Prepared by Julian García, based on material by
Arun Konagurthu

Lecture 1

Introduction to
Computational Modelling and Simulation

FIT 3139

Computational Modelling and Simulation

MONASH University
Information Technology

Teaching team

Julian García

- Room 206F, Woodside, CL
  Email: julian.garcia@monash.edu

Simon Teshuva

- Email: simon.teshuva@monash.edu

FIT3139.Clayton-x@monash.edu

Announcements

Unit Announcements

Discussions

Ed Discussion — class Q&A forum

Use this
unless personal matter

Computational science

From Wikipedia, the free encyclopedia

Not to be confused with computer science.

Computational science (also **scientific computing** or **scientific computation**) is concerned with constructing mathematical models and quantitative analysis techniques and using computers to analyze and solve scientific problems.[1] In practical use, it is typically the application of computer simulation and other forms of computation from numerical analysis and theoretical computer science to problems in various scientific disciplines.

Modelling and simulation

From *SIAM News*, Volume 46, Number 10, December 2013

The 2013 Nobel Prize in Chemistry Celebrates Computations in Chemistry and Biology

By Tamar Schlick

Progress in science depends on new techniques, new discoveries and new ideas, probably in that order.

—Sidney Brenner (2002 Nobel laureate in physiology or medicine)

It was an exciting day for the fields of computational chem ... early on
the morning of October 9, 2013. Awarded to Martin Karplus, ... for com-
plex chemical systems," the prize provided a welcome "seal ... not unlike
the traditional division between pure and applied mathemati ... problems
versus *approximating* complex physical systems to solve *rea* ... l systems
were judged by their ability to *reproduce* experimental data. T ... o employ
direct, though not error-free, measurements, especially for la ... onsidered
uninformative, but if they predicted properties ahead of expe

Photo: A. Mahmoud

Martin Karplus

Prize share: 1/3

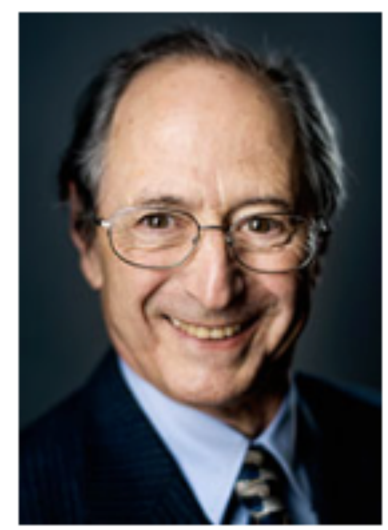


Photo: A. Mahmoud

Michael Levitt

Prize share: 1/3

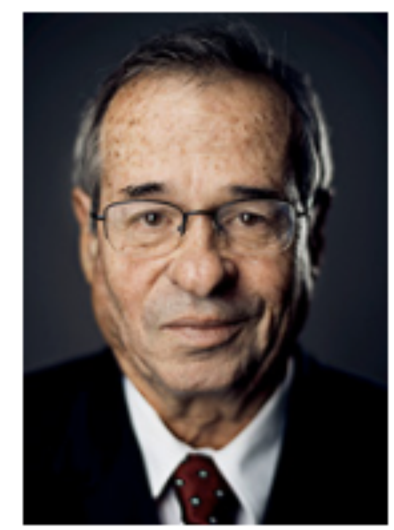


Photo: A. Mahmoud

Arieh Warshel

Prize share: 1/3

Scientific Method

- Observe event.
- Formulate questions.
- Propose hypothesis.
- Make prediction using the hypothesis.
- Test hypothesis.
- Analyse results.

Scientific Method

Statistical Models

Hypothesise - Model - Test

Mathematical Models

Models

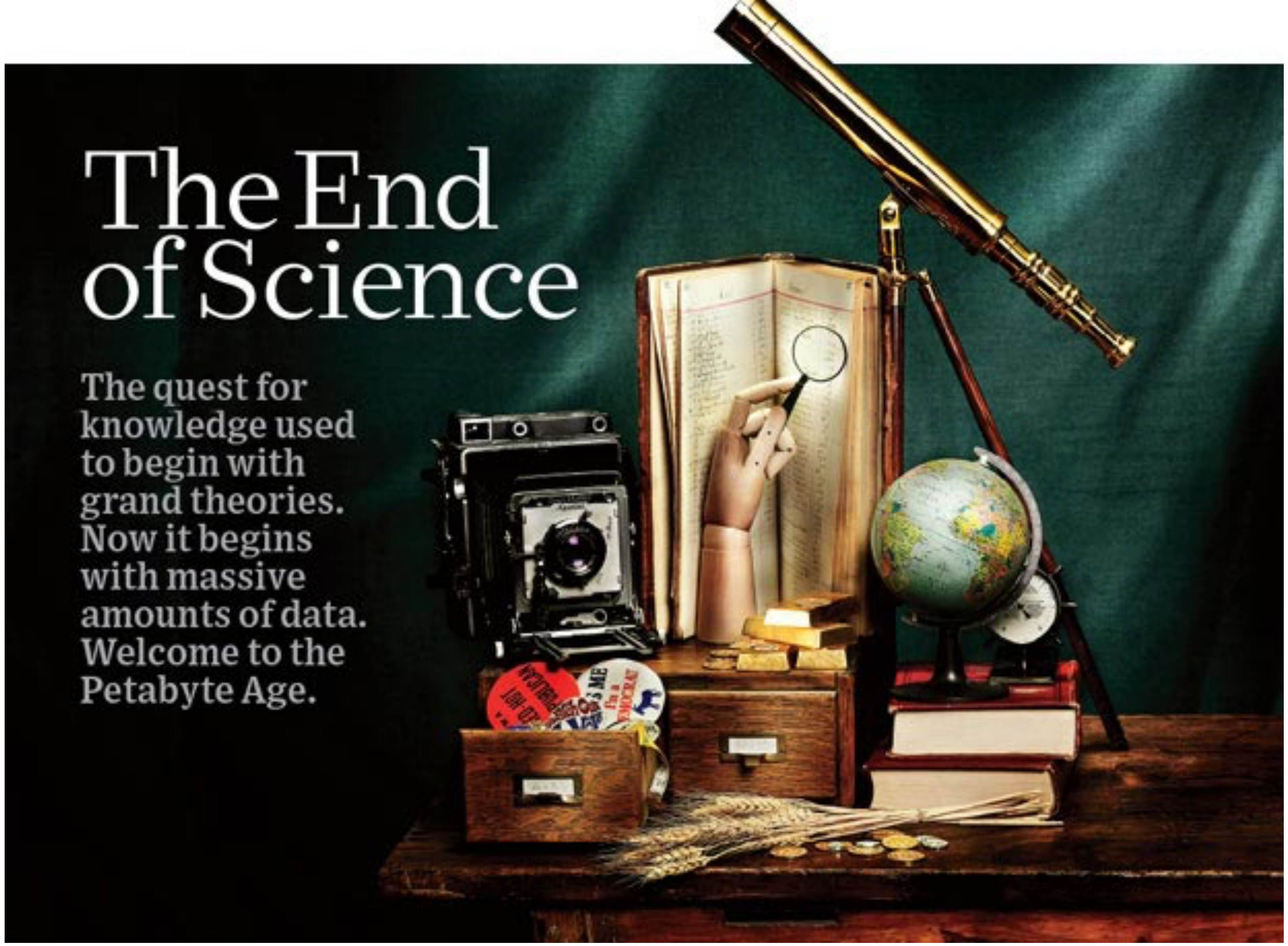
Mathematical models:

If the world would be like this
(assumption A, assumption B, etc)
then....

Statistical models:

Explaining data
(e.g. testing two competing models).

Predicting outcomes based on existing data.



Source: Wired Magazine - http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory

"All models are wrong, and increasingly you can succeed without them."

wrongly attributed to Peter Norving.

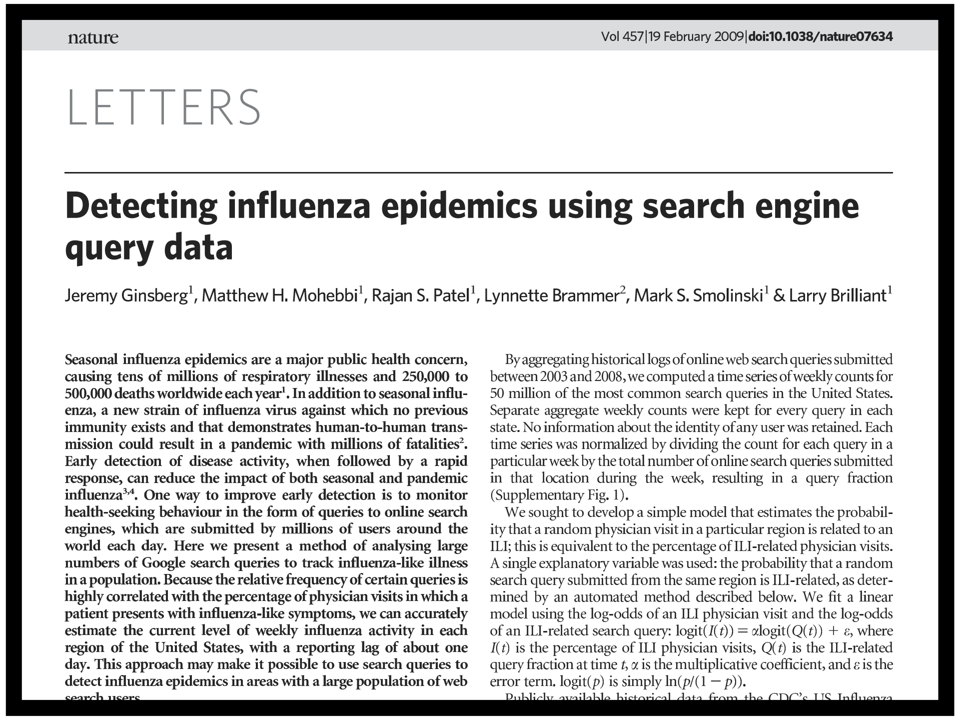
“That's a silly statement, I didn't say it, and I disagree with it. ”

https://norvig.com/fact-check.html

Google Flu Trends

https://www.google.org/flutrends/au/#AU

Ginsberg, Jeremy, et al. "Detecting influenza epidemics using search engine query data." *Nature* 457.7232 (2009): 1012-1014.



