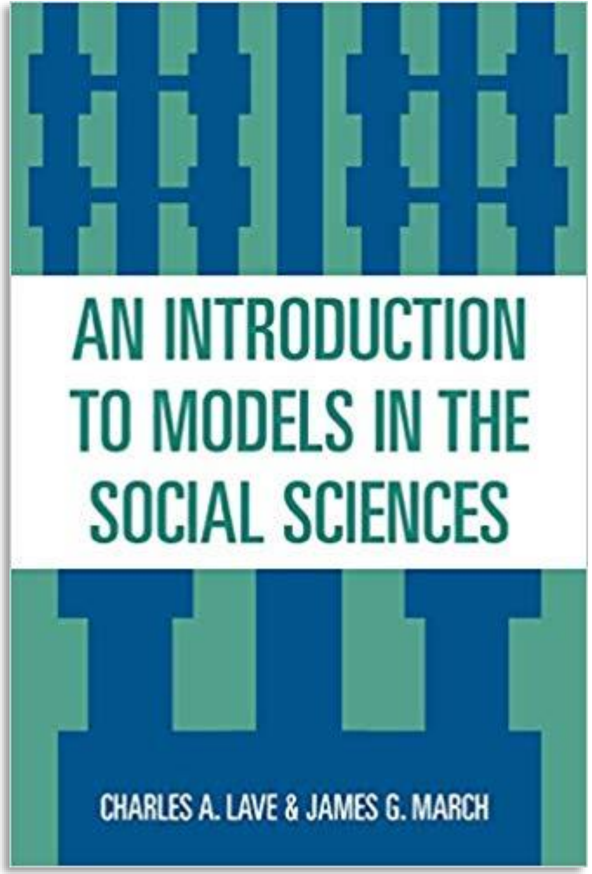
Organised complexity problems where "*sizeable number of factors which are interrelated into an organic whole*"

Common in life and social sciences

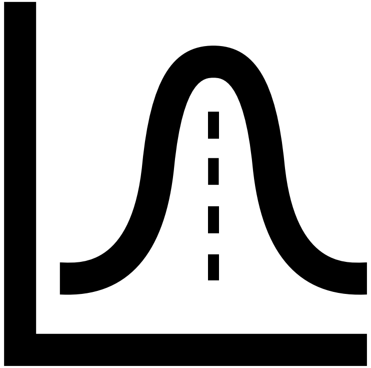
- Why is one chemical a poison and the mirror image molecule harmless?
- How can currency be stabilised?
- How to bring about a stable, decent and peaceful world?

Different types of sciences have different types of problems

It is a nuisance, but God has chosen to give the easy problems to the physicists.



Methods are specific to the type of problem



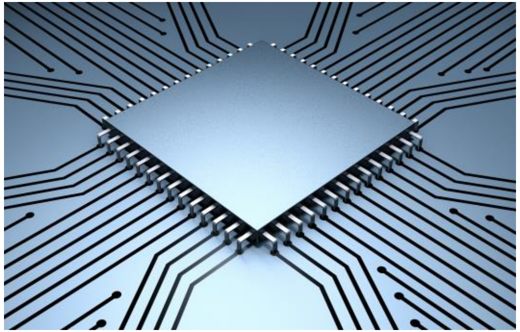
Simple

- Experiments
- Mathematical formulas

Disorganised complex

- Statistics

Organised complex: Promise of future (in 1948) methods



- Computing power
- Multidisciplinary teams

Understanding complexity

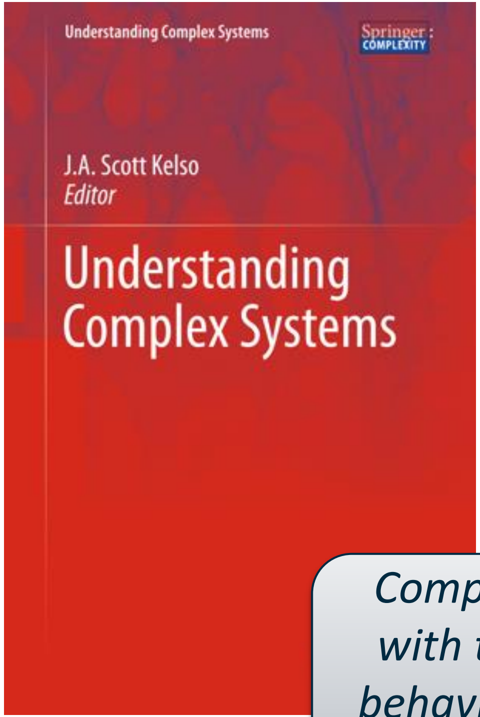
Durham University

What makes many cars a traffic jam?

Some definitions

sizeable number of factors which are interrelated into an organic whole

Weaver (1948)



that which is compounded out of something so that the whole is one, not like a heap but like a syllable-now the syllable is not its elements, ba is not the same as b and a [perhaps better to think of 's' and 'h' and 'sh']

Aristotle (~350 BCE), Metaphysics VII:17
Translation by WD Ross http://classics.mit.edu/Aristotle/metaphysics.7.vii.html

Kelso

Complex Systems are systems that comprise many interacting parts with the ability to generate a new quality of macroscopic collective behavior the manifestations of which are the spontaneous formation of distinctive temporal, spatial or functional structures.

Behaviour of system driven by interaction

A system is complex if the ***system's behaviour of interest is driven by the interactions*** between the individual parts, not simply the behaviour of the parts independently.

Something 'more' than having many parts that are linked in complicated ways

Traffic jam arises from the interactions between cars (drivers really), also interactions with environmental factors such as traffic lights

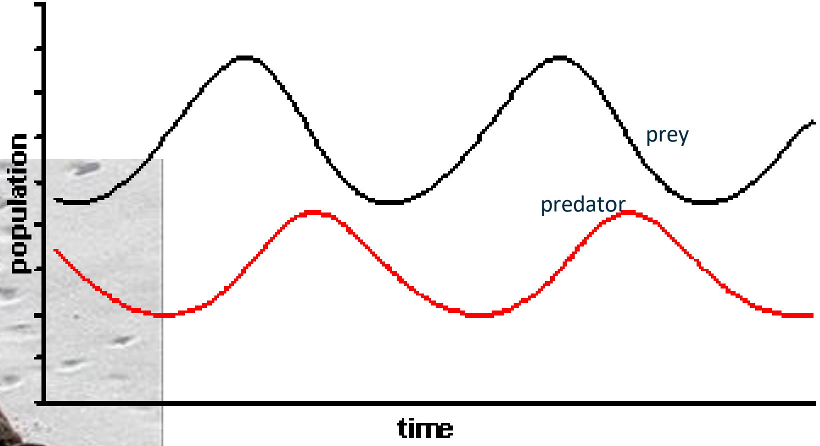
Durham University

Example: Standing ovation is cascade behaviour

What if first person to stand was at the back?

Durham University

Example: Predator and prey population sizes cycle



Predator cycle follows prey cycle with delay, as prey available the predators have children and that leads to a food shortage

Characteristics: Collective Behaviour

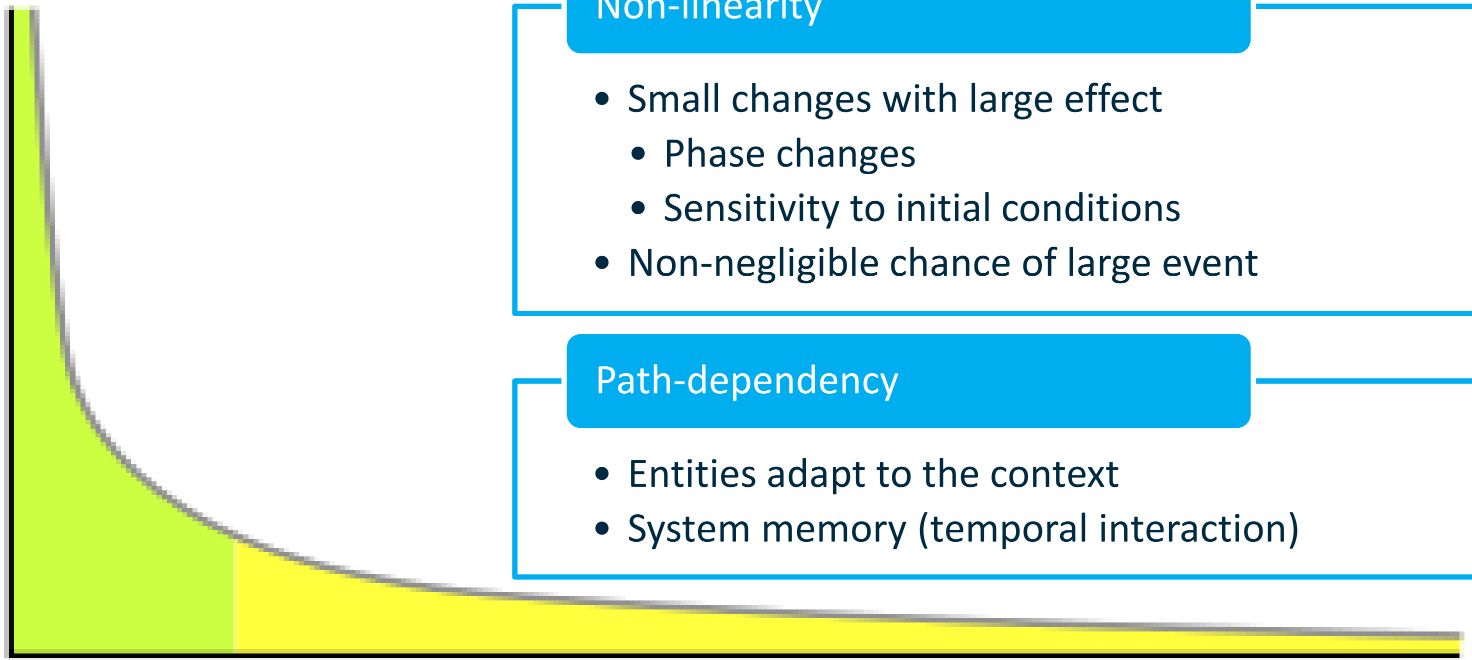
Emergence

- Formal term for the system having properties that the parts do not have

Self-organisation

- The system exhibits order
- Parts (entities) are autonomous, without central control

