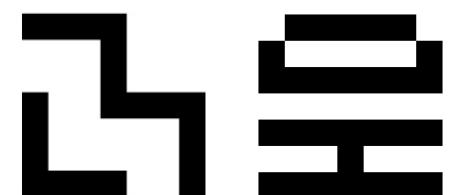


High-throughput behavioral science using virtual labs

Summer Institute in Computational Social Science

June 22nd , 2019



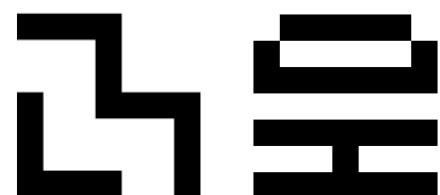
High-throughput behavioral science using
virtual labs

Abdullah M. Almaatouq
amaatouq@mit.edu

Different ways that you can make your experiment happen

	Cost	Control	Realism	Ethics
Partner with the powerful	low	medium	high	potentially complex
Use existing systems	low	low	high	potentially complex
Just do it yourself → Build an experiment	medium	high	medium	relatively easy
Build a product	high	high	high	relatively easy

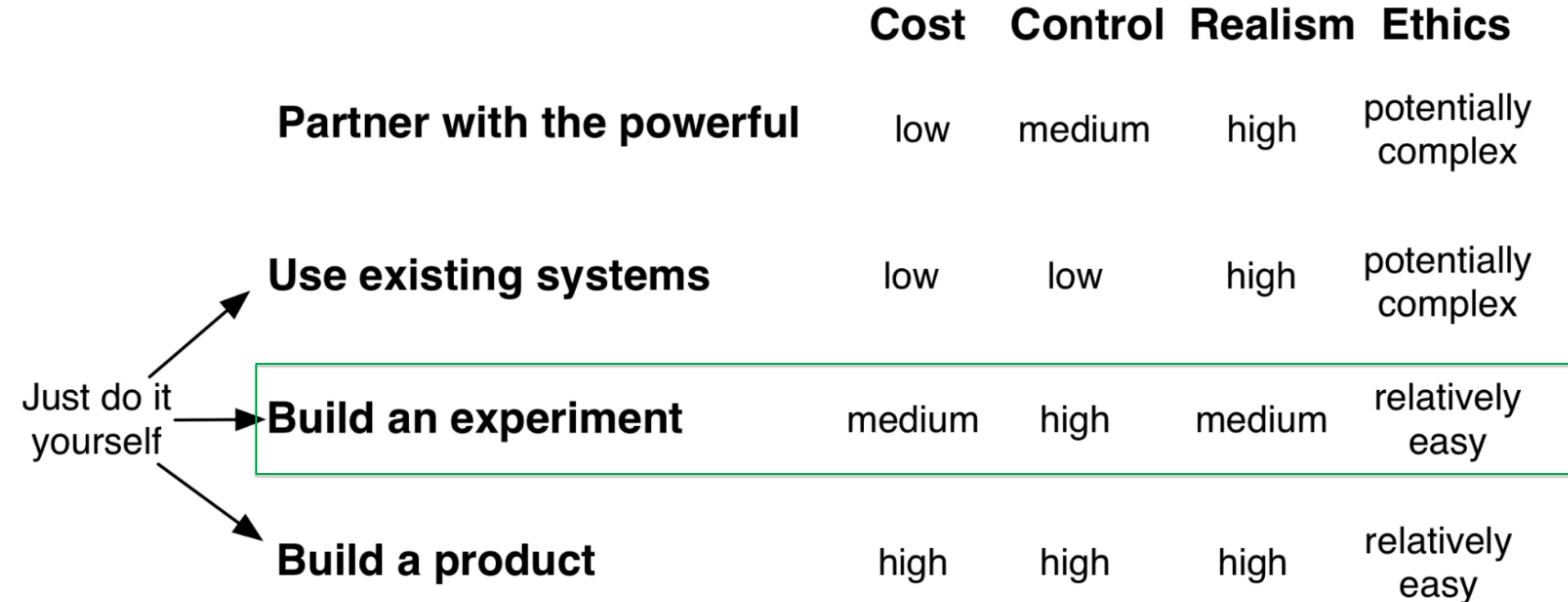
Salganik, Matthew J. *Bit by bit: social research in the digital age*. Princeton University Press, 2017.



Different ways that you can make your experiment happen

	Cost	Control	Realism	Ethics
Partner with the powerful	low	medium	high	potentially complex
Use existing systems	low	low	high	potentially complex
Build an experiment	medium	high	medium	relatively easy
Build a product	high	high	high	relatively easy

Just do it yourself → **Build an experiment**



Salganik, Matthew J. *Bit by bit: social research in the digital age*. Princeton University Press, 2017.