nature Vol 457|19 February 2009|doi:10.1038/nature07634

LETTERS

Detecting influenza epidemics using search engine query data

Jeremy Ginsberg[1], Matthew H. Mohebbi[1], Rajan S. Patel[1], Lynnette Brammer[2], Mark S. Smolinski[1] & Larry Brilliant[1]

Seasonal influenza epidemics are a major public health concern, causing tens of millions of respiratory illnesses and 250,000 to 500,000 deaths worldwide each year[1]. In addition to seasonal influenza, a new strain of influenza virus against which no previous immunity exists and that demonstrates human-to-human transmission could result in a pandemic with millions of fatalities[2]. Early detection of disease activity, when followed by a rapid response, can reduce the impact of both seasonal and pandemic influenza[3,4]. One way to improve early detection is to monitor health-seeking behaviour in the form of queries to online search engines, which are submitted by millions of users around the world each day. Here we present a method of analysing large numbers of Google search queries to track influenza-like illness in a population. Because the relative frequency of certain queries is highly correlated with the percentage of physician visits in which a patient presents with influenza-like symptoms, we can accurately estimate the current level of weekly influenza activity in each region of the United States, with a reporting lag of about one day. This approach may make it possible to use search queries to detect influenza epidemics in areas with a large population of web search users.

By aggregating historical logs of online web search queries submitted between 2003 and 2008, we computed a time series of weekly counts for 50 million of the most common search queries in the United States. Separate aggregate weekly counts were kept for every query in each state. No information about the identity of any user was retained. Each time series was normalized by dividing the count for each query in a particular week by the total number of online search queries submitted in that location during the week, resulting in a query fraction (Supplementary Fig. 1).

We sought to develop a simple model that estimates the probability that a random physician visit in a particular region is related to an ILI; this is equivalent to the percentage of ILI-related physician visits. A single explanatory variable was used: the probability that a random search query submitted from the same region is ILI-related, as determined by an automated method described below. We fit a linear model using the log-odds of an ILI physician visit and the log-odds of an ILI-related search query: $\mathrm{logit}(I(t)) = \alpha\,\mathrm{logit}(Q(t)) + \varepsilon$, where $I(t)$ is the percentage of ILI physician visits, $Q(t)$ is the ILI-related query fraction at time t, α is the multiplicative coefficient, and ε is the error term. $\mathrm{logit}(p)$ is simply $\ln(p/(1-p))$.

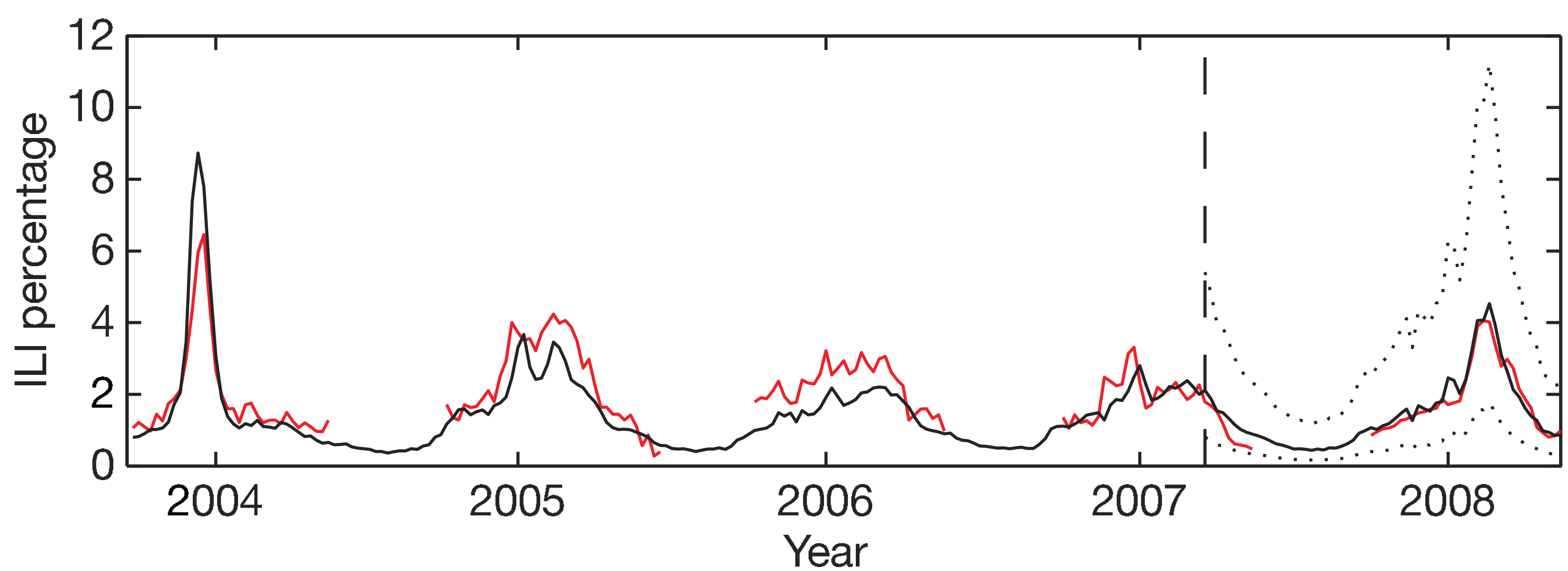
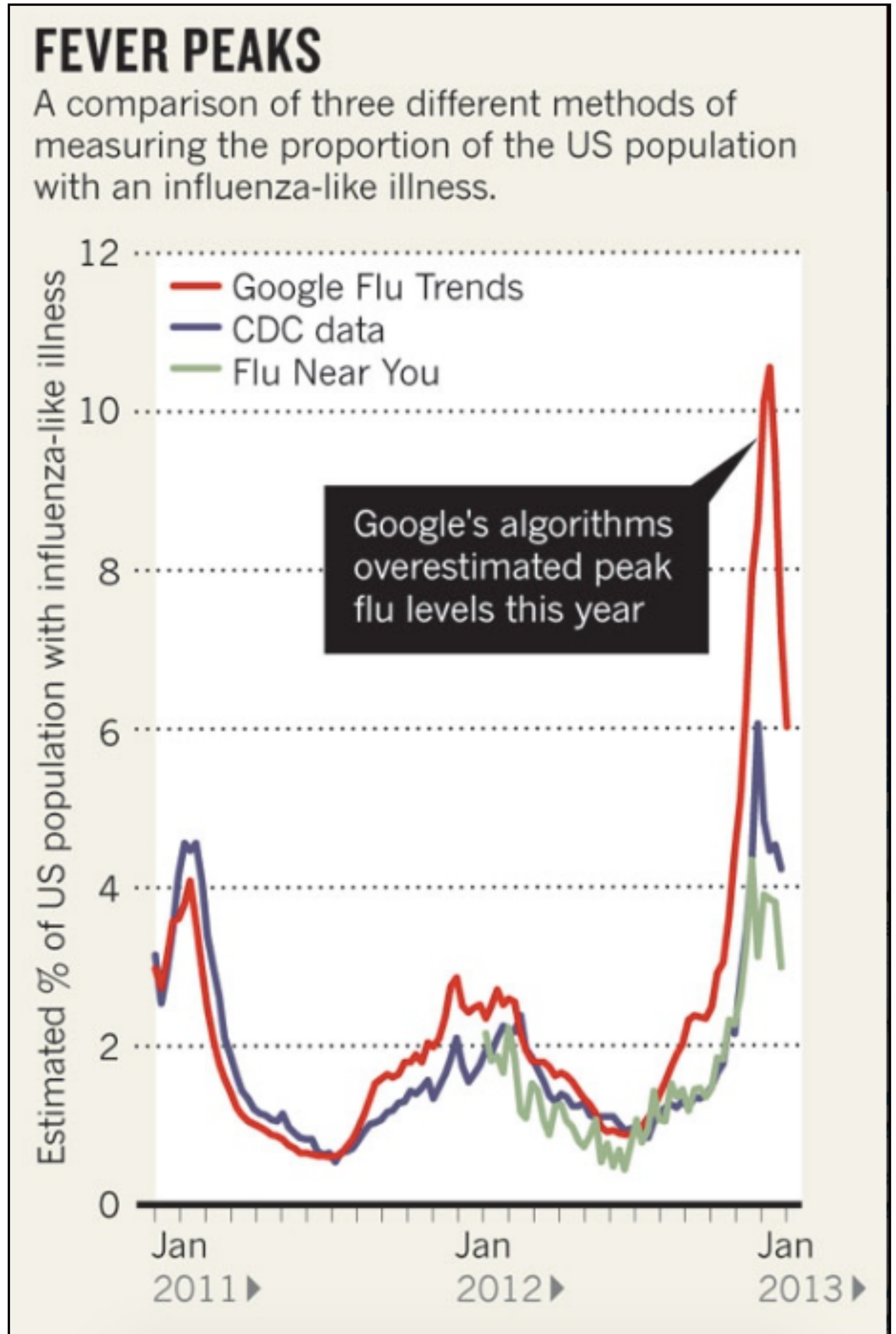


Figure 2 | A comparison of model estimates for the mid-Atlantic region (black) against CDC-reported ILI percentages (red), including points over which the model was fit and validated. A correlation of 0.85 was obtained over 128 points from this region to which the model was fit, whereas a correlation of 0.96 was obtained over 42 validation points. Dotted lines indicate 95% prediction intervals. The region comprises New York, New Jersey and Pennsylvania.