Characteristics: Feedback



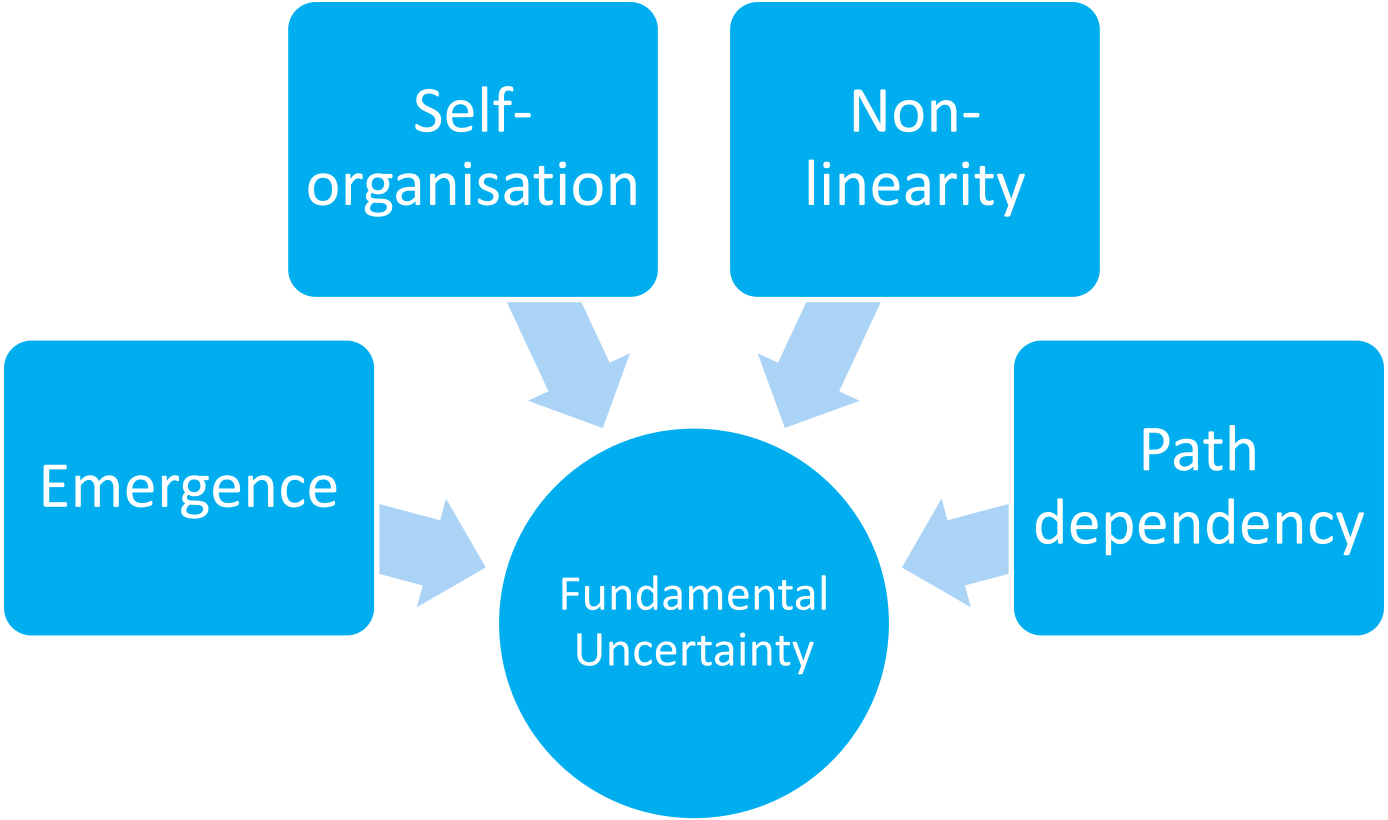
Non-linearity

- Small changes with large effect
  - Phase changes
  - Sensitivity to initial conditions
- Non-negligible chance of large event

Path-dependency

- Entities adapt to the context
- System memory (temporal interaction)

Characteristics: Limited predictability



Durham University

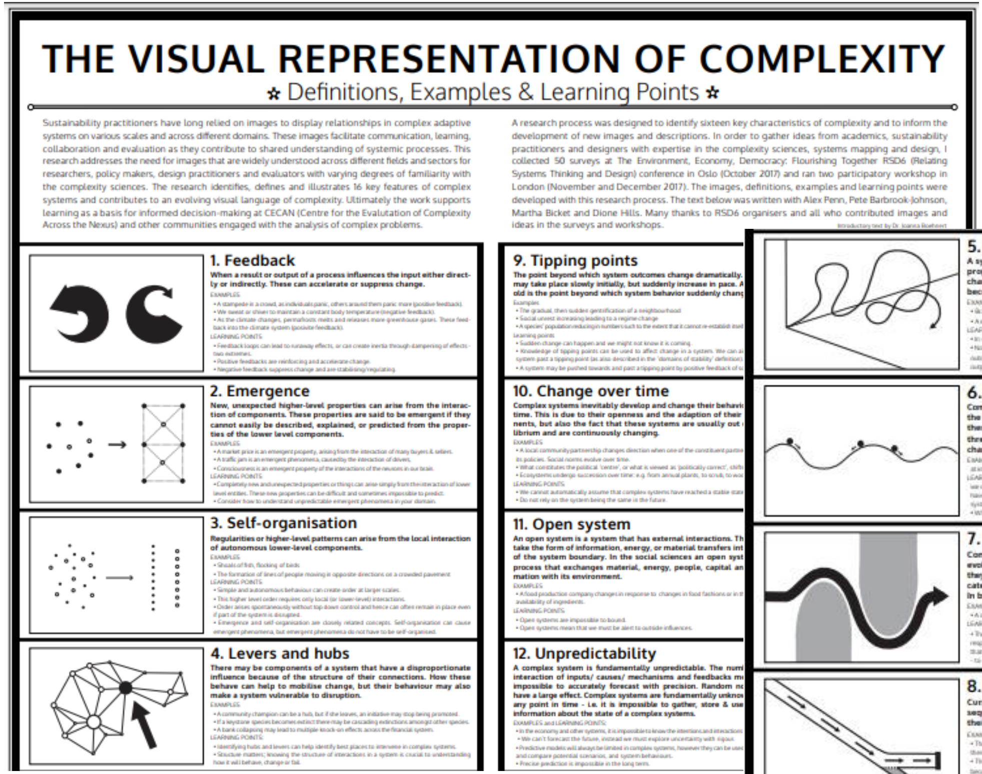
NOT all properties in all systems

You may see definitions of complex systems that include these properties

Many complex systems display several of these properties, but only interaction driven causal process is truly required

- Some older definitions of emergence describe the emergent behaviour as ‘surprising’, but better to think of it as different

A broader collection of complex problem properties



Boehnert (2018). “The Visual Representation of Complexity". *Proceedings of Relating Systems Thinking and Design (RSDX) Symposium*

Durham University

1. Feedback

When a result or output of a process influences the input either directly or indirectly. These can accelerate or suppress change.

EXAMPLES

- A stampede in a crowd, as individuals panic, others around them panic more (positive feedback).
- We sweat or shiver to maintain a constant body temperature (negative feedback).
- As the climate changes, permafrosts melts and releases more greenhouse gases. These feedback into the climate system (posivite feedback).

LEARNING POINTS

- Feedback loops can lead to runaway effects, or can create inertia through dampening of effects - two extremes.
- Positive feedbacks are reinforcing and accelerate change.
- Negative feedback suppress change and are stabilising/regulating.

2. Emergence

New, unexpected higher-level properties can arise from the interaction of components. These properties are said to be emergent if they cannot easily be described, explained, or predicted from the properties of the lower level components.

EXAMPLES

- A market price is an emergent property, arising from the interaction of many buyers & sellers.
- A traffic jam is an emergent phenomena, caused by the interaction of drivers,
- Consciousness is an emergent property of the interactions of the neurons in our brain.

LEARNING POINTS

- Completely new and unexpected properties or things can arise simply from the interaction of lower level entities. These new properties can be difficult and sometimes impossible to predict.
- Consider how to understand unpredictable emergent phenomena in your domain.

3. Self-organisation

Regularities or higher-level patterns can arise from the local interaction of autonomous lower-level components.

EXAMPLES

- Shoals of fish, flocking of birds
- The formation of lines of people moving in opposite directions on a crowded pavement

LEARNING POINTS

- Simple and autonomous behaviour can create order at larger scales.
- This higher level order requires only local (or lower-level) interactions.
- Order arises spontaneously without top down control and hence can often remain in place even if part of the system is disrupted.
- Emergence and self-organisation are closely related concepts. Self-organisation can cause emergent phenomena, but emergent phenomena do not have to be self-organised.

4. Levers and hubs

There may be components of a system that have a disproportionate influence because of the structure of their connections. How these behave can help to mobilise change, but their behaviour may also make a system vulnerable to disruption.

EXAMPLES

- A community champion can be a hub, but if she leaves, an initiative may stop being promoted.
- If a keystone species becomes extinct there may be cascading extinctions amongst other species.
- A bank collapsing may lead to multiple knock-on effects across the financial system.

LEARNING POINTS:

- Identifying hubs and levers can help identify best places to intervene in complex systems.
- Structure matters; knowing the structure of interactions in a system is crucial to understanding how it will behave, change or fail.

5. Non-linearity

A system is non-linear when the effect of inputs on outcomes are not proportional. The behaviour of a system may exhibit exponential changes, or changes in direction (i.e., increases in some measure becoming decreases), despite small or consistent changes in inputs.

EXAMPLES

- Braking distance in a car at 30MpH is more than twice that at 20MpH
- A new product may be slow to take-off but after a point sales will accelerate, before slowing again.

LEARNING POINTS

- In social settings, few things are actually linear
- Non-linearity can mean that the relationships between things can be just as powerful in determining outcomes as the structure of interactions. • In non-linear systems when we double or half an input, the output will not be double or half its original value, and may be completely different.

6. Domains of stability

Complex systems may have multiple stable states which can change as the context evolves. Systems gravitate towards such states, remaining there unless significantly perturbed. If change in a system passes a threshold, it may slide rapidly into another stable state, making change very difficult to reverse.

EXAMPLES • The melting of Antarctic ice: The planet may be stable with or without ice caps, but not at intermediate states. • Poverty traps: Low or reasonable incomes are stable, but not intermediates

LEARNING POINTS • Knowledge of domains of stability can be used to effect change in a system. If we can push a system into a different, more desirable, stable state with a policy intervention then we have changed the system in a robust way. • We do not need to put in continuous effort to keep the system in the new state. • We may try to use policy to change the positions of domains of stability. • What is possible in a system is often discontinuous and sticky. Not everything is stable.

7. Adaptation

Components or actors within the system are capable of learning or evolving, changing how the system behaves in response to interventions as they are applied. So, for example, in social systems people may communicate, interpret and behave strategically to anticipate future situations. In biological systems, species will evolve in response to change.

EXAMPLES • Bacteria evolve to become resistant to antibiotics.

- A new tax/regulation is circumvented.