The 1960's Physical Labs



FIGURE 2. THE LARGE GROUP LABORATORY.



FIGURE 3. OBSERVATION DECK.



High-throughput behavioral science using
virtual labs

Abdullah M. Almaatouq
amaatouq@mit.edu



Fromkin, Howard L. "The behavioral science laboratories at Purdue's Krannert School." *Administrative Science Quarterly* (1969)

The 2000's Physical Labs



High-throughput behavioral science using
virtual labs

Abdullah M. Almaatouq
amaatouq@mit.edu

The 2000's Physical Labs



Limitations

- Subjects are homogeneous
- Limited number of people
- Short amount of time
- Simple task
- expensive, difficult to set up, and hard to replicate



Physical Labs: Limitations

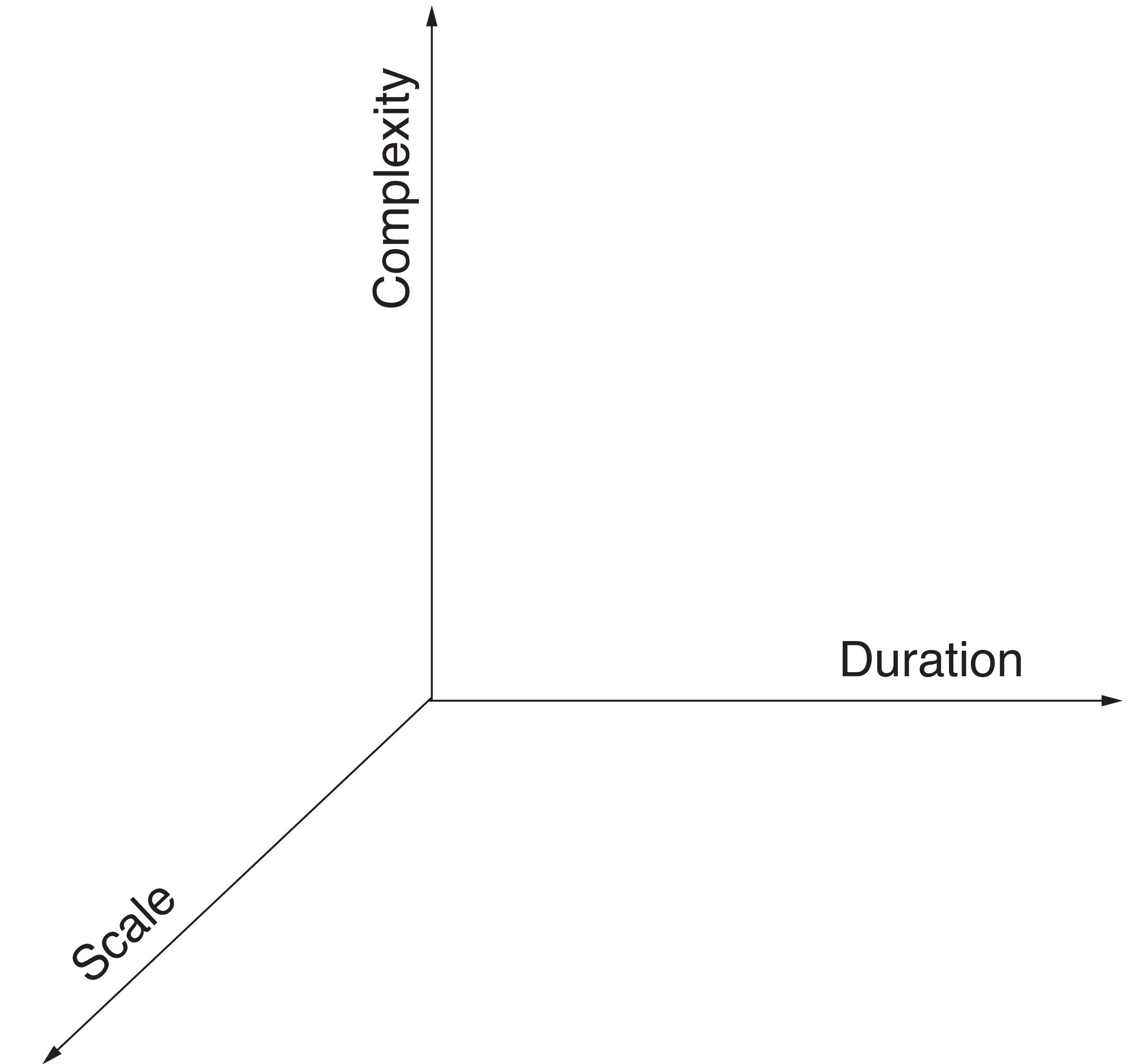


Figure Inspired by Duncan Watts

High-throughput behavioral science using
virtual labs

Abdullah M. Almaatouq
amaatouq@mit.edu

Physical Labs: Limitations

Physical labs

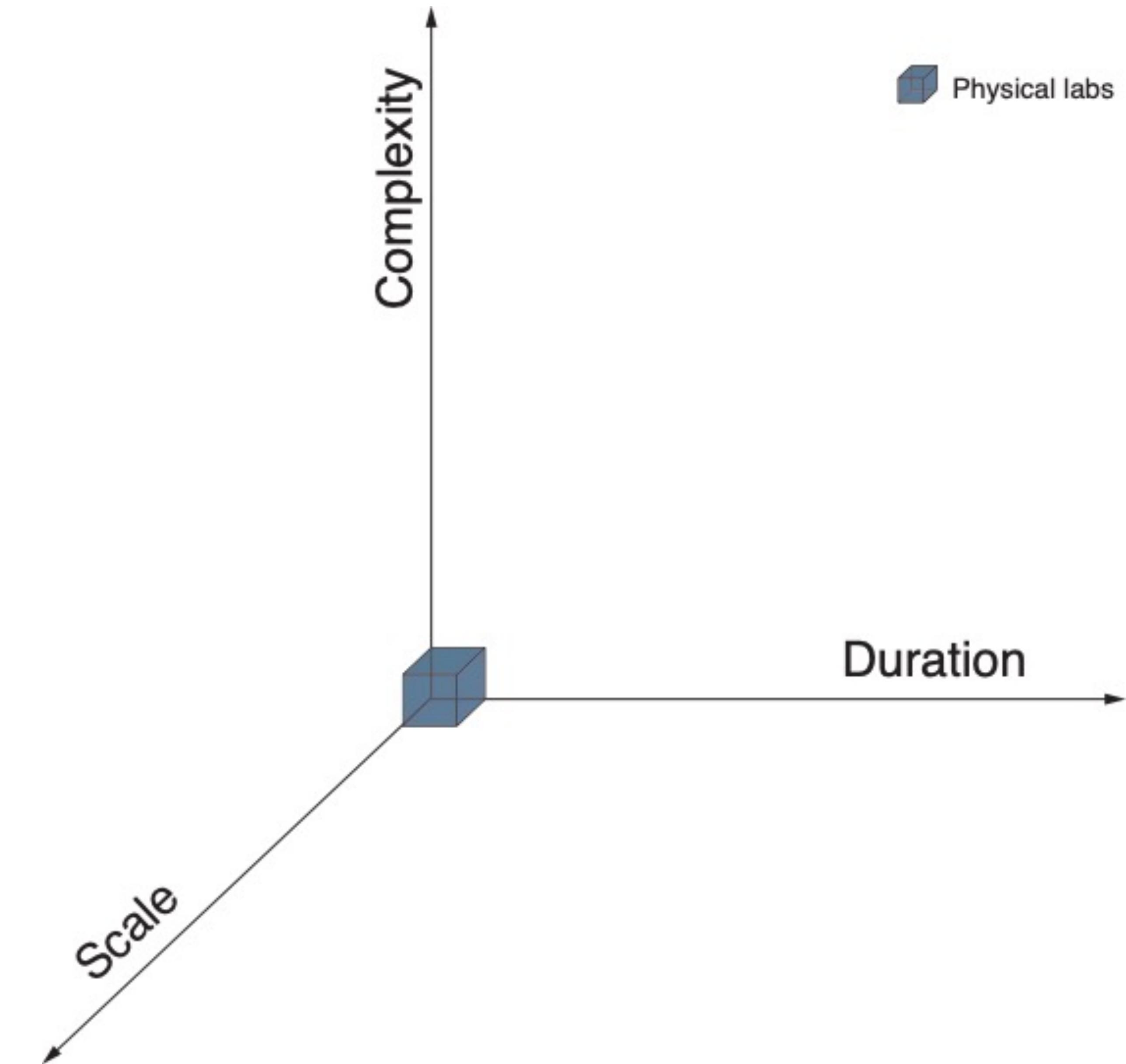


Figure Inspired by Duncan Watts

High-throughput behavioral science using
virtual labs

Abdullah M. Almaatouq
amaatouq@mit.edu

Virtual Labs: Lifting historical barriers

Physical labs

Scale (size and participation)