Butler, Declan. "When Google got flu wrong." *Nature* 494.7436 (2013): 155.



Big Data, are we making a big mistake:
http://www.ft.com/intl/cms/s/2/21a6e7d8-b479-11e3-a09a-00144feabdc0.html#axzz3h3L4uJxD

Scientific Method

Statistical Models

Hypothesise - Model - Test

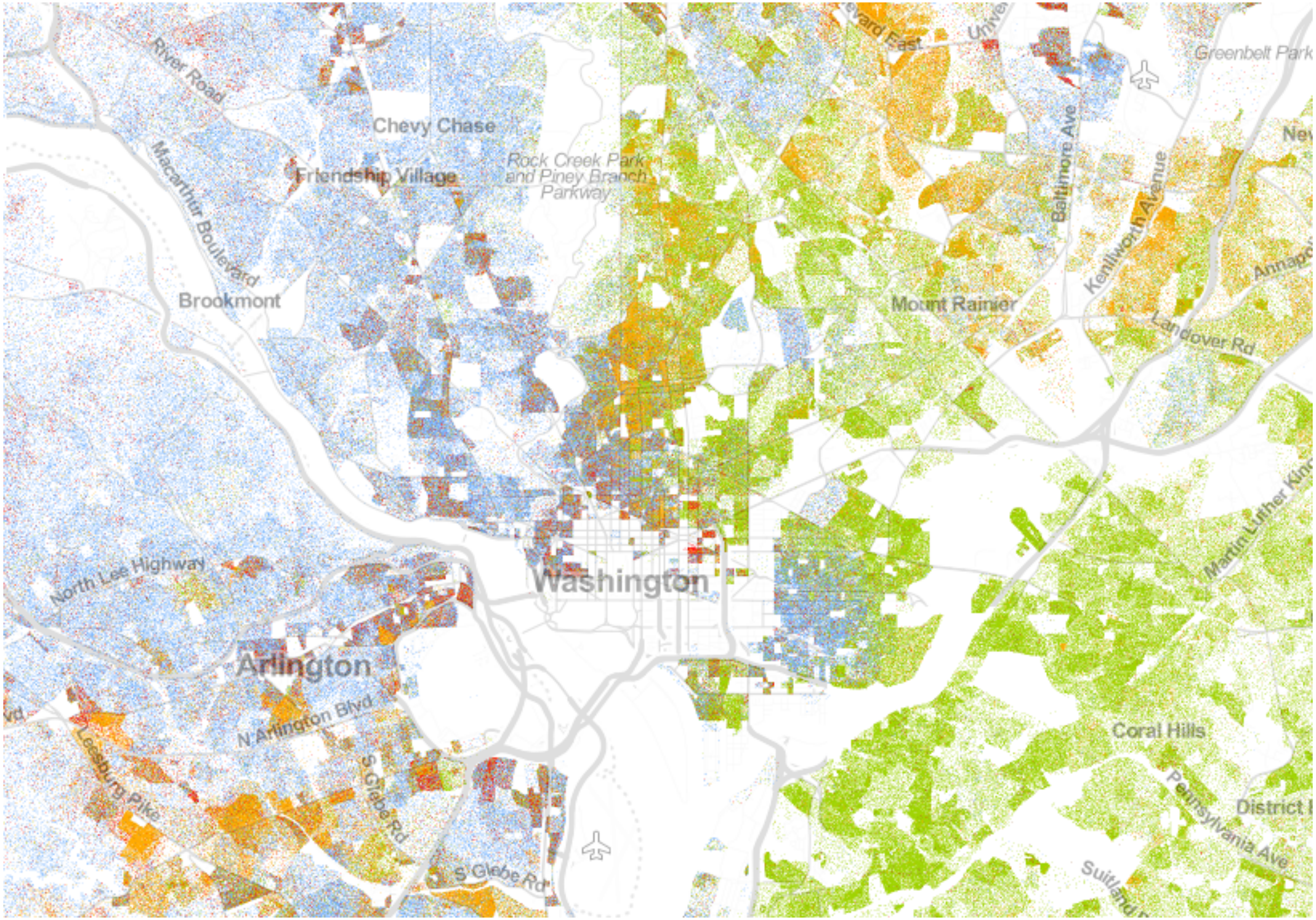
Mathematical Models

Computational Models

Science.

- Observation
- Theoretical Reasoning
- Experiment
- Computation

Washington

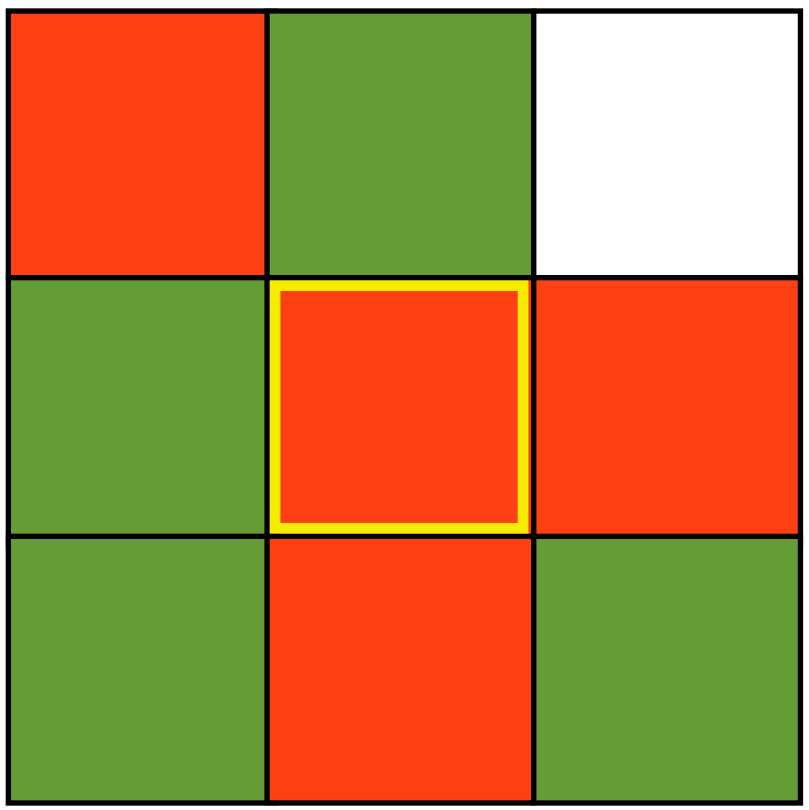

River Road
Chevy Chase
Greenbelt Park
Rock Creek Park
and Piney Branch
Parkway
Friendship Village
Macarthur Boulevard
Baltimore Ave
Kenilworth Avenue
Brookmont
Mount Rainier
Landover Rd
North Lee Highway
Washington
Martin Luther King
Arlington
N Arlington Blvd
Leesburg Pike
S Glebe Rd
S Glebe Rd
Coral Hills
Pennsylvania Ave
Suitland

Why are cities often segregated?

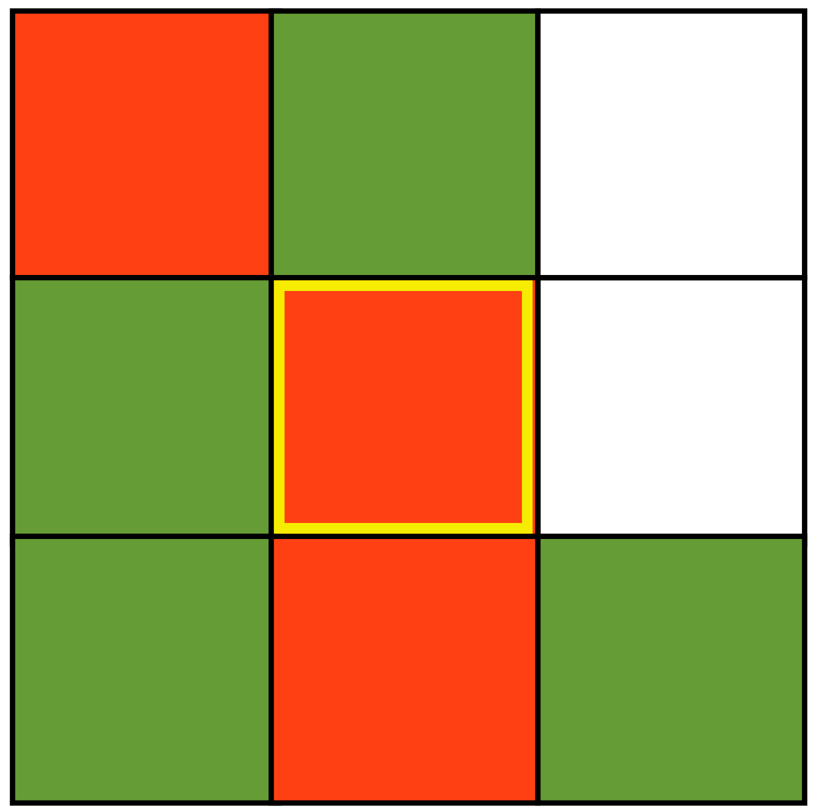
Science.