LEARNING POINTS:

- The rules of the game change as you play it. • We have to be prepared to adapt our interventions in response to how the system reacts to previous input. • We should be aware of the pressures to adapt that we are putting in place in systems. • We also need to be prepared for individuals - and systems - to adapt in response to an intervention in ways we didn't anticipate.

8. Path dependency

Current and future states, actions, or decisions depend on the sequence of states, actions, or decisions that preceded them – namely their (typically temporal) path.

EXAMPLES

- The first fold of a piece of origami paper will determine which final shapes are possible; origami is therefore a path dependent art.
- The organisation chosen to lead a new policy initiative influences which other organisations also become involved.
- VHS + Betamax, or railways + gauges -> once one option is adopted it would be impractical to switch.

LEARNING POINTS

- What paths are we 'locked-into'? What paths might our actions lock us into? What is it that makes a particular change impossible because of path dependency? Which 'lock-ins' might shift soon?

9. Tipping points

The point beyond which system outcomes change dramatically. Change may take place slowly initially, but suddenly increase in pace. A threshold is the point beyond which system behavior suddenly changes.

Examples

- The gradual, then sudden gentrification of a neighbourhood
- Social unrest increasing leading to a regime change
- A species' population reducing in numbers such to the extent that it cannot re-establish itself in the wild.

Learning points

- Sudden change can happen and we might not know it is coming.
- Knowledge of tipping points can be used to affect change in a system. We can aim to get a system past a tipping point (as also described in the 'domains of stability' definition).
- A system may be pushed towards and past a tipping point by positive feedback of some kind.

10. Change over time

Complex systems inevitably develop and change their behaviour over time. This is due to their openness and the adaption of their components, but also the fact that these systems are usually out of equilibrium and are continuously changing.

EXAMPLES

- A local community partnership changes direction when one of the constituent partners changes its policies. Social norms evolve over time.
- What constitutes the political 'centre', or what is viewed as 'politically correct', shifts over time.
- Ecosystems undergo succession over time: e.g. from annual plants, to scrub, to woodland.

LEARNING POINTS

- We cannot automatically assume that complex systems have reached a stable state.
- Do not rely on the system being the same in the future.

11. Open system

An open system is a system that has external interactions. These can take the form of information, energy, or material transfers into or out of the system boundary. In the social sciences an open system is a process that exchanges material, energy, people, capital and information with its environment.

EXAMPLES

- A food production company changes in response to changes in food fashions or in the cost and availability of ingredients.

LEARNING POINTS

- Open systems are impossible to bound.
- Open systems mean that we must be alert to outside influences.

12. Unpredictability

A complex system is fundamentally unpredictable. The number and interaction of inputs/ causes/ mechanisms and feedbacks mean it is impossible to accurately forecast with precision. Random noise can have a large effect. Complex systems are fundamentally unknowable at any point in time - i.e. it is impossible to gather, store & use all the information about the state of a complex systems.

EXAMPLES and LEARNING POINTS:

- In the economy and other systems, it is impossible to know the intentions and interactions of all actors.
- We can't forecast the future, instead we must explore uncertainty with rigour.
- Predictive models will always be limited in complex systems, however they can be used to explore and compare potential scenarios, and system behaviours.
- Precise prediction is impossible in the long term.

13. Unknowns

Because of their complex causal structure and openness, there are many factors which influence (or can influence) a system of which we are not aware. The inevitable existence of such unknowns mean we often see unexpected indirect effects of our interventions.

EXAMPLES

- A powerful social grouping operating in a policy area not anticipated by a policy maker.
- An undiscovered plant in a rainforest with numerous potential health applications.

LEARNING POINTS

- Expect the unexpected.
- Be prepared to learn as the system unfolds it will become apparent that it might influence or be influenced by completely unexpected things.
- A new technology might enable a fundamental change, leading to widespread social effects.

14. Distributed control

Control of a system is distributed amongst many actors. No one actor has total control. Each actor may only have access to local information.

EXAMPLES

- A smoking cessation intervention's success may be determined by the many health professionals 'on the ground' running events and offering advice, rather than the central agency.
- Political parties' local groups and government may have differing views to the central parliamentary party. The central and distributed groups may conduct poltical work in contradictory ways.

LEARNING POINTS

- There is no top down control in complex systems. Decisions and reactions happen locally and the interactions of all these lower-level decisions can give us system-level properties such as stability, resilience, adaptation or whole system emergent regulation.
- The best we can do is to "steer" the system.

15. Nested systems

Complex systems are often nested hierarchies of complex systems (so-called 'systems of systems').

EXAMPLES

- Brain -> person -> society -> planet
- An ecosystem is made up of organisms, made up of cells, made up of organelles which were once free-living bacteria, made up of complex metabolic processes intertwined with genetic systems (each nested level is a complex system).

LEARNING POINTS

- When studying a particular system, it is useful to be aware of the larger system of which it is part, or the smaller systems operating within it.
- Mechanisms of change (as in realist evaluation) may be taking place at a higher or lower level to the one where an intervention is taking place.

16. Multiple scales and levels

Actors and interactions in complex systems can operate across scales and levels. For this reason systems must be studied and understood from multiple perspectives simultaneously.

EXAMPLES

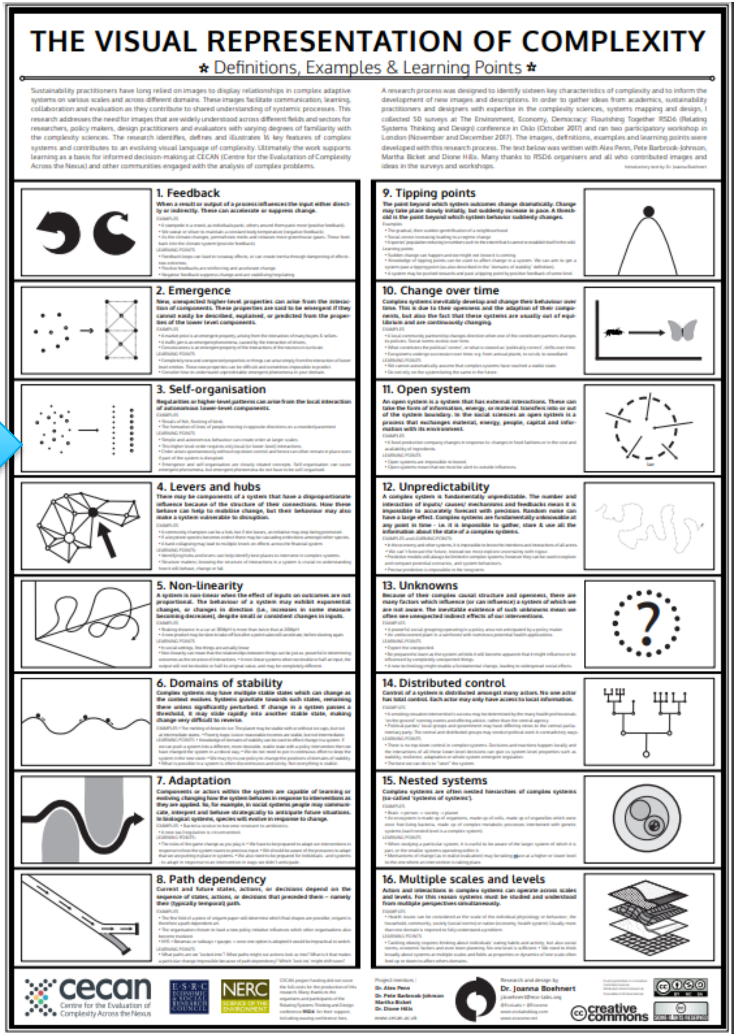
- Health issues can be considered at the scale of the individual physiology or behaviour, the household, community, society (social norms) or nation (economy, health system). Usually more than one domain is required to fully understand a problem.

LEARNING POINTS

- Tackling obesity requires thinking about individuals' eating habits and activity, but also social norms, economic factors and even town planning. No one level is sufficient. • We need to think broadly about systems at multiple scales and fields as properties or dynamics of one scale often feed up or down to affect others domains.

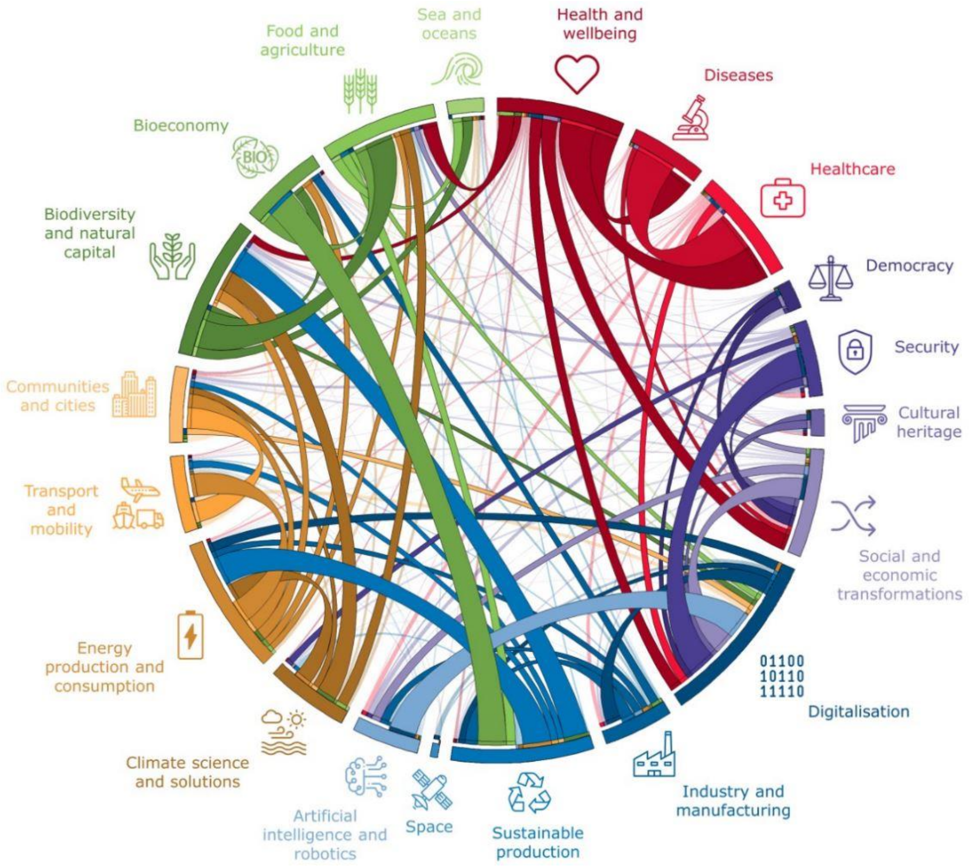
Discussion

How does the example system(s) display complex properties?