- Fewer constraints on location
- Large-scale social interactions
- More samples of data

Time (duration and synchronicity)

- Longer periods of time
- Real time interaction

Complexity (realism and granularity)

- More realistic tasks
- Immersive environments
- Precise instrumentation

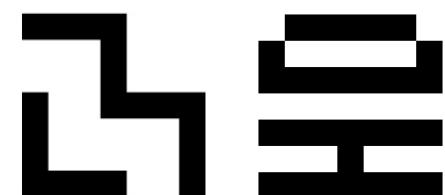
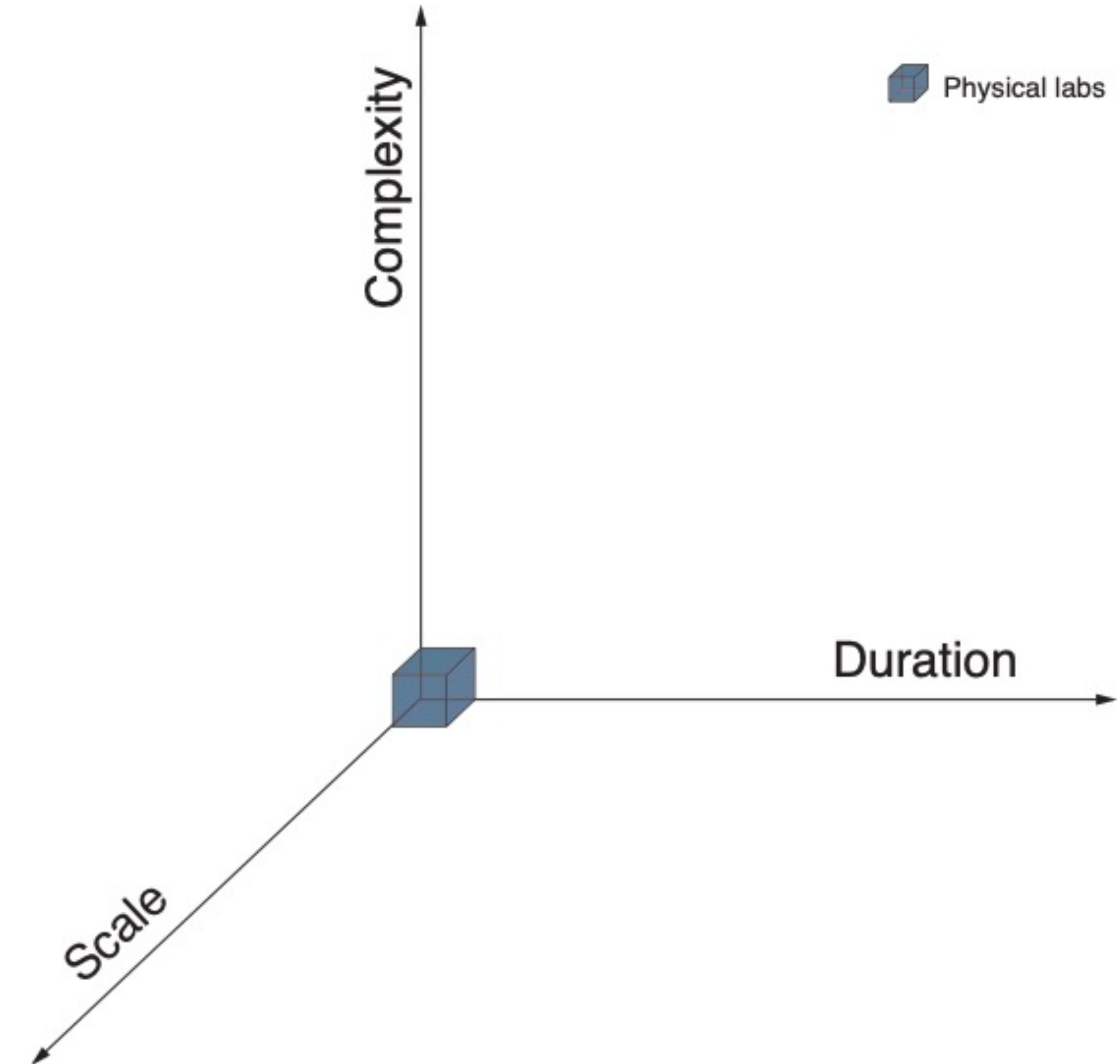