- Observation
- Theoretical Reasoning
- Experiment
- Computation

Should I stay or should I go



3/7



2/6

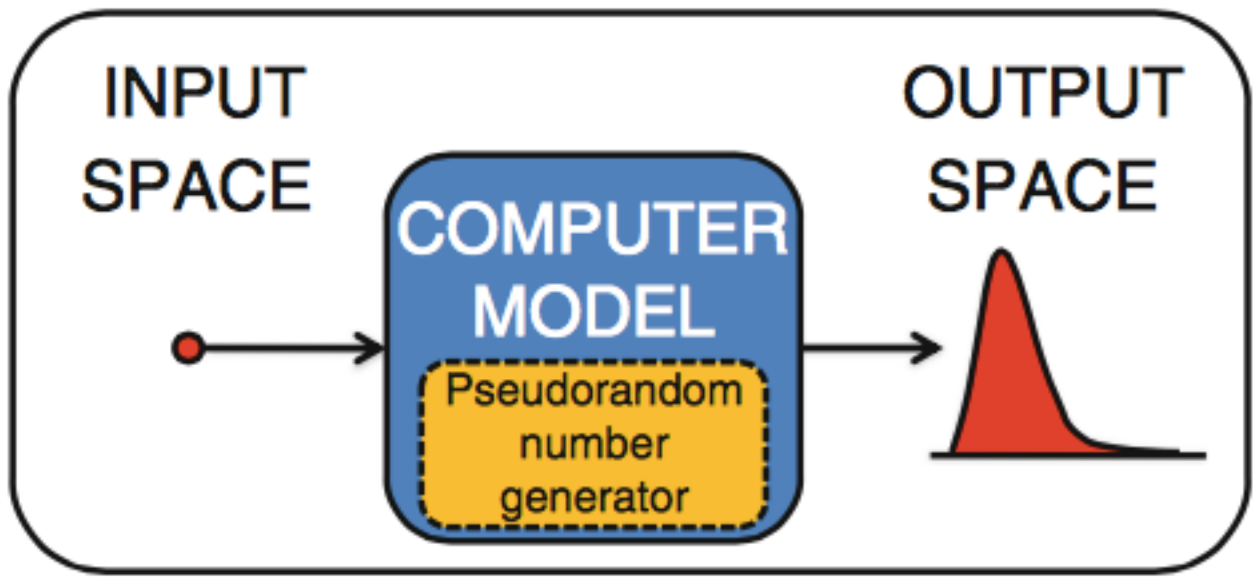
Dynamic Models of Segregation (1971)

Schelling's segregation model.

- A number of agents are distributed on a grid (with more spots than agents).
- Types: Green and Red agents (50% each).
  Neighbour preferences ($0.0 < h < 1.0$)
- States: Happy and Unhappy.
- Happy if alike neighbors $\geq$ h.
- If happy, stay. If unhappy move to a random vacant spot.

Science.

- Observation
- Theoretical Reasoning
- Experiment
- Computation



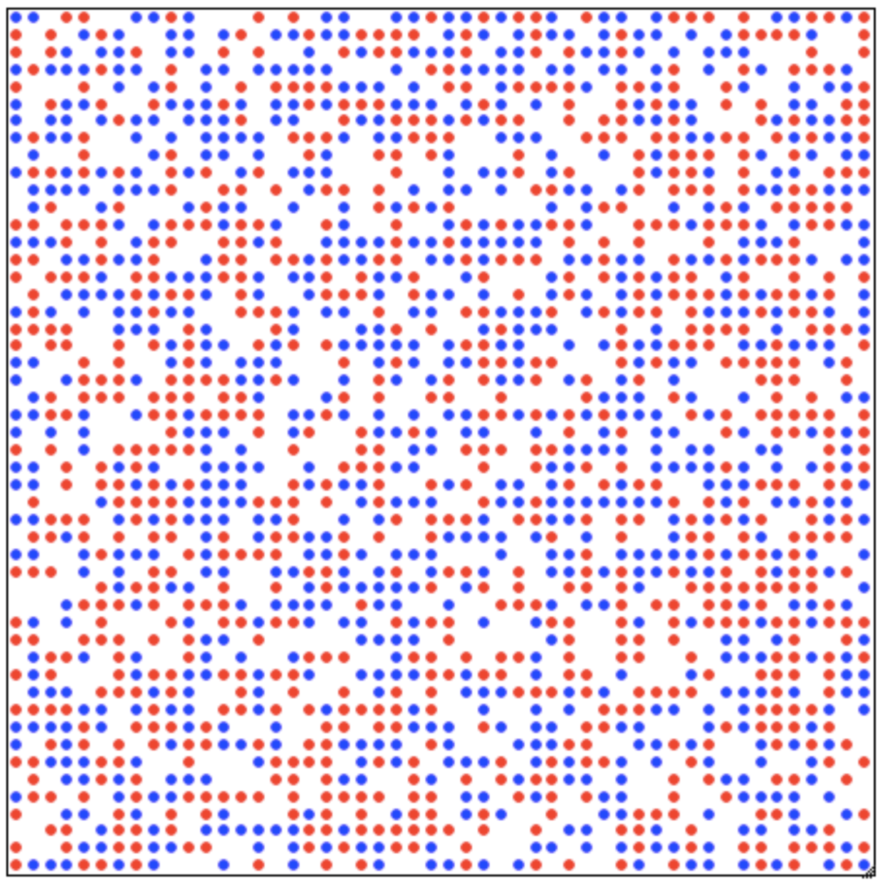
Izquierdo, Luis R., et al. "Combining Mathematical and Simulation Approaches to Understand the Dynamics of Computer Models." *Simulating Social Complexity*. Springer Berlin Heidelberg, 2013. 235-271.

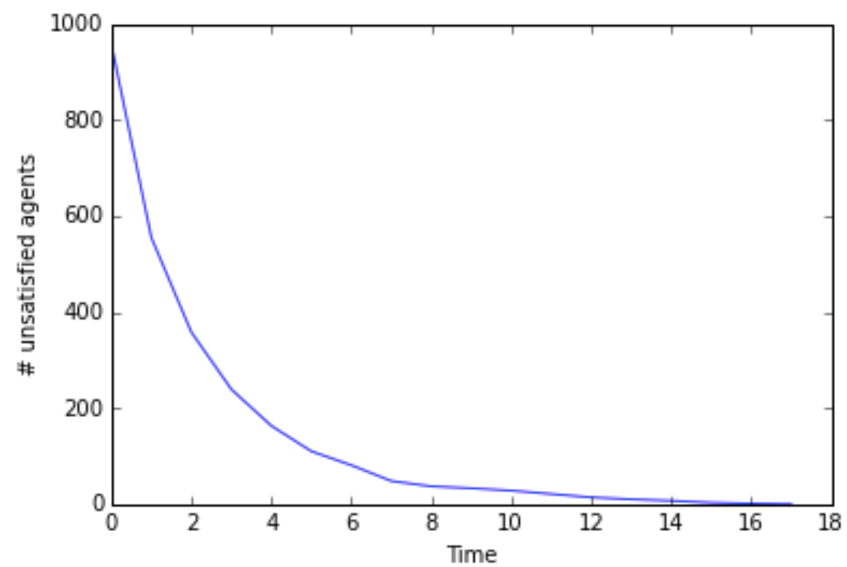
Schelling’s segregation model.

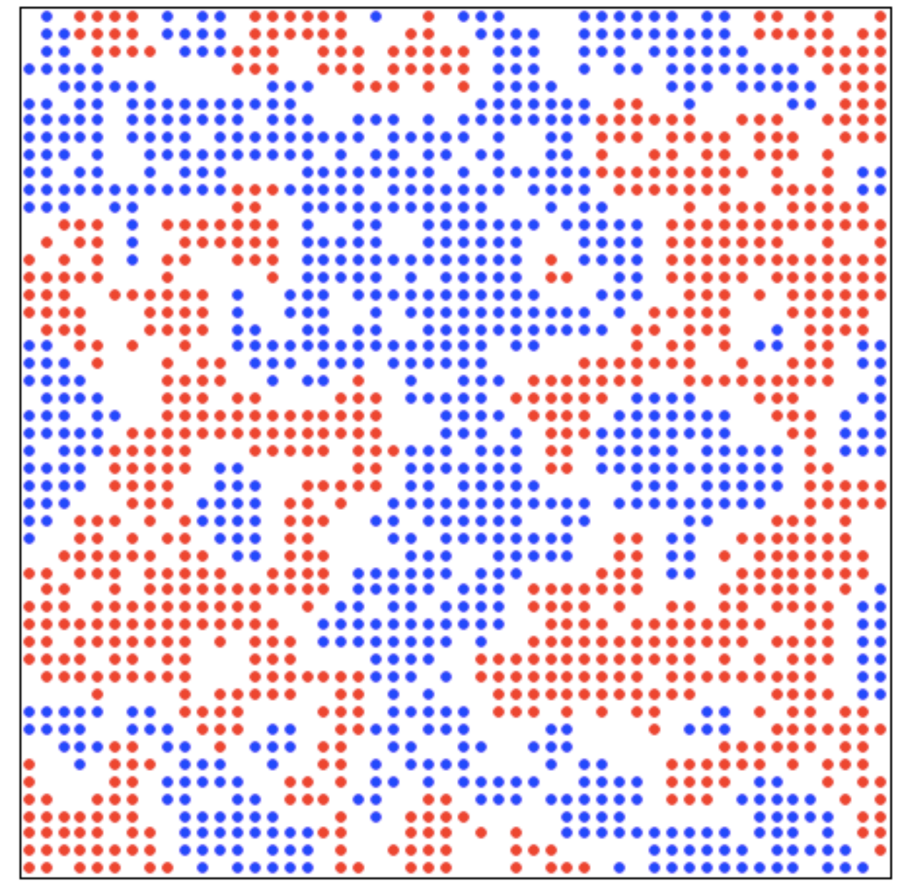
```
place agents on the grid.
for each agent, calculate happiness.
while (maximum number of iterations not reached)
do
        for each unhappy agent:
                    find a new location randomly.
        for each agent:
            calculate happiness.
        if (everyone is happy):
            return
end
```

Questions you could ask from the model?