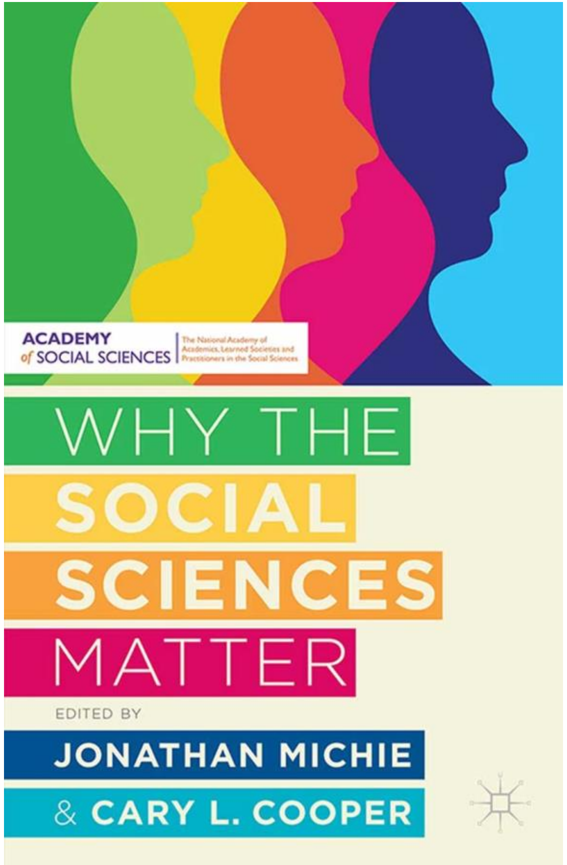
Durham University

Complexity in social science



Durham University

Social science deals with important issues



Social science deals with important issues

Social scientists help us imagine alternative futures

Social science can help us make sense of our finances (including social welfare)

Social scientists contribute to our health and well-being

Social science might save your life (eg workplace safety)

Social science can make your neighbourhood safer

We need social scientists as public intellectuals

Social science can improve our children's lives and education

Social science can change the world for the better (eg human rights)

Social science can broaden your horizons

We need social science to guarantee our democracy

Experimental methods help us to understand people

ECONOMETRICA

VOLUME 47 MARCH, 1979 NUMBER 2

PROSPECT THEORY: AN ANALYSIS OF DECISION UNDER RISK

BY DANIEL KAHNEMAN AND AMOS TVERSKY[1]

This paper presents a critique of expected utility theory as a descriptive model of decision making under risk, and develops an alternative model, called prospect theory. Choices among risky prospects exhibit several pervasive effects that are inconsistent with the basic tenets of utility theory. In particular, people underweight outcomes that are merely probable in comparison with outcomes that are obtained with certainty. This tendency, called the certainty effect, contributes to risk aversion in choices involving sure gains and to risk seeking in choices involving sure losses. In addition, people generally discard components that are shared by all prospects under consideration. This tendency, called the isolation effect, leads to inconsistent preferences when the same choice is presented in different forms. An alternative theory of choice is developed, in which value is assigned to gains and losses rather than to final assets and in which probabilities are replaced by decision weights. The value function is normally concave for gains, commonly convex for losses, and is generally steeper for losses than for gains. Decision weights are generally lower than the corresponding probabilities, except in the range of low probabilities. Overweighting of low probabilities may contribute to the attractiveness of both insurance and gambling.

Example: Many biases in the way people make decisions

Prospect theory

- Loss is more important than gain
- Small probabilities are overweighted
- Definite outcomes are overweighted

Qualitative methods help us to understand people

How to ask about gender in the census?

We explored relevant issues, including responses to sex and gender questions and potential barriers to answering, terminology, privacy, burden and acceptability. The Office for National Statistics (ONS) Data Collection Methodology branch undertook 4 focus groups and 18 one-to-one in-depth interviews with members of the public, including those with a trans identity and those who are cisgender (people whose self-identity conforms to the sex or gender assigned at birth). These were conducted in March and April 2017.

Three question designs were explored:

- the 2011 Census "Sex" question
- a hybrid design with the addition of "Other" to the sex question
- a two-step design with separate sex and gender identity questions

None of the question designs, as they were presented, would meet the requirement for a reliable estimate of the trans population. Furthermore, none of the question designs would fully meet respondent needs for questions that are easily understood and answered. It is therefore recommended that none of these designs be used in the 2021 Census.

https://www.ons.gov.uk/methodology/classificationsandstandards/measuringequality/genderidentity/qualitativeresearchongenderidentityphase1summaryreport

Why do people refuse covid vaccination?

1 . Main points

- Many participants who were unwilling or uncertain about receiving a coronavirus (COVID-19) vaccine expressed concerns about their safety; these included concerns about immediate side effects and longer-term impacts that participants felt could not yet be known.
- Fears about the safety of COVID-19 vaccines were often linked with how quickly they had been developed; participants perceived this as a sign that the COVID-19 vaccines could not be as safe as other vaccines that had been developed and used over several years.
- Some participants did not perceive catching COVID-19 as a significant risk; typically, this was because they were younger and felt they were unlikely to either catch or develop serious symptoms from catching COVID-19, or because they felt they were already taking adequate steps to avoid catching COVID-19.
- Those who were unable to receive a COVID-19 vaccine cited barriers including: not being able to find childcare to attend the vaccination appointment; not being able to travel to the vaccination centre; or having existing physical or mental health conditions that prevented them from receiving a COVID-19 vaccine.
- There was an appetite for more information about COVID-19 vaccines, particularly: side effects; contents; how they had been developed; and differences between, and safety of, the various COVID-19 vaccines.
- Some participants accessed information about COVID-19 vaccines from social media or unverified sources as well as, or instead of, mainstream media; this gave them cause for concern, for example, about the contents or side effects of COVID-19 vaccines.

https://www.ons.gov.uk/releases/covid19vaccinationqualitativestudy

Statistical methods help us to understand society

Example: Understanding disadvantage

In the US, intergenerational mobility differs by race: With parents in the top 20% of income

- 41% of white children end up in top 20% of income
- 18% of black children end up in top 20% of income

But a long way to go on improving equity

Race and Economic Opportunity in the United States: an Intergenerational Perspective*

Raj Chetty, Nathaniel Hendren, Maggie R Jones, Sonya R Porter

The Quarterly Journal of Economics, Volume 135, Issue 2, May 2020, Pages 711–783,
https://doi.org/10.1093/qje/qjz042
Published: 26 December 2019

PDF Split View Cite Permissions Share ▾

Abstract

We study the sources of racial disparities in income using anonymized longitudinal data covering nearly the entire U.S. population from 1989 to 2015. We document three results. First, black Americans and American Indians have much lower rates of upward mobility and higher rates of downward mobility than whites, leading to persistent disparities across generations. Conditional on parent income, the black-white income gap is driven by differences in wages and employment rates between black and white men; there are no such differences between black and white women. Hispanic Americans have rates of intergenerational mobility more similar to whites than blacks, leading the Hispanic-white income gap to shrink across generations. Second, differences in parental marital status, education, and wealth explain little of the black-white income gap conditional on parent income. Third, the black-white gap persists even among boys who grow up in the same neighborhood. Controlling for parental income, black boys have lower incomes in adulthood than white boys in 99% of Census tracts. The few areas with small black-white gaps tend to be low-poverty neighborhoods with low levels of racial bias among whites and high rates of father presence among blacks. Black males who move to such neighborhoods earlier in childhood have significantly better outcomes. However, less than 5% of black children grow up in such areas. Our findings suggest that reducing the black-white income gap will require efforts whose impacts cross neighborhood and class lines and increase upward mobility specifically for black men.

Social science has improved people's lives

Public health is credited with adding 25 years to the life expectancy of people in the United States in this century. Yet, ask the average person what public health is and their reply might be limited to: "healthcare for low-income families." CDC's Ten Great Public Health Achievements in the 20th Century was created to remind us of how far we've come, how we got here, and exactly what public health is: the active protection of our nation's health and safety, credible information to enhance health decisions, and partnerships with local minorities and organizations to promote good health.

Learn more about how far we've come in the Morbidity and Mortality Weekly Report (MMWR):

Ten Great Public Health Achievements in the 20th Century

- Immunizations
- Motor-Vehicle Safety
- Workplace Safety
- Control of Infectious Diseases
- Declines in Deaths from Heart Disease and Stroke
- Safer and Healthier Foods
- Healthier Mothers and Babies
- Family Planning
- Fluoridation of Drinking Water
- Tobacco as a Health Hazard
- Future Directions of Public Health

Additional 25 years of life due to public health policies

Details at:

https://sphweb.bumc.bu.edu/otlt/mph-modules/ph/publichealthhistory/publichealthhistory9.html

CDC Centers for Disease Control and Prevention
CDC 24/7: Saving Lives, Protecting People™

http://www.cdc.gov/about/history/tengpha.htm

Durham University

Much of the contribution uses surveys and statistical analysis

National regular surveys

- Describe society
- Inform government plans

Familiar uses

- Estimate economic indicators such as inflation
- Measure social inequalities
- Track changes in social attitudes
- Identify gaps in services

Adhoc surveys and analysis

Government surveys on critical issues

- COVID-19 testing

Individual researchers testing a theory

Evaluating interventions

- New healthcare technology
- Changes in policy

Durham University

Alternative data sources now available

Social scientists had distinctive methods through latter half of 20th century that justified claims of expertise in social problems

Large transactional data makes such expertise irrelevant for some questions

- No need to use proxy variables (eg class) to infer similar expected transactions
- Interpretation does not rely on random selection and good response rates

The Coming Crisis of Empirical Sociology

Mike Savage
University of Manchester

Roger Burrows
University of York

ABSTRACT
This article argues that in an age of *knowing capitalism*, sociologists have not adequately thought about the challenges posed to their expertise by the proliferation of 'social' transactional data which are now routinely collected, processed and analysed by a wide variety of private and public institutions. Drawing on British examples, we argue that whereas over the past 40 years sociologists championed innovative methodological resources, notably the sample survey and the in-depth interviews, which reasonably allowed them to claim distinctive expertise to access the 'social' in powerful ways, such claims are now much less secure. We argue that both the sample survey and the in-depth interview are increasingly dated research methods, which are unlikely to provide a robust base for the jurisdiction of empirical sociologists in coming decades. We conclude by speculating how sociology might respond to this coming crisis through taking up new interests in the 'politics of method'.