Figure Inspired by Duncan Watts

High-throughput behavioral science using
virtual labs

Abdullah M. Almaatouq
amaatouq@mit.edu

Virtual Labs: Lifting historical barriers

Physical labs
Virtual labs

Scale (size and participation)

- Fewer constraints on location
- Large-scale social interactions
- More samples of data

Time (duration and synchronicity)

- Longer periods of time
- Real time interaction

Complexity (realism and granularity)

- More realistic tasks
- Immersive environments
- Precise instrumentation

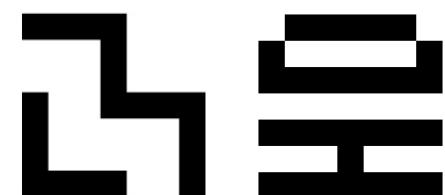
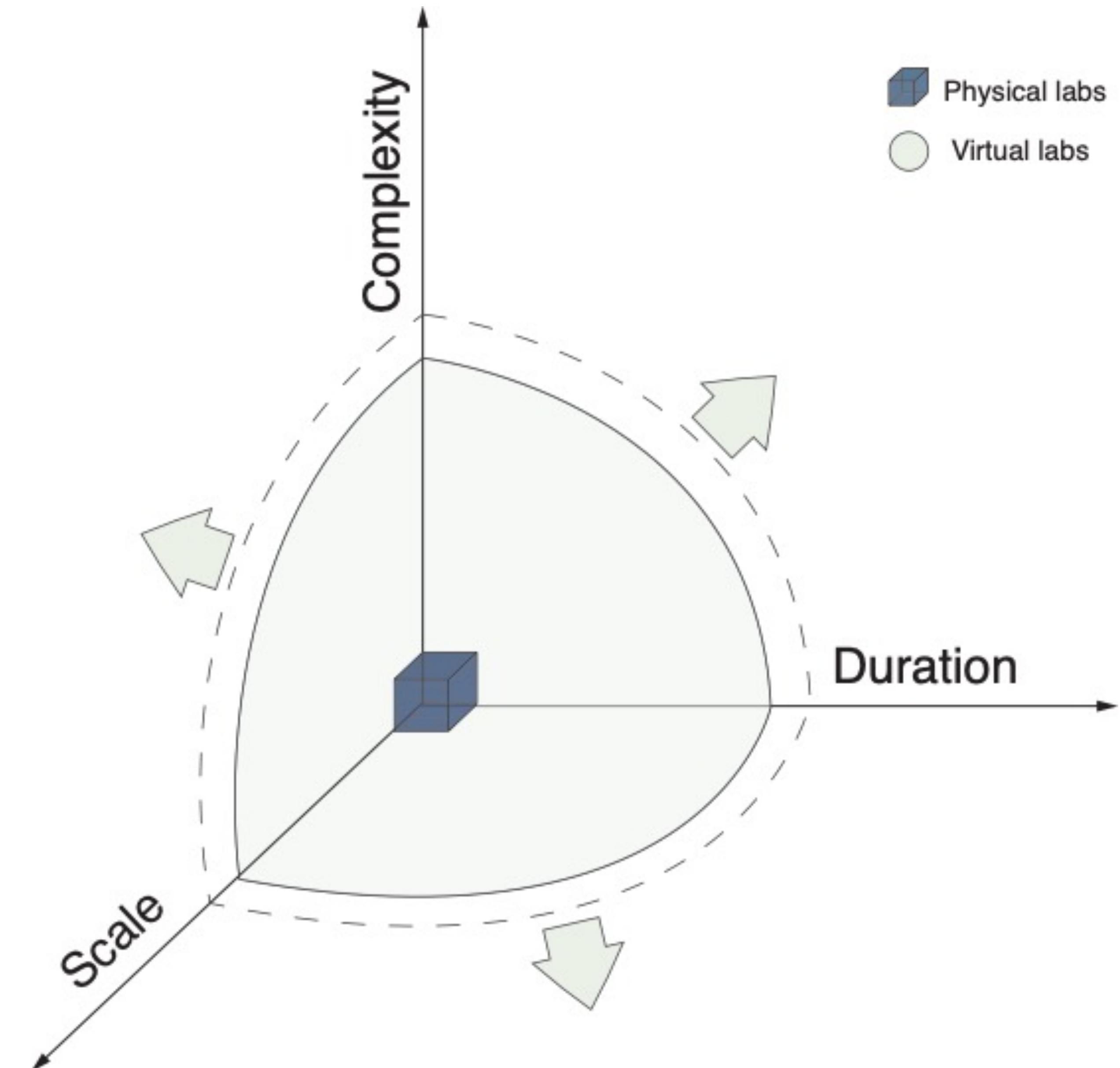