Demo.

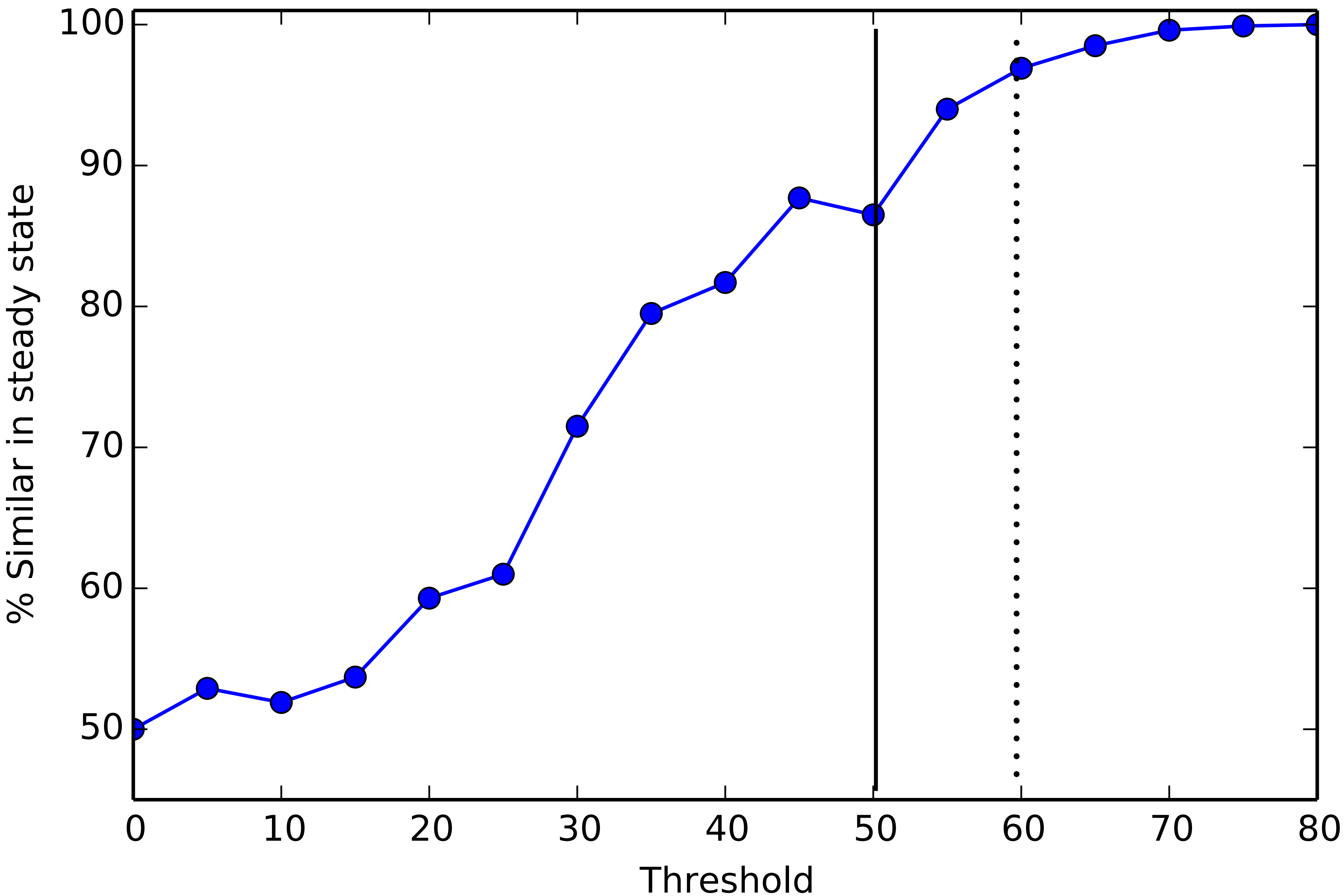



1000
800
600
400
200
0
unsatisfied agents
0
2
4
6
8
10
12
14
16
18
Time



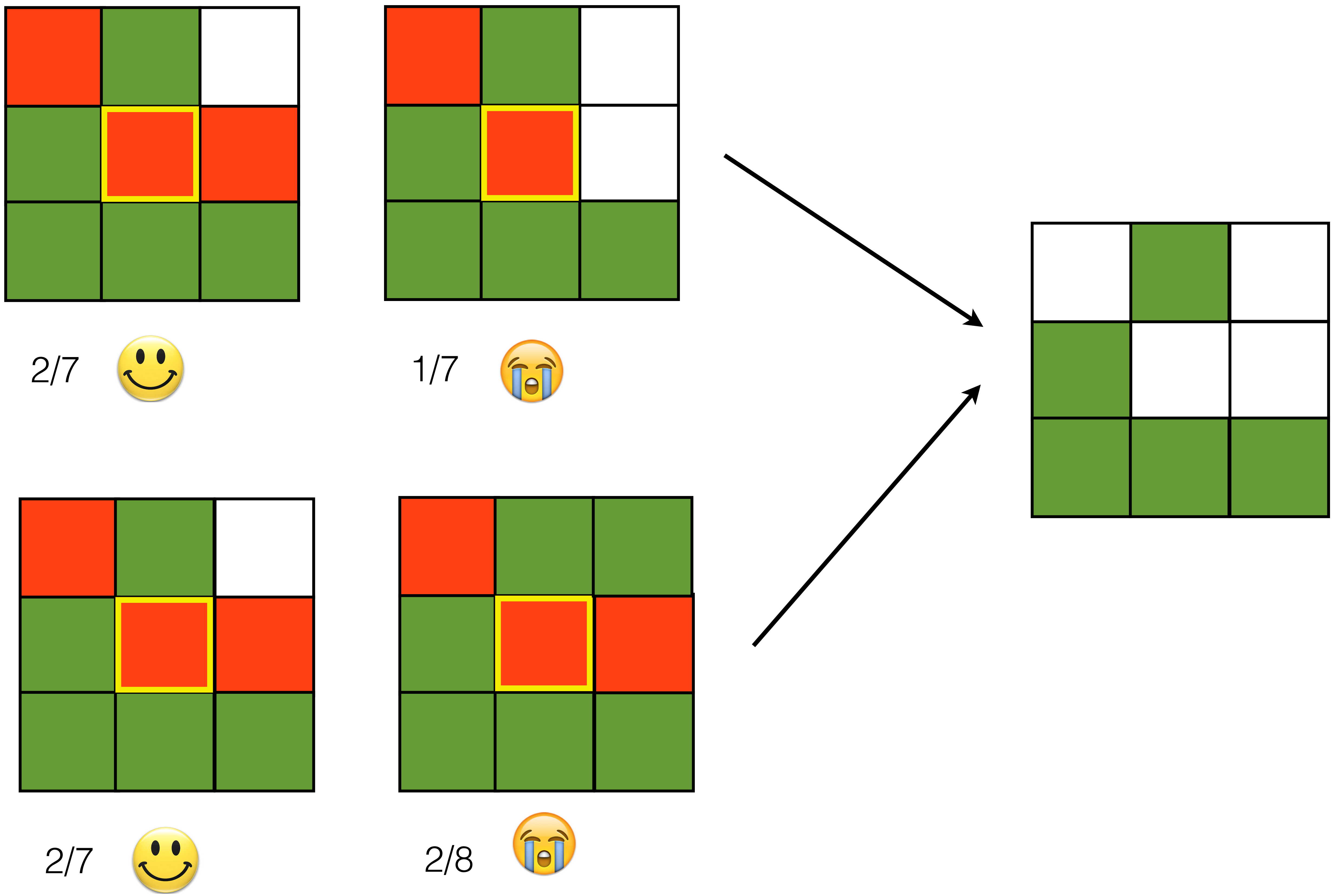
Main results.