Current social problems less amenable to survey methods

Global population overload

Global warming and climate change

Increasing ecological challenges – access to water, food

Entrenched inequality, within and between countries

Cultural conflict and terrorism

Fast moving epidemics, pandemics

Fragile infrastructure such as supply chain breakdowns

Durham University

In social policy, may be referred to as ‘wicked problems’

There is no definitive formulation of the problem

- Competing perspectives
- No specific end point

Solutions are relative; good or bad. The set of potential solutions (or even approaches) is not well described.

Every wicked problem is essentially unique

- Interventions cannot be readily undone
- No opportunity to learn by trial and error
- May be no public tolerance of failure

Every wicked problem can be considered to be a symptom of another problem

- Intervening in one area creates other problems

Policy Sciences 4 (1973), 155–169

Dilemmas in a General Theory of Planning*

HORST W. J. RITTEL
Professor of the Science of Design, University of California, Berkeley

MELVIN M. WEBBER
Professor of City Planning, University of California, Berkeley

ABSTRACT

The search for scientific bases for confronting problems of social policy is bound to fail, because of the nature of these problems. They are “wicked” problems, whereas science has developed to deal with “tame” problems. Policy problems cannot be definitively described. Moreover, in a pluralistic society there is nothing like the undisputable public good; there is no objective definition of equity; policies that respond to social problems cannot be meaningfully correct or false; and it makes no sense to talk about “optimal solutions” to social problems unless severe qualifications are imposed first. Even worse, there are no “solutions” in the sense of definitive and objective answers.

Characteristics of wicked problems are features of complexity

Wicked problems	Complexity framing
Entrenched problems resistant to change	Domains of stability
Poorly formulated with competing perspectives	Open system
Problems are symptoms of other problems	Feedback Interactions

Durham University

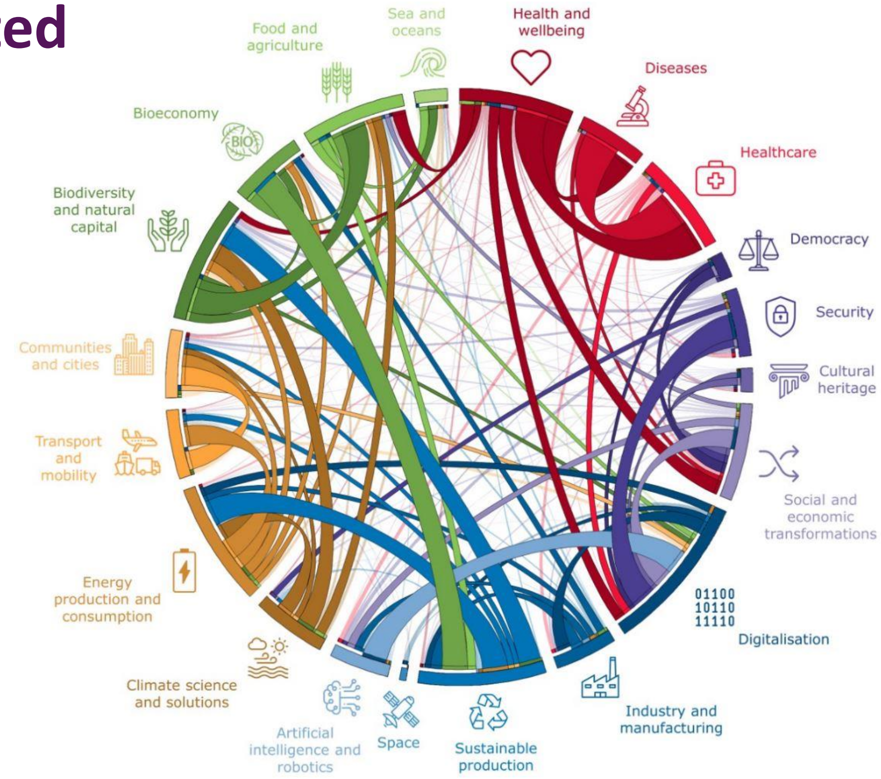
Social domains are interconnected

EU sought suggested missions for their research agenda

Comments by research organisations, NGOs, businesses and other organisations

Contributions addressed multiple intervention domains simultaneously

- Mapping of co-occurrence



European Commission, Directorate-General for Research and Innovation, Sorokins, J., Griniece, E., *Responses to the call for feedback on "Mission-Oriented Research & Innovation in the European Union" by Mariana Mazzucato : analysis report*, Dudek, P.(editor), Publications Office, 2018, https://data.europa.eu/doi/10.2777/870760 (Figure 6)

Durham University

Survey methods typically ignore context… (1968 quote, but still valid)

As usually practiced, using random sampling of individuals, the survey is a sociological meatgrinder, tearing the individual from his social context… If our aim is to understand people’s behaviour rather than simply record it, we want to know about primary groups, neighborhoods, social circles and communities; about interaction, communication, role expectations, and social control.


American Behavioral Scientist

Impac

Available access | Other | First published November 1968

Survey Research and Macro-Methodology

Allen H. Barton View all authors and affiliations

Volume 12, Issue 2 | https://doi.org/10.1177/000276426801200201

Contents | PDF / ePub | Cite article | Share options | Information, rig

BRINGING SOCIETY BACK IN

Survey Research and Macro-Methodology

ALLEN H. BARTON

Durham University

Survey methods can include context
But analysis is not a simple comparison of variables

Barton (1968) also describes methods to incorporate context

Ask about the characteristics of their friends, co-workers, neighbours, family…

- Perceived characteristics can also be compared to real characteristics if both surveyed

Use cluster sampling

- Measure for both cluster and individuals

Include ‘interpersonal environment’ in survey

- Network data: relationship and influence measures such as frequency/duration of contact

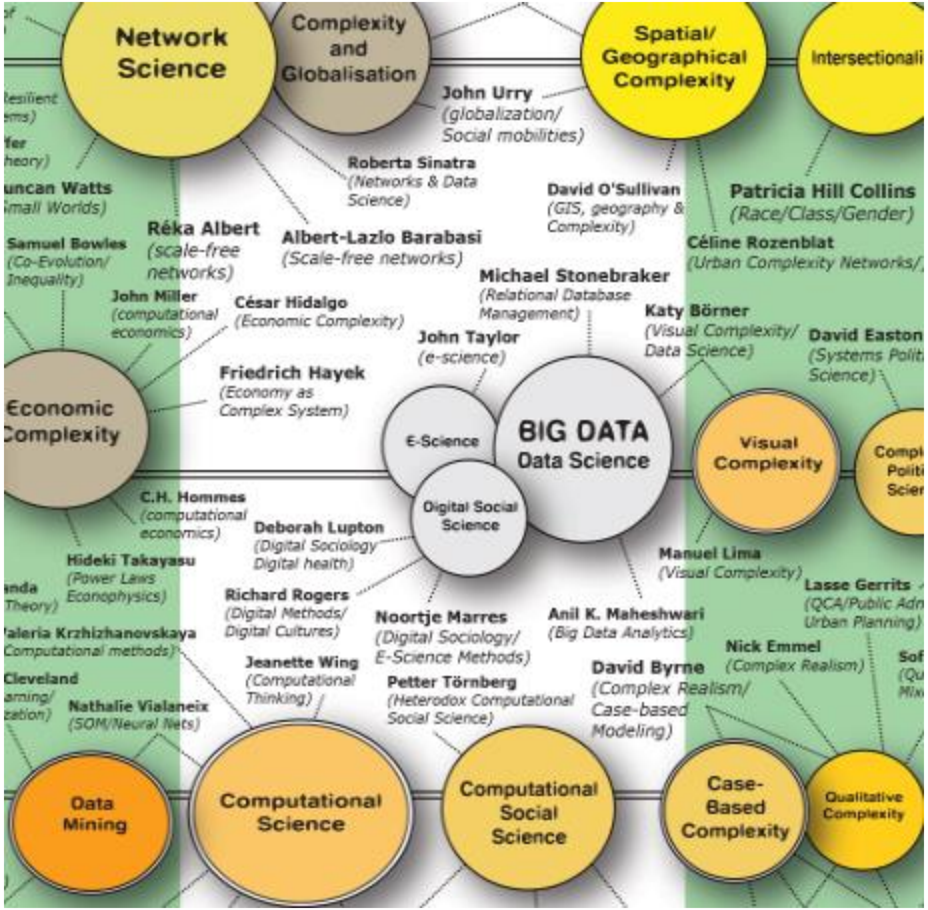
Supplement with information about institutional connections

- Enriches the analysis, for example by comparing opportunities

Barton, A. H. (1968). Bringing society back in: Survey research and macro-methodology. *American Behavioral Scientist*, 12(2),

Durham University

Complexity methods



Durham University

Why are complex problems difficult?

Reductionist training

Focus on ‘the’ (singular) cause

- Legal system: Who is at fault?

Scientific training:

- Experiments seek to isolate one difference
- Narrative with reasons is ‘subjective’
  - Low in the evidence hierarchy
- Randomised control trials report average
  - Diversity not interesting

Why doesn’t this work?

Reductionist techniques look at only one part

- Predator population has cyclic pattern but reason unclear without the prey
- Music cannot be analysed by looking at one instrument in isolation