Figure Inspired by Duncan Watts

High-throughput behavioral science using
virtual labs

Abdullah M. Almaatouq
amaatouq@mit.edu

Virtual Labs: The reality

- Mostly traditional experiments in online environments
 - Very simple tasks
 - Mostly do not support “synchronous” designs
 - No real-time interaction
- Requires a lot of customized programming
 - Shady randomization procedures
- Results still hard to generalize to real-world scenarios
 - Inefficient in data collection
 - Less statistical power than physical labs
 - Goodman, et al. "Data collection in a flat world: The strengths and weaknesses of Mechanical samples." *Journal of Behavioral Decision Making* 26.3 (2013): 213-224.

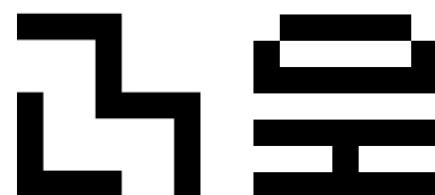
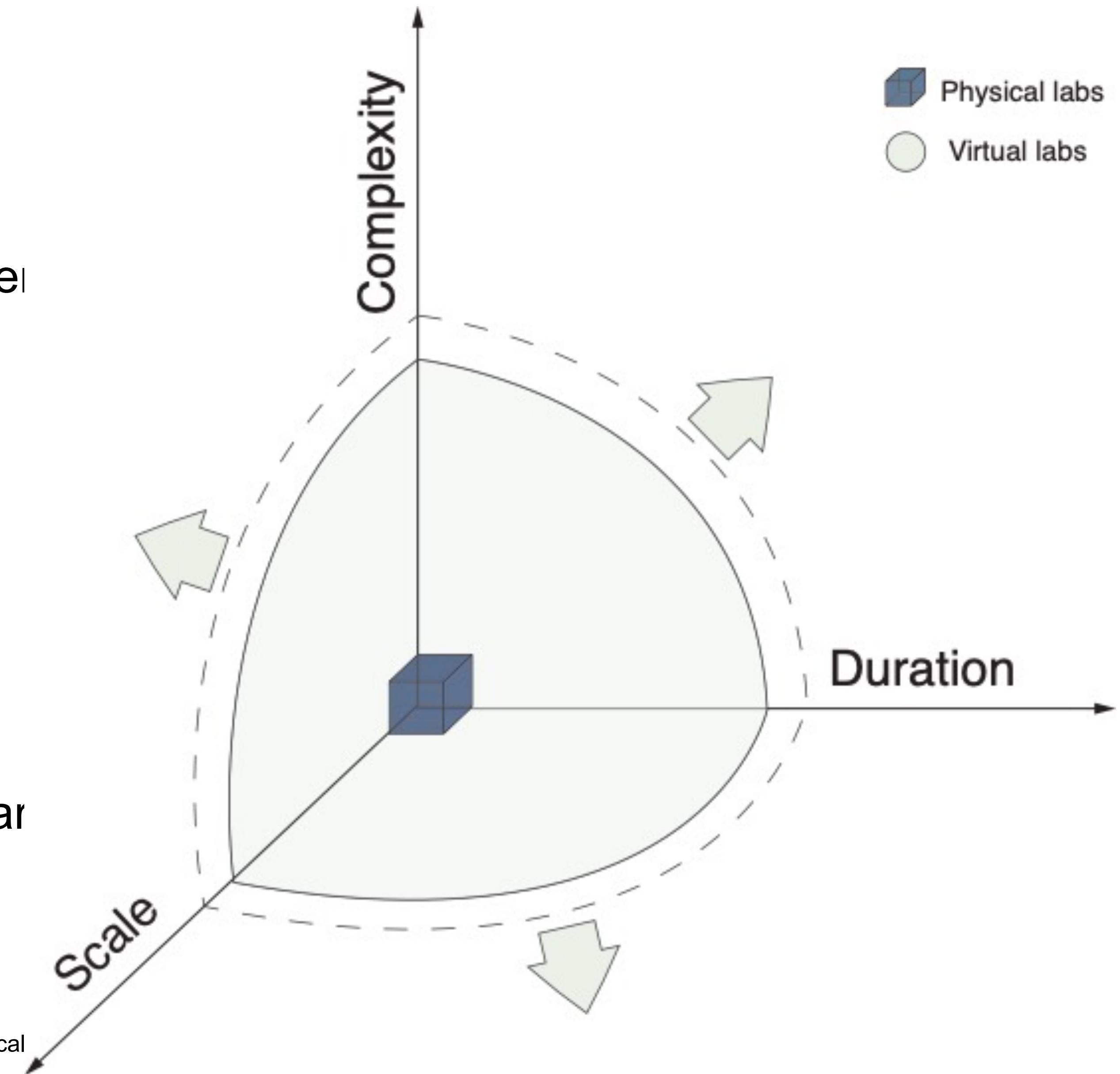