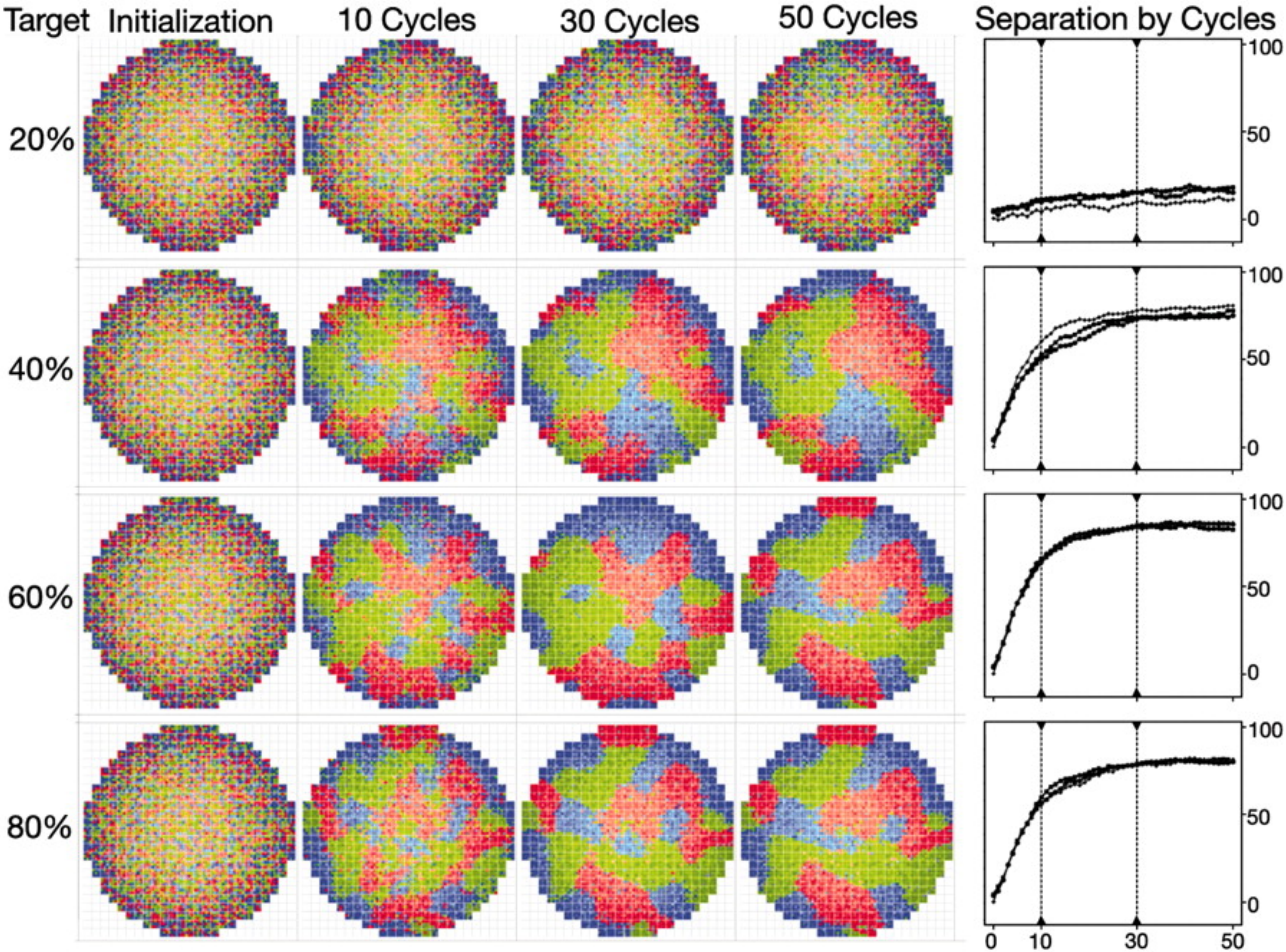
- Only a slight threshold bias leads to wholesale segregation.
- Large thresholds need not lead to segregation.
- Micromotives $\neq$ Macrobehaviors
- This is consistent across different variations of the model and for a wide range of parameters.



100
90
80
70
60
50
0
10
20
30
40
50
60
70
80
% Similar in steady state
Threshold

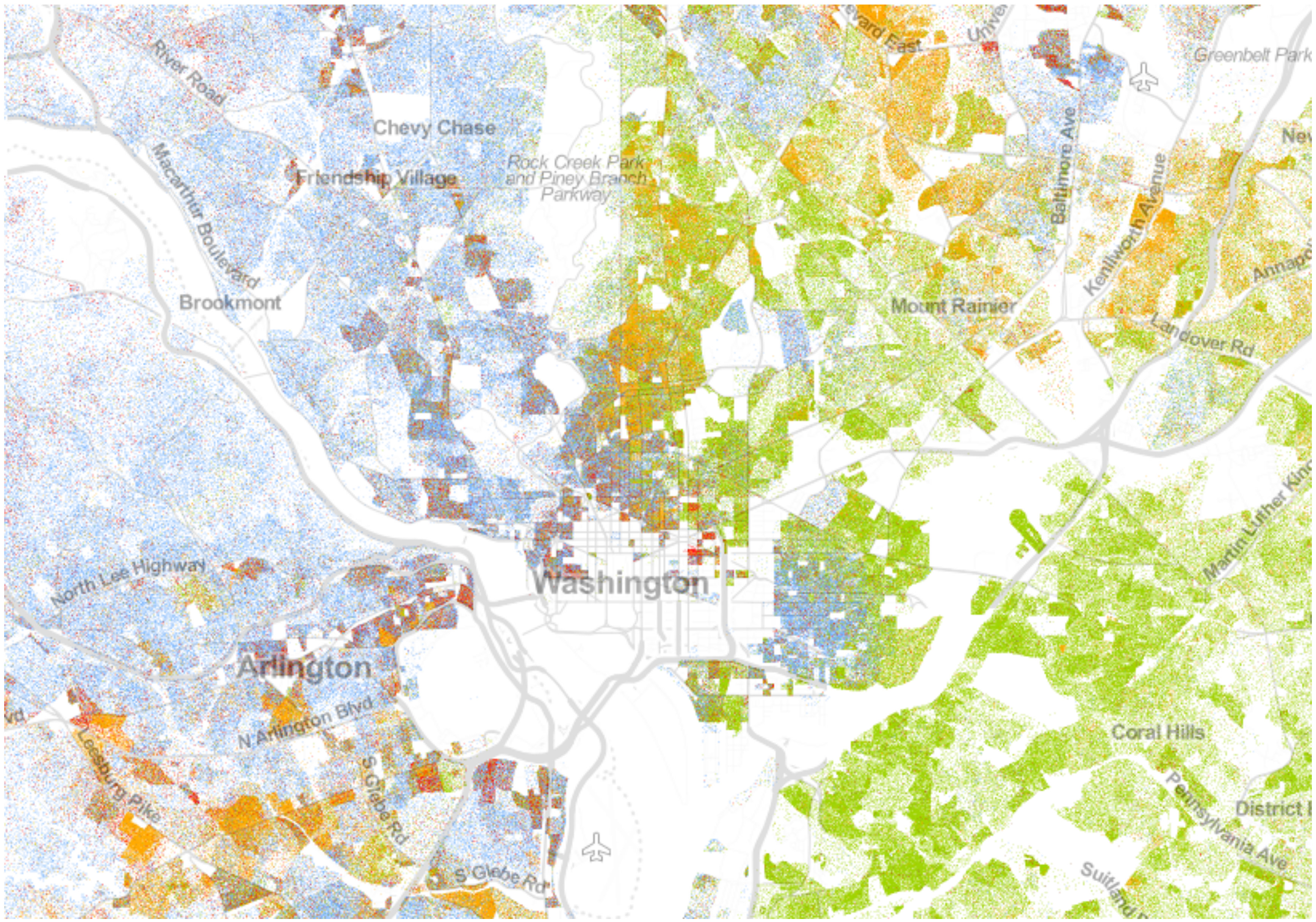
Tipping: Small decisions, big effects





Understanding the social context of the Schelling segregation model
William A. V. Clark and Mark Fossett
PNAS March 18, 2008 105 (11) 4109-4114

A new perspective…



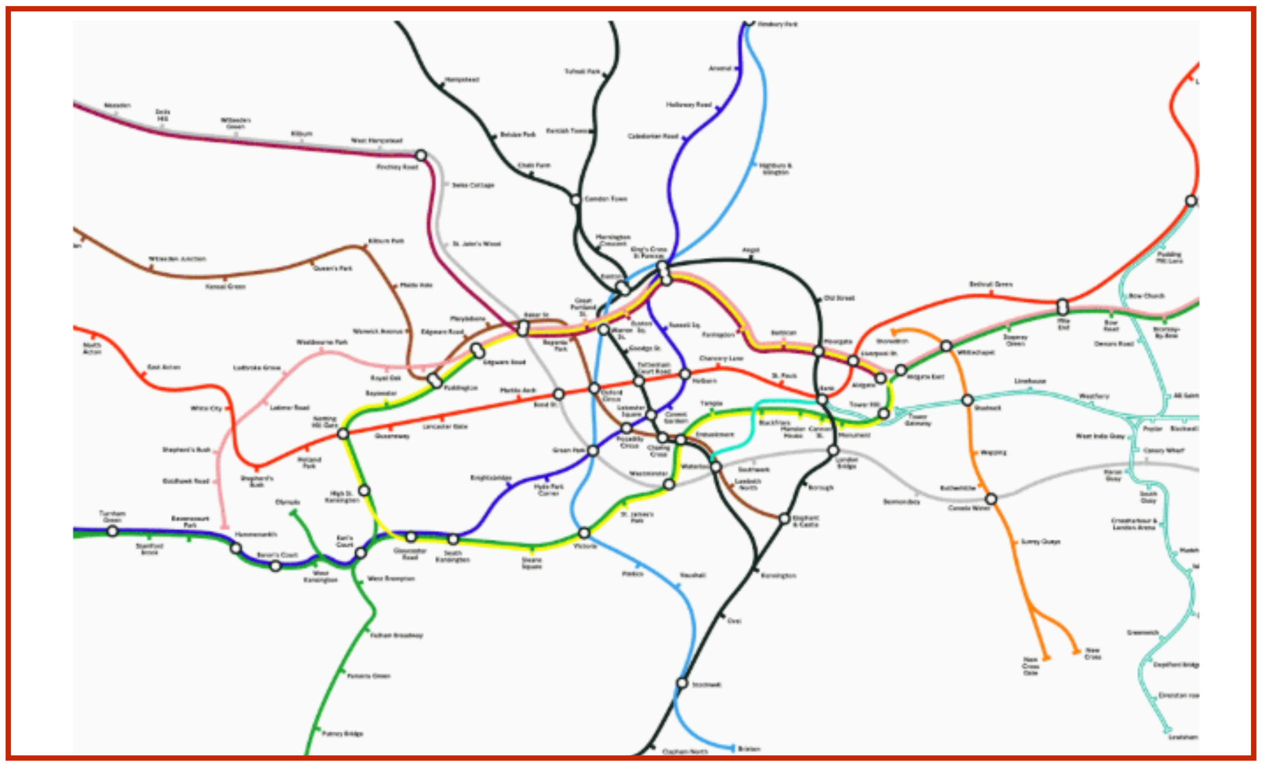
What people want as individuals
may not lead to what they want as a group.