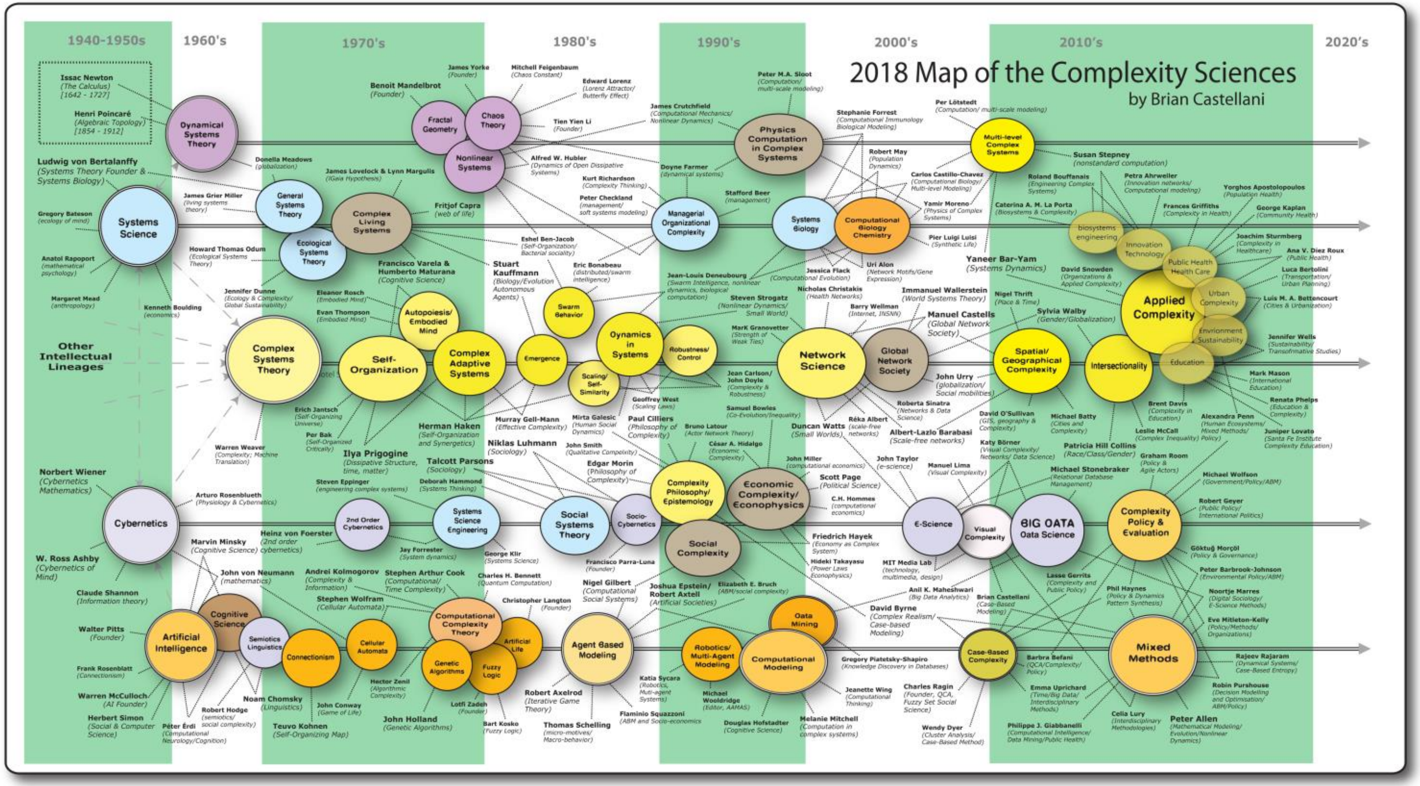
Expect `cause’ to be near ‘effect’

Even harder if people are involved

- Preferences and situation change

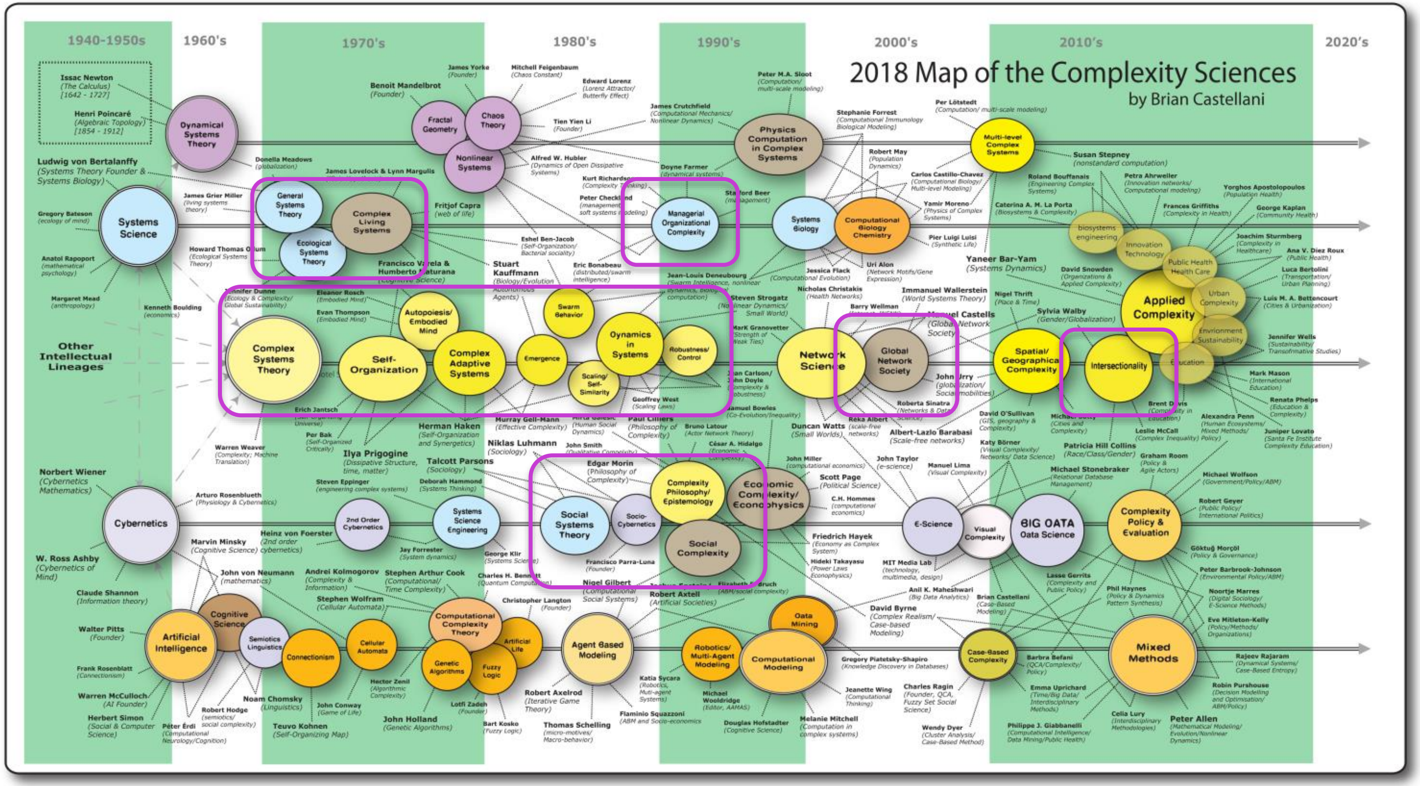
Many social issues have multiple contributing factors

Complexity sciences is not a single field



http://www.art-sciencefactory.com/complexity-map_feb09.html

Some of the complexity oriented theories



What makes social complexity theories different?

Focus on the intersection of factors

- Example: intersectionality theory (Patricia Hill Collins)

Study social complexity in systems terms (paths, relationships)

- Example: world systems theory (Immanuel Wallerstein)

See the structure of society in network terms

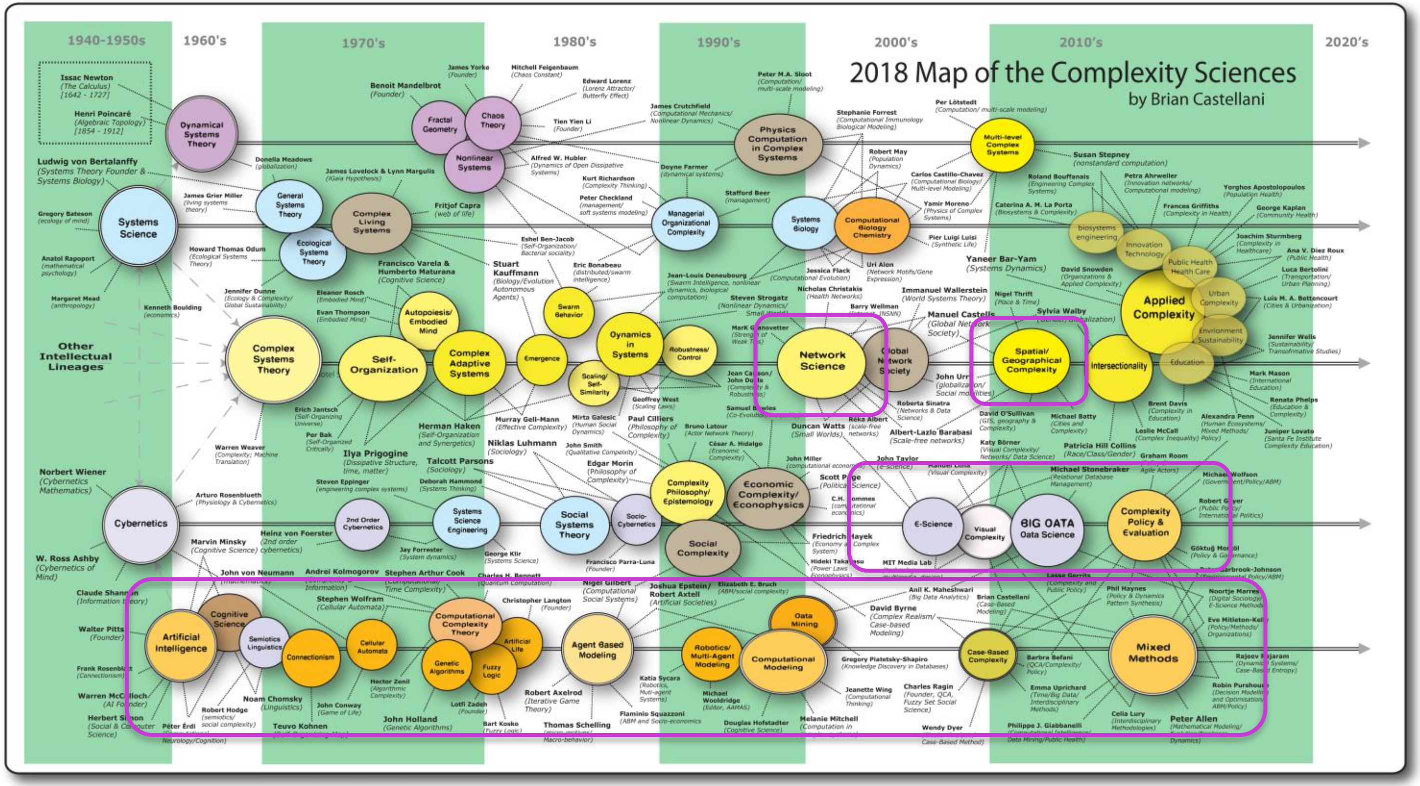
- Example: global network society (Manuel Castells)

Interested in how social complexity evolves across time.

- Focus is on dynamics, emergence, self-organisation, etc.

Interdisciplinary focused

There are many complexity oriented methods



So, what makes these methods so different?

What they are not	What they are
Not focused on variables Not focused on simple aggregate averages Not reductionist Not static Not focused on one-size-fits-all models, interventions or services	Use computation and machine intelligence • Hence, computational social science Focus on cases, interactions and process • What do cases have in common and their differences? • What explanations are consistent with the cases?