Figure Inspired by Duncan Watts

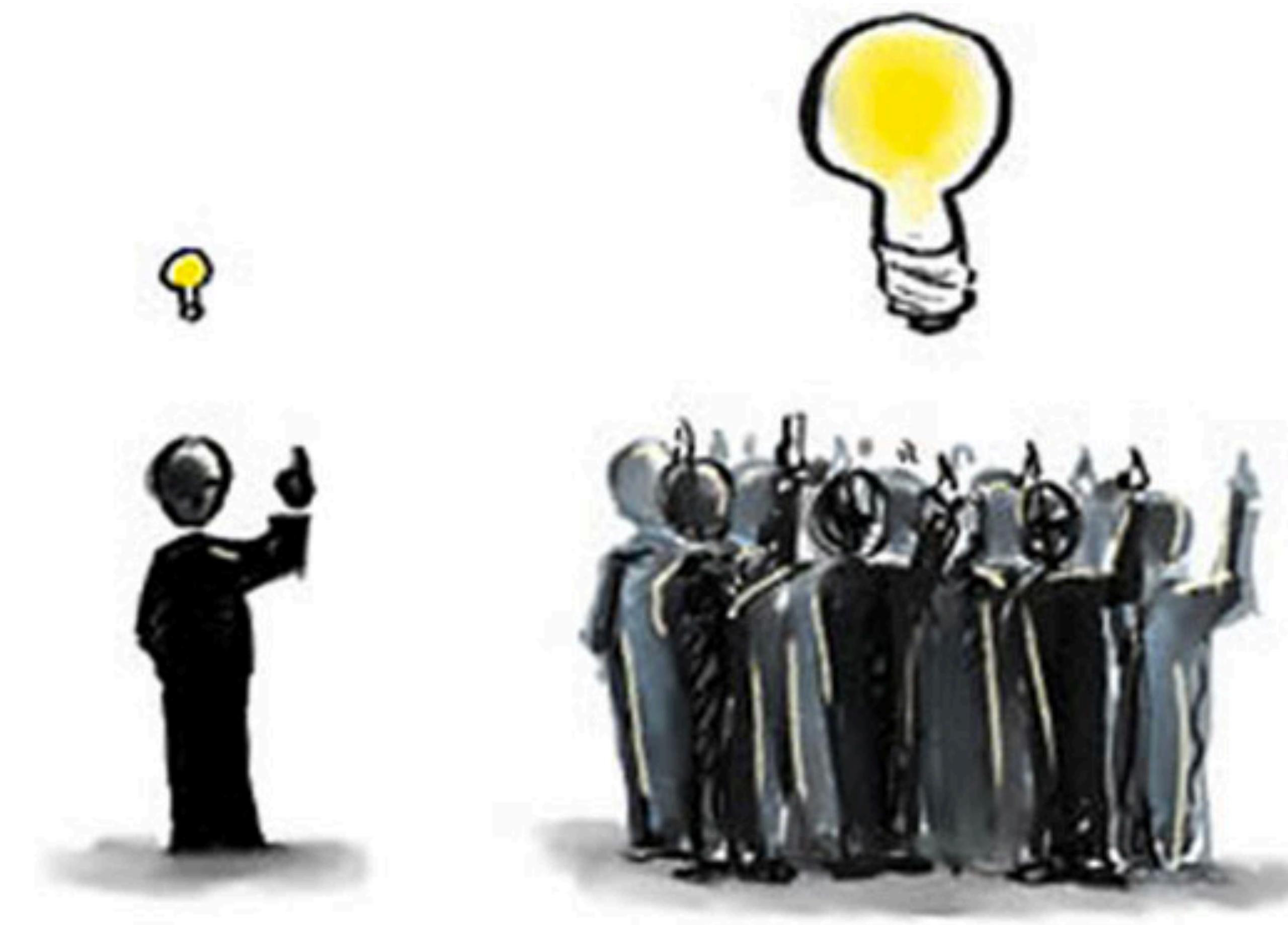
High-throughput behavioral science using
virtual labs