"**All models are wrong**, but some are useful."

Occam's razor:
due to William of Ockham (c. 1287–1347).

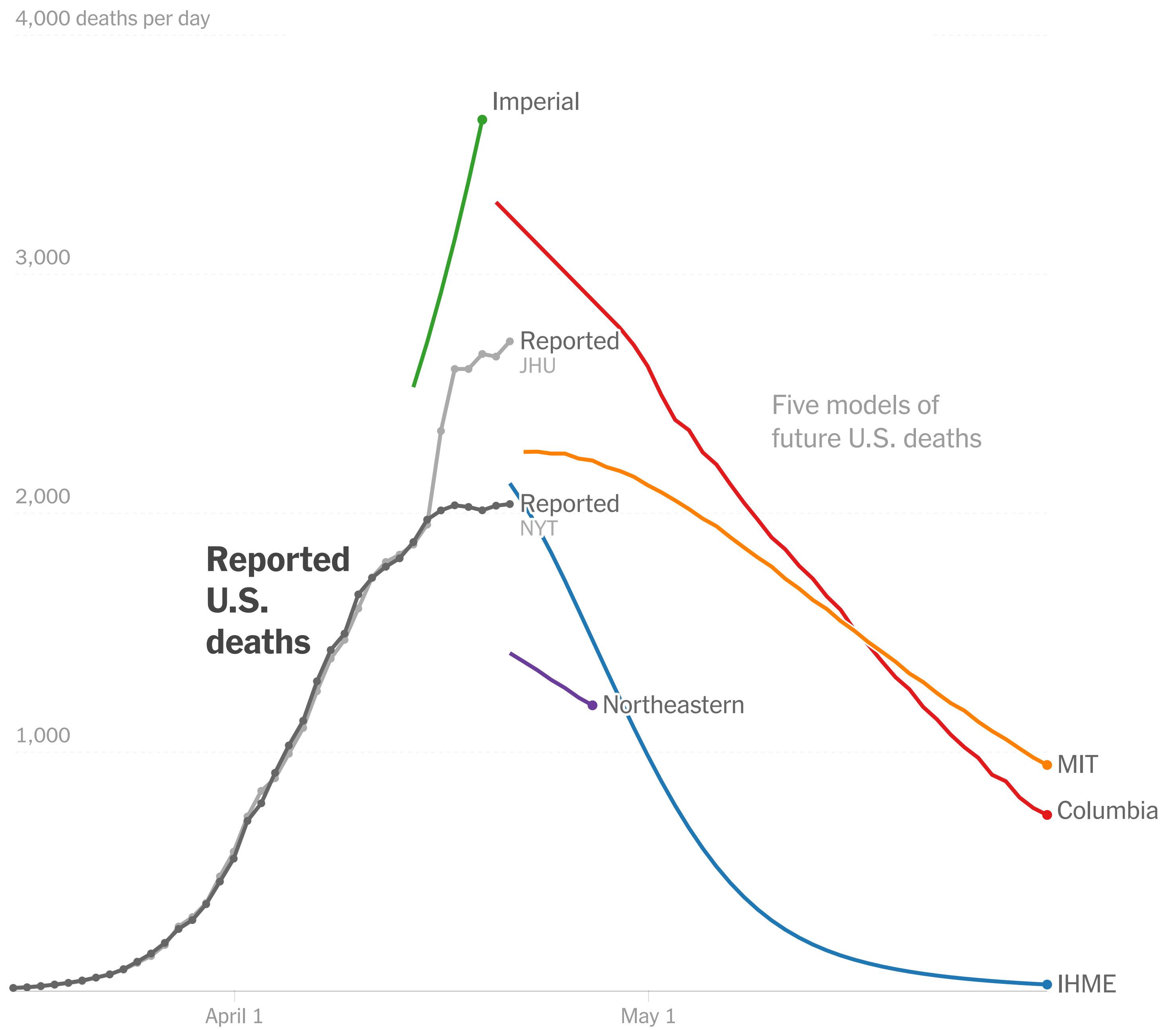
Among competing hypotheses, the one with the fewest assumptions should be selected.
Other, more complicated solutions may ultimately prove correct, but—in the absence of certainty—
the fewer assumptions that are made, the better.

Models are by definition, simple



https://www.timeout.com/london/things-to-do/23-alternative-tube-maps

What 5 Coronavirus Models Say the Next Month Will Look Like



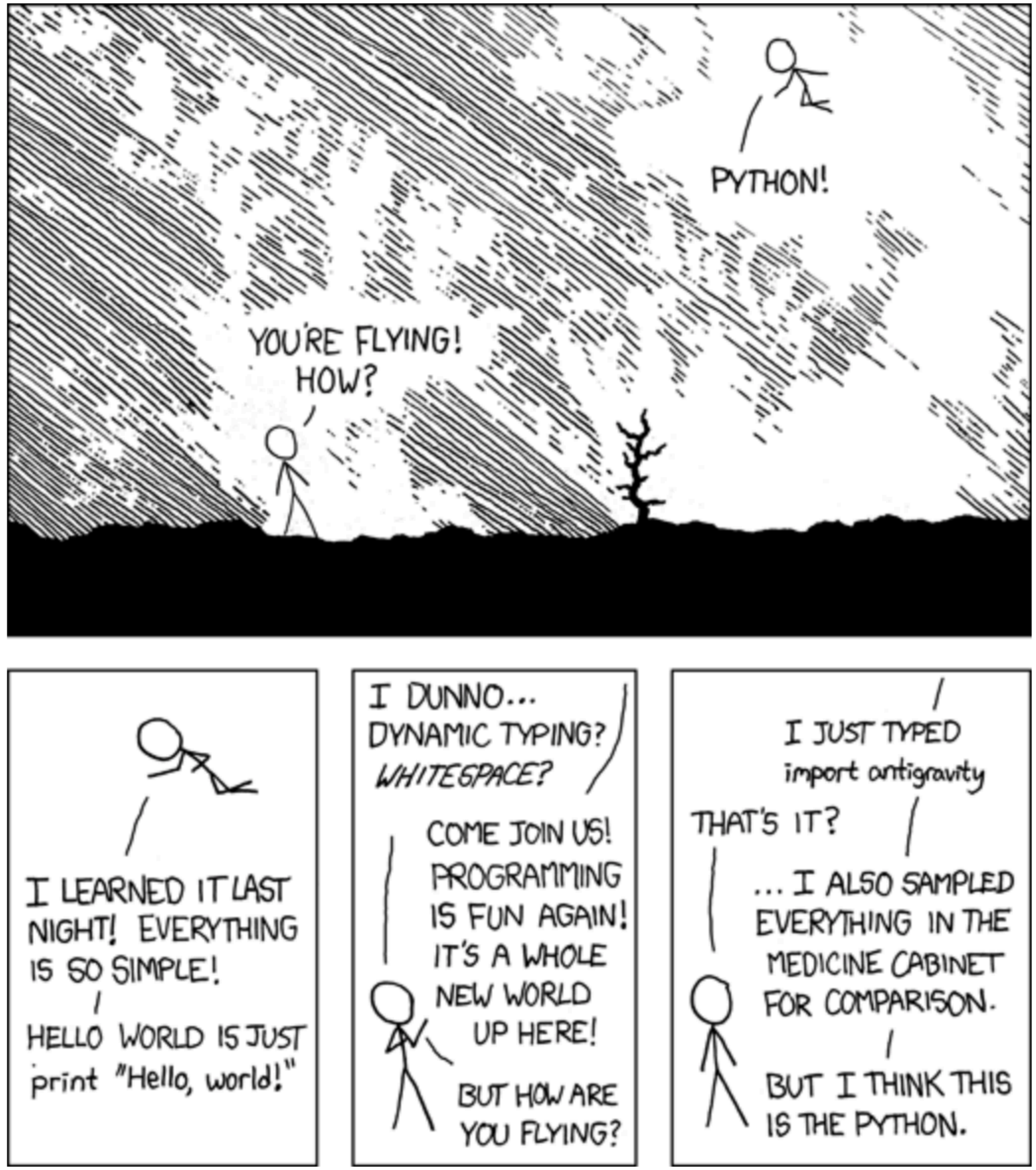
Reported deaths are rolling 7-day averages. Lines differ on whether to include roughly 5,000 probable deaths in New York City.