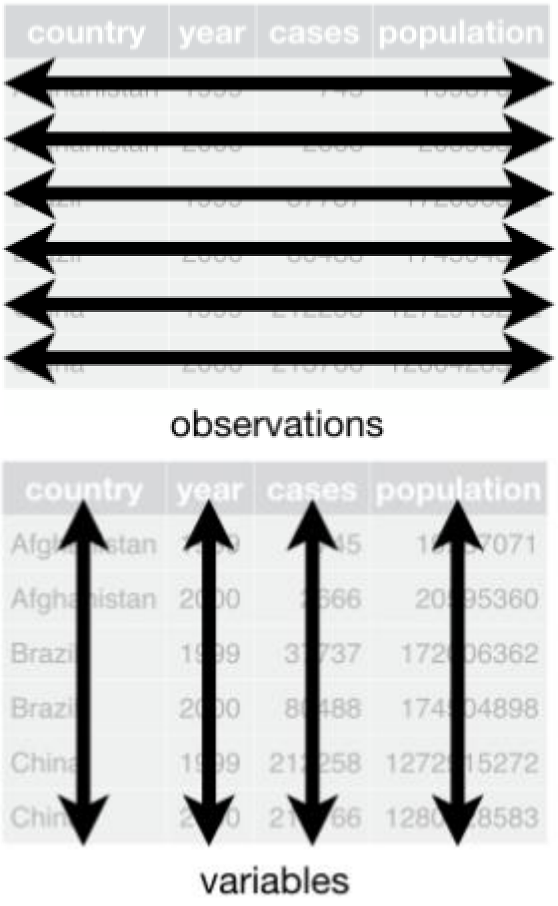
What is a variable?

Dataset contains values about

- Cases (observations)
- Values of attributes for each case (variables)

The same attributes are measured for each case

Each variable aggregates the values of one attribute for all cases



Wickham, *R for Data Science* (Figure 12.1)
https://r4ds.had.co.nz/tidy-data.html

What does it mean for two variables to be related?

Variables are associated is shorthand for "the values of two attributes across all cases have a pattern"

- Statistical models summarise cases
- Big values with big values and small values with small (or vice versa)

Correlation is not causation

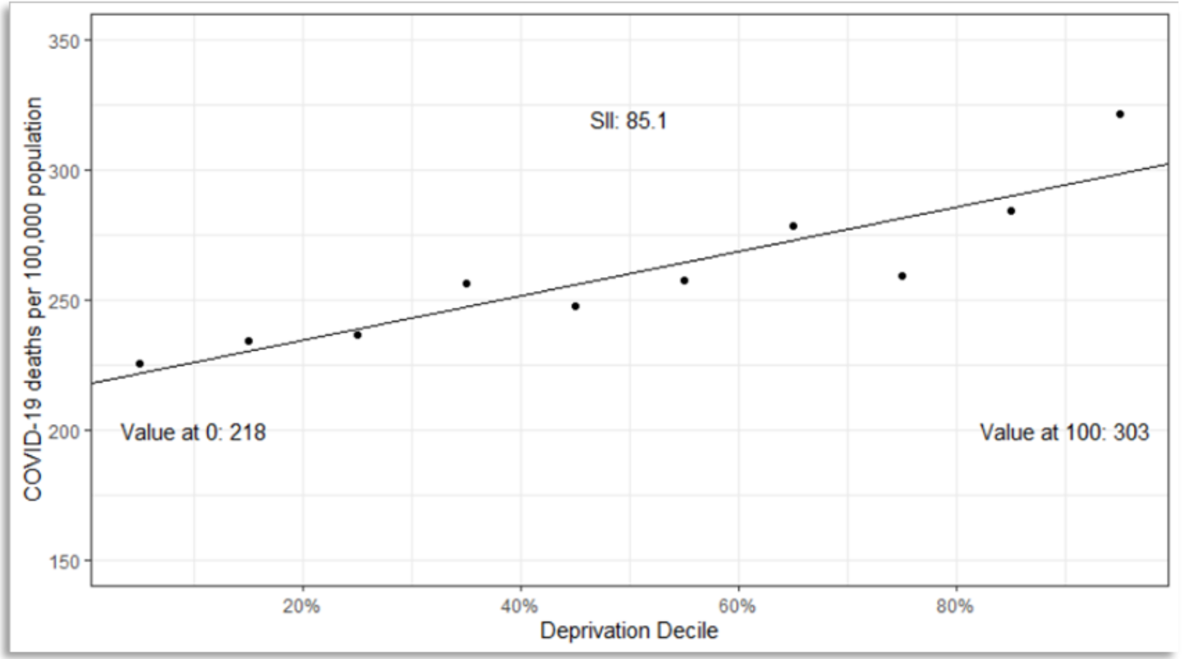
- Much repeated phrase
- Patterns are descriptive, not causative

Variable reports the state of one attribute of some case(s) at some point in time

- They are not 'real', they cannot exert effects
- Highlighting one attribute hides the many ways that the cases are different

Statistical models suggest what may be true for other cases based on the observed cases

What does it mean for two variables to be related?



Each data point represents a group of local areas and covid mortality is higher in more deprived areas

Durham University

Process is key in complex systems

Statistical models describe an observed pattern but do not explain why that pattern exists

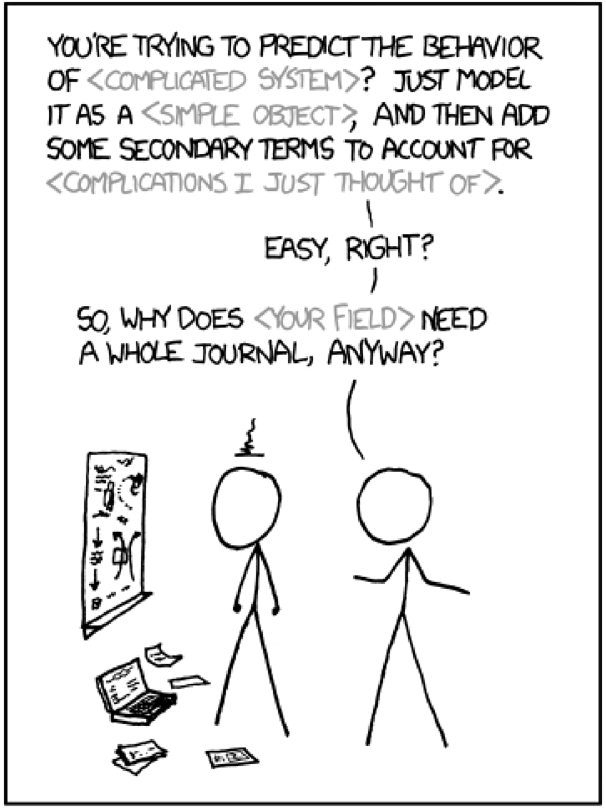
Social reality is organised

- Only by understanding ‘why’ can we make claims about cases that differ from those observed
  - Extrapolating beyond the data

Specific methods, not just ‘add stuff’

- Methods are complexity aware

Source: xkcd webcomic [http://xkcd.com/793/]



LIBERAL-ARTS MAJORS MAY BE ANNOYING SOMETIMES, BUT THERE'S *NOTHING* MORE OBNOXIOUS THAN A PHYSICIST FIRST ENCOUNTERING A NEW SUBJECT.

Data are not sufficient, no matter how big

EVERY DAY, MILLIONS of people use Google to dig up information that drives their daily lives, from how long their commute will be to how to treat their child's illness. This search data reveals a lot about the searchers: their wants, their needs, their concerns—extraordinarily valuable information. If these searches accurately reflect what is happening in people's lives, analysts could use this information to track diseases, predict sales of new products, or even anticipate the results of elections.

WIRED OPINION

ABOUT
David Lazer is a professor in the Department of Political Science and in the College of Computer and Information Sciences at Northeastern University. Ryan Kennedy is an associate professor of political science at the University of Houston.

In 2008, researchers from Google explored this potential, claiming that they could "nowcast" the flu based on people's searches. The essential idea, published in a paper in *Nature*, was that when people are sick with the flu, many search for flu-related information on Google, providing almost instant signals of overall flu prevalence. The paper demonstrated that search data, if properly tuned to the flu tracking information from the Centers for Disease Control and Prevention, could produce accurate estimates of flu prevalence two weeks earlier than the CDC's data—turning the digital refuse of people's searches into potentially life-saving insights.

BIG DATA

The Parable of Google Flu: Traps in Big Data Analysis

David Lazer,[1,2*] Ryan Kennedy,[1,3,4] Gary King,[3] Alessandro Vespignani[3,5,6]

In February 2013, Google Flu Trends (GFT) made headlines but not for a reason that Google executives or the creators of the flu tracking system would have hoped. *Nature* reported that GFT was predicting more than double the proportion of doctor visits for influenza-like illness (ILI) than the Centers for Disease Control and Prevention (CDC), which bases its estimates on surveillance reports from laboratories across the United States (*1*, *2*). This happened despite the fact that GFT was built to predict CDC reports. Given that GFT is often held up as an exemplary use of big data (*3*, *4*), what lessons can we draw from this error?

The problems we identify are

Durham University

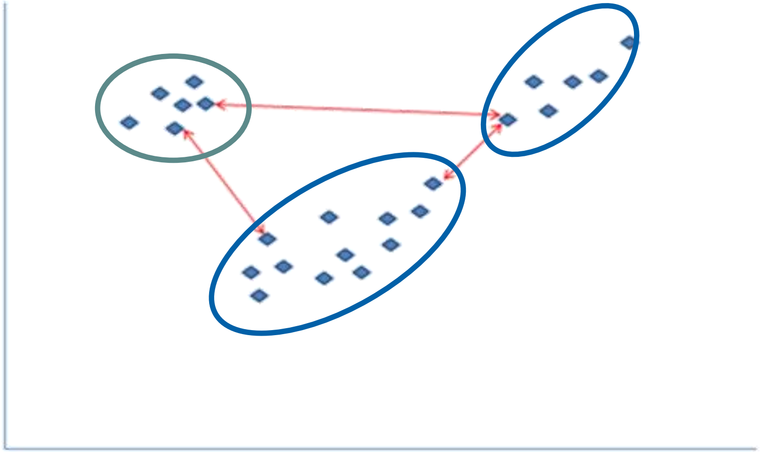
https://www.science.org/doi/10.1126/science.1248506

Clustering: Case-Based Complexity

Area Code	Income	Employment	Health						
	People in income deprivation (%)	Working-age people in employment deprivation (%)	GP-recorded chronic condition (rate per 100)	Limiting long-term illness (rate per 100)	Premature death (rate per 100,000)	GP-recorded mental health condition (rate per 100)	Cancer incidence (rate per 100,000)	Low birth weight (live single births less than 2.5kg) (%)	Children aged 4-5 who are obese (%)
	16	10	14.3	22.7	382.4	23.2	611.9	5.5	11.8
Isle of Anglesey	15	10	13.4	20.8	359.6	23.0	601.1	5.5	12.7
Gwynedd	13	8	12.9	19.5	348.9	20.3	604.0	5.0	13.0
Conwy	15	10	12.9	20.6	375.6	23.7	593.6	5.1	11.4
Denbighshire	17	11	14.7	21.8	397.4	28.4	639.3	6.1	12.2
Flintshire	12	8	14.1	19.7	358.1	23.1	647.8	5.4	11.2
Wrexham	15	9	14.3	21.5	393.7	24.3	637.3	6.4	12.4
Powys	11	7	12.8	18.8	309.1	19.0	579.8	4.7	10.5
Ceredigion	12	8	12.7	20.0	322.4	19.9	545.5	4.8	10.5
Pembrokeshire	15	10	13.1	20.5	345.8	22.1	606.1	5.2	12.5
Carmarthenshire	15	11	13.9	23.7	365.5	20.0	602.6	5.4	12.8