Abdullah M. Almaatouq
amaatouq@mit.edu

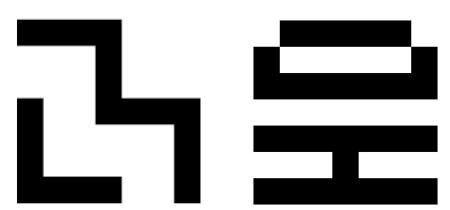
We needed more replicable, reliable, flexible solution



We needed more replicable, reliable, flexible solution



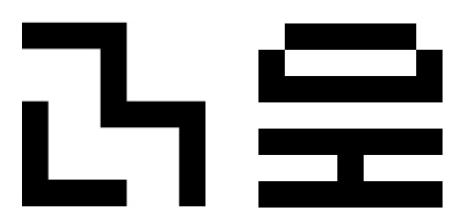
Collective Intelligence



High-throughput behavioral science using
virtual labs

Abdullah M. Almaatouq
amaatouq@mit.edu

Collective Intelligence



Collective Intelligence



Credit: Getty Images/Brand X



Collective Failure



Collective Failure

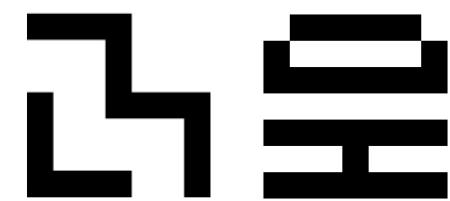
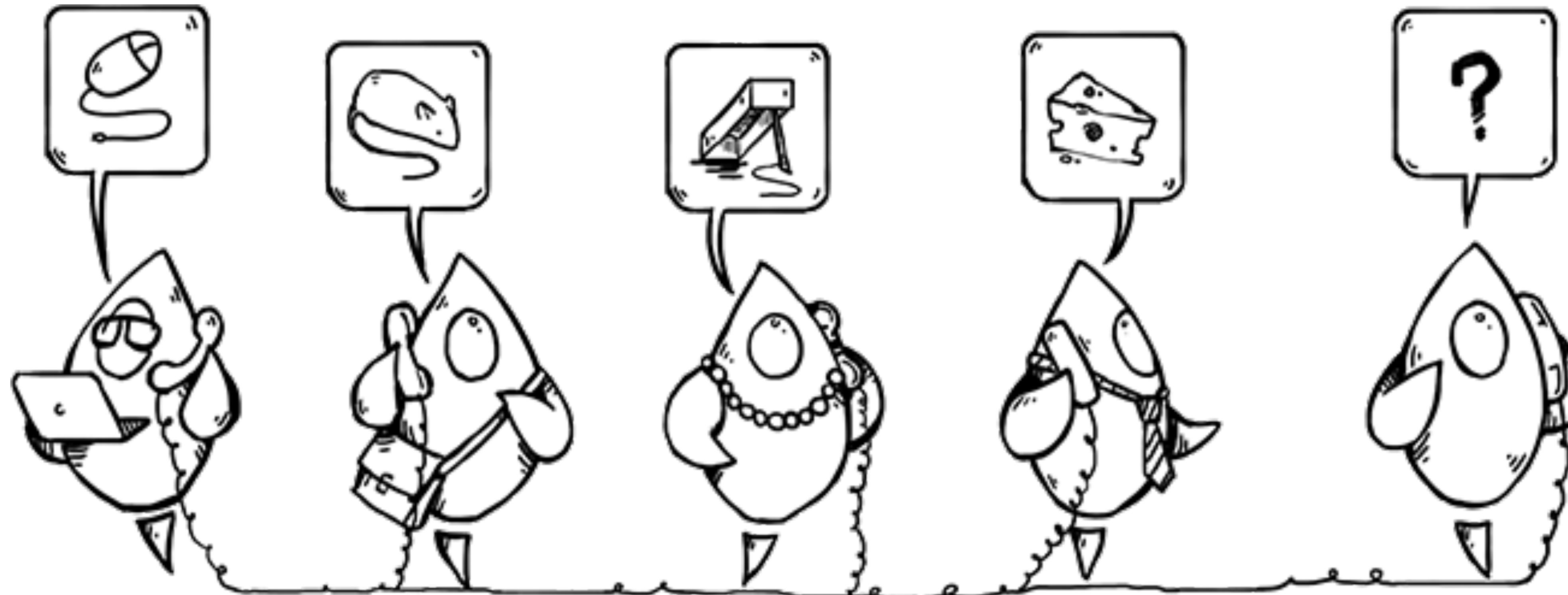


Collective Failure

How lies spread