Latest model projections for Northeastern, I.H.M.E. and M.I.T. are April 21; Columbia is April 19; Imperial is April 13.

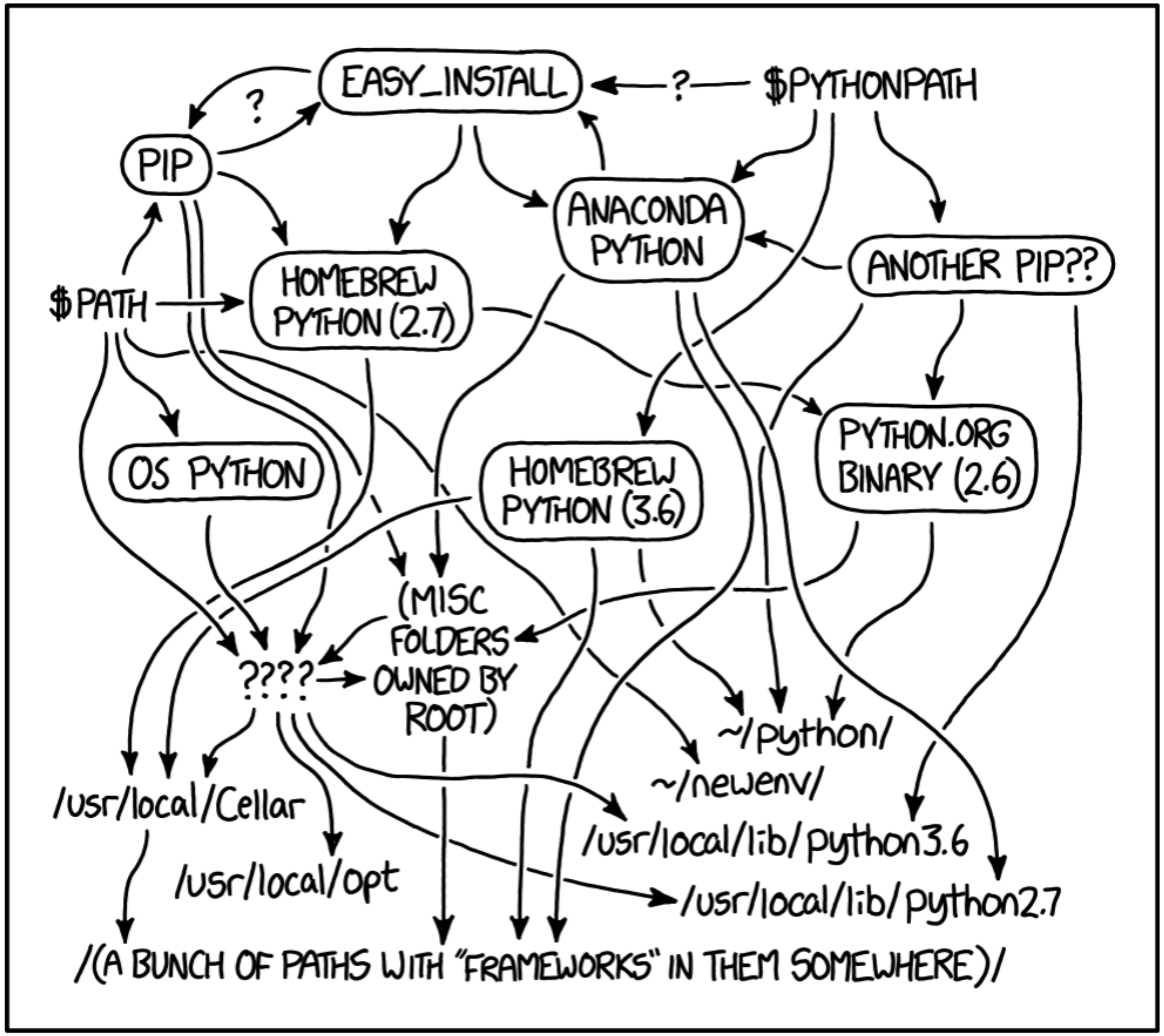
https://www.nytimes.com/interactive/2020/04/22/upshot/coronavirus-models.html

What is this course about?

- The subject is about **modelling and simulation** for science/engineering.
- The subject is about **solving computational problems** that often arise doing science/engineering.
- Mathematics will provide a necessary foundation.
- The subject is **not about programming**, although you will practically work out things you learn in this unit.
- You can use **Python/Matlab**. We won’t teach you the language. You will have to pick it up.



https://xkcd.com/353/



MY PYTHON ENVIRONMENT HAS BECOME SO DEGRADED THAT MY LAPTOP HAS BEEN DECLARED A SUPERFUND SITE.

https://xkcd.com/1987/

Content at glance

- General **strategies/techniques for modelling**:
  - Linear vs Non-linear models
  - Discrete vs Continuous models
  - Deterministic vs Stochastic Models

and **model analysis**

- Phase-plane analysis
- Steady state solutions
- Linearisation around fixed points, etc

Content at glance

- A selection of **numerical methods**
  - For solving linear systems
  - For solving non-linear equations
  - For solving first-order differential equations.

- A selection of **computational strategies**
  - Combinatorial search
  - Montecarlo Simulation
  - Heuristics

Although this is a 3rd-year unit, it is an introductory unit.
Focuses on **breadth** of techniques.

Introduction to Computational Models

Linear Systems

Non-Linear Systems and Dynamics

Stochastic Systems and Simulation

Heuristics for Optimisation

Game Theory

Unit schedule

Week 1	1	Introduction to Modelling and Simulation	
	2	Floating point number systems	
Week 2	3	Errors in Models	Quizzes begin - Assignment 1 released
	4	LU factorisation and partial pivoting	
Week 3	5	Eigensystems	
	6	Power iteration methods	
Week 4	7	Non-linear equations	
	8	Dynamical systems and difference equations	Assignment 1 due
Week 5	9	Coupled models	Assignment 2 released
	10	Continuous models and Lotka Volterra	
MID SEMESTER BREAK			
Week 6	11	Solving ODE's	
	12	Runge Kutta	
Week 7	13	Montecarlo	
	14	Sampling	
Week 8	15	Discrete event simulation	Assignment 2 due
	16	Discrete Markov chains	[ANZAC DAY - Workshop replacement online]
Week 9	17	Final project Overview	
	18	Heuristics + Simulated Annealing	Final Project released
Week 10	19	Evolutionary Computation	
	20	Heuristics case study	
Week 11	21	Games in Normal Form	
	22	Nash Equilibrium	
Week 12	23	Support Enumeration	
	24	Game Theory Applications	
Week 14			Final project due

Unit structure

- **Pre-workshop:**
  Watch videos on your own time before the workshop

- **Workshop:**
  Mondays 11:00 (S4)
  Thursdays 11 AM (S3).

- **Applied sessions:**
  2 hours during the week - weeks 1 to 12.
  Theory and practical — worksheet published week before

Teaching team

Julian García

- Room 331, 29 Ancora Imparo, CL
  Email: julian.garcia@monash.edu

Simon Teshuva

- Email: simon.teshuva@monash.edu

FIT3139.Clayton-x@monash.edu

Assessment

Assessment task	Value	Due date
Assignment 1	15%	Week 4
Assignment 2	25%	Week 8
Quizzes	10%	Week 12 released weekly from week 2, best attempt
Final project	50%	Week 14

Cheating, Collusion, Plagiarism

- **Cheating:** Seeking to obtain an unfair advantage in an examination or in other written or practical work required to be submitted or completed for assessment.
- **Collusion:** Unauthorised collaboration on assessable work with another person or persons.
- **Plagiarism:** To take and use another person's ideas and or manner of expressing them and to pass them off as one's own by failing to give appropriate acknowledgement. This includes material from any source, staff, students or the Internet – published and un-published works.

http://infotech.monash.edu.au/resources/student/assignments/policies.html

http://bit.ly/plagiarism_video

We run plagiarism detection on all assignments.

Special consideration

http://www.monash.edu.au/exams/special-consideration.html

Centrally Managed
No exception

Recommended Reading

- **Applied Numerical Methods with MATLAB for Engineers & Scientists** 3ed Chapra S.

- **Scientific Computing : An Introductory Survey**, 2ed Heath M.

We have **no textbook**, due to structure of the course. Recommended reading will also be posted on Moodle, per week/topic.

If possible take notes…

The Pen Is Mightier Than the Keyboard

Advantages of Longhand Over Laptop Note Taking

Pam A. Mueller[1]

Daniel M. Oppenheimer[2]

[1]Princeton University
[2]University of California, Los Angeles

Pam A. Mueller, Princeton University, Psychology Department, Princeton, NJ 08544 E-mail: pamuelle@princeton.edu

Author Contributions Both authors developed the study concept and design. Data collection was supervised by both authors. P. A. Mueller analyzed the data under the supervision of D. M. Oppenheimer. P. A. Mueller drafted the manuscript, and D. M. Oppenheimer revised the manuscript. Both authors approved the final version for submission.

(slides are usually not self-standing **on purpose**)

Do you have any form of condition (medical, disability other) that impacts on your ability to study?

Disability Support Services provides a range of services for registered students including:

- Notetakers and Auslan interpreters
- Readings in alternative formats
- Adaptive equipment and software
- Alternative arrangements for exams

For further information and details about how to register:

Email: disabilitysupportservices@monash.edu
Phone: 03 9905 5704
Web: monash.edu/disability

ALSO:

http://www.monash.edu/health/counselling

Looking forward!