Case →

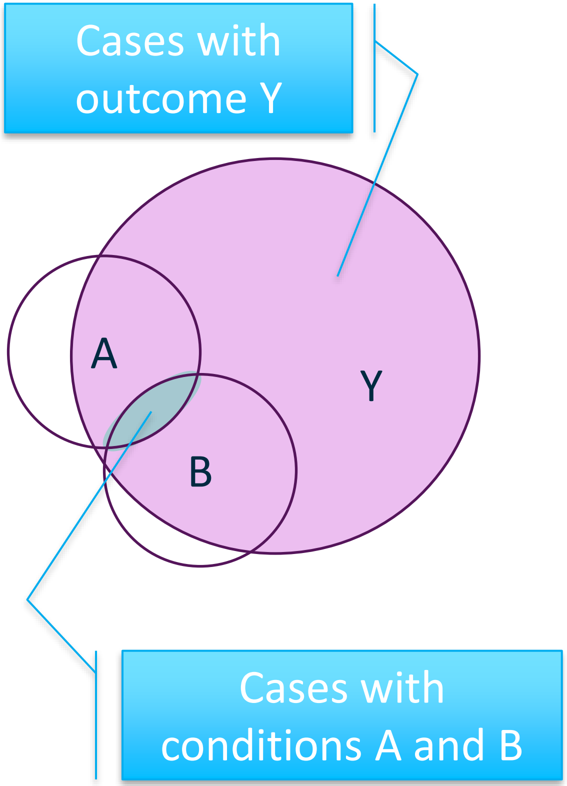
← Case Profile



Workshop case study: Criminology
Patterns of arrests in USA States

Durham University

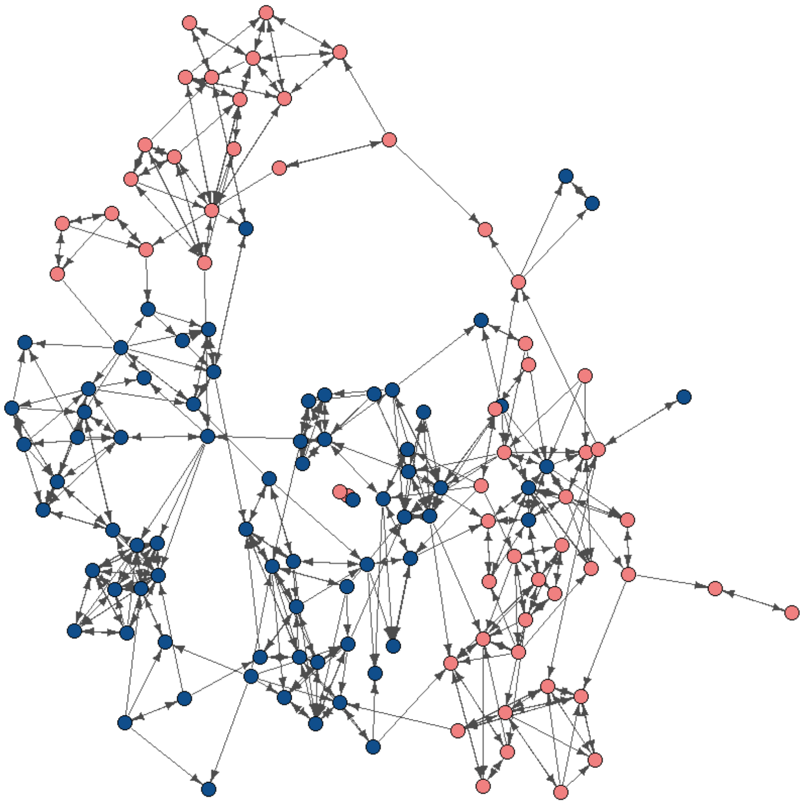
Qualitative Comparative Analysis: Configurational Complexity

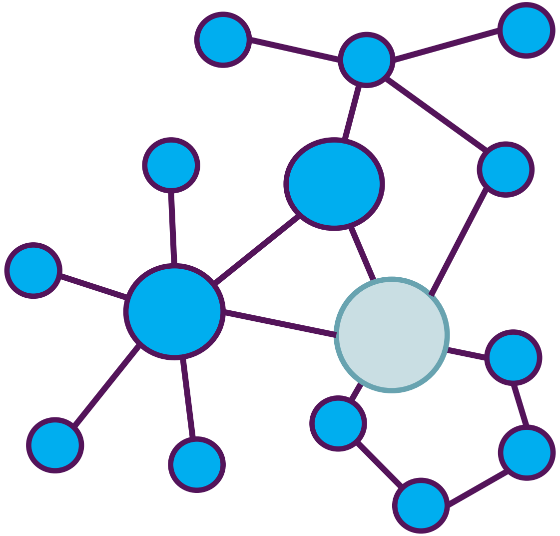


Conditions		Cases by outcome	
A	B	Y	Y
T	T	1	0
T	F	any	1
F	T	any	1
F	F	any	any

Workshop case study: Political Science
Explain survival/breakdown of democracy

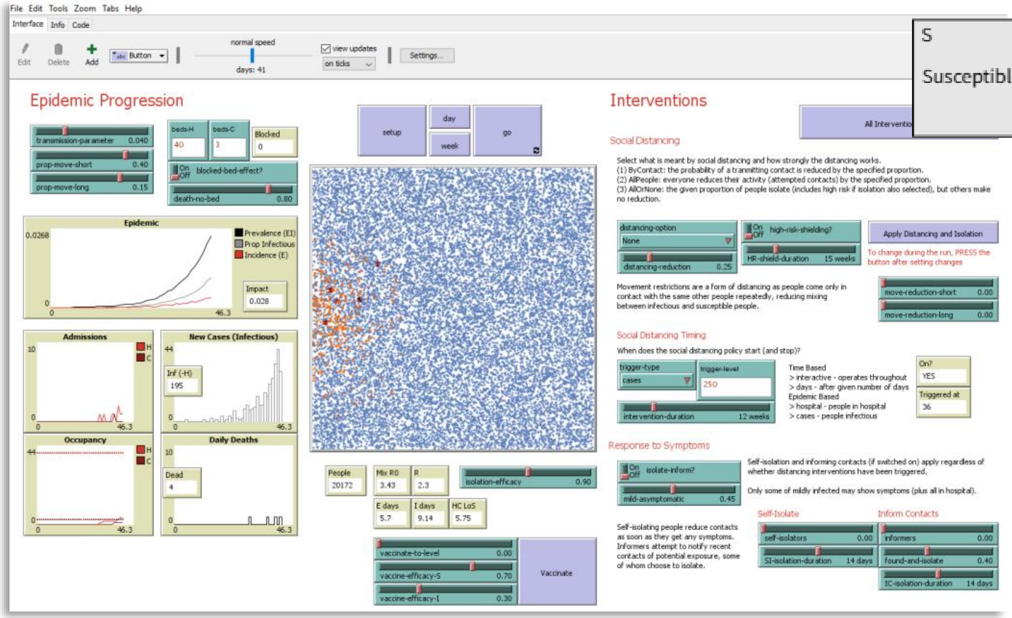
Social network analysis: Modelling relationships

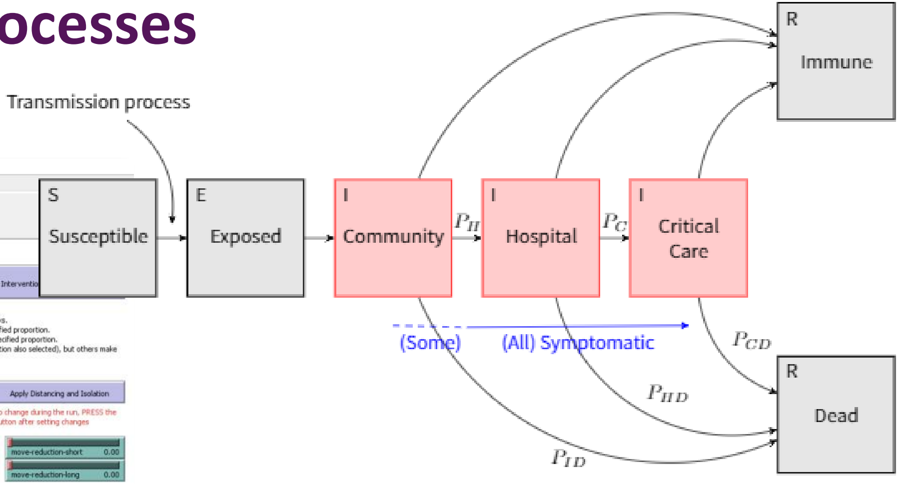




Workshop case study: Sociology
Diffusion of family planning practices

Social Simulation: Modelling processes





Workshop case study: Social Psychology
Conformity and segregation

Durham University

Theory and method work together

Computational Social Science is truly interdisciplinary

- Mathematical/computer scientists have analytical skills and training to develop methods and work with data
- Social scientists bring ideas, applications and interpretation
- Theory organises the ideas and interpretation

> As much as it bears similarities to previous work, however, what is new about the current generation of network-related research is a rapidly emerging and highly interdisciplinary synthesis of new analytical techniques, enormously greater computing power, and an unprecedented volume of empirical data. Sociologists have much to gain from this progress, and also much to contribute. Most of the work discussed in this review is taking place in the mathematical sciences, particularly in physics. That is hardly surprising, as physics and mathematics typically lead the way methodologically, whereas biological and social sciences follow with applications. But in this particular case, the flow of ideas has been bidirectional, with many of the core ideas—not just applications—having come from sociology. Physicists may be marvelous technicians, but they are mediocre sociologists. Thus, if the science of networks is to live up to its early promise, then the other disciplines—sociology in particular—must offer guidance in, for example, the interpretation of empirical and theoretical findings, particularly in the context of policy applications, and also in suggesting measures and models that are increasingly relevant to the important problems at hand.

Durham University

Watts (2004), The "New" Science of Networks, *Annual Review of Sociology*. 30:1, 243-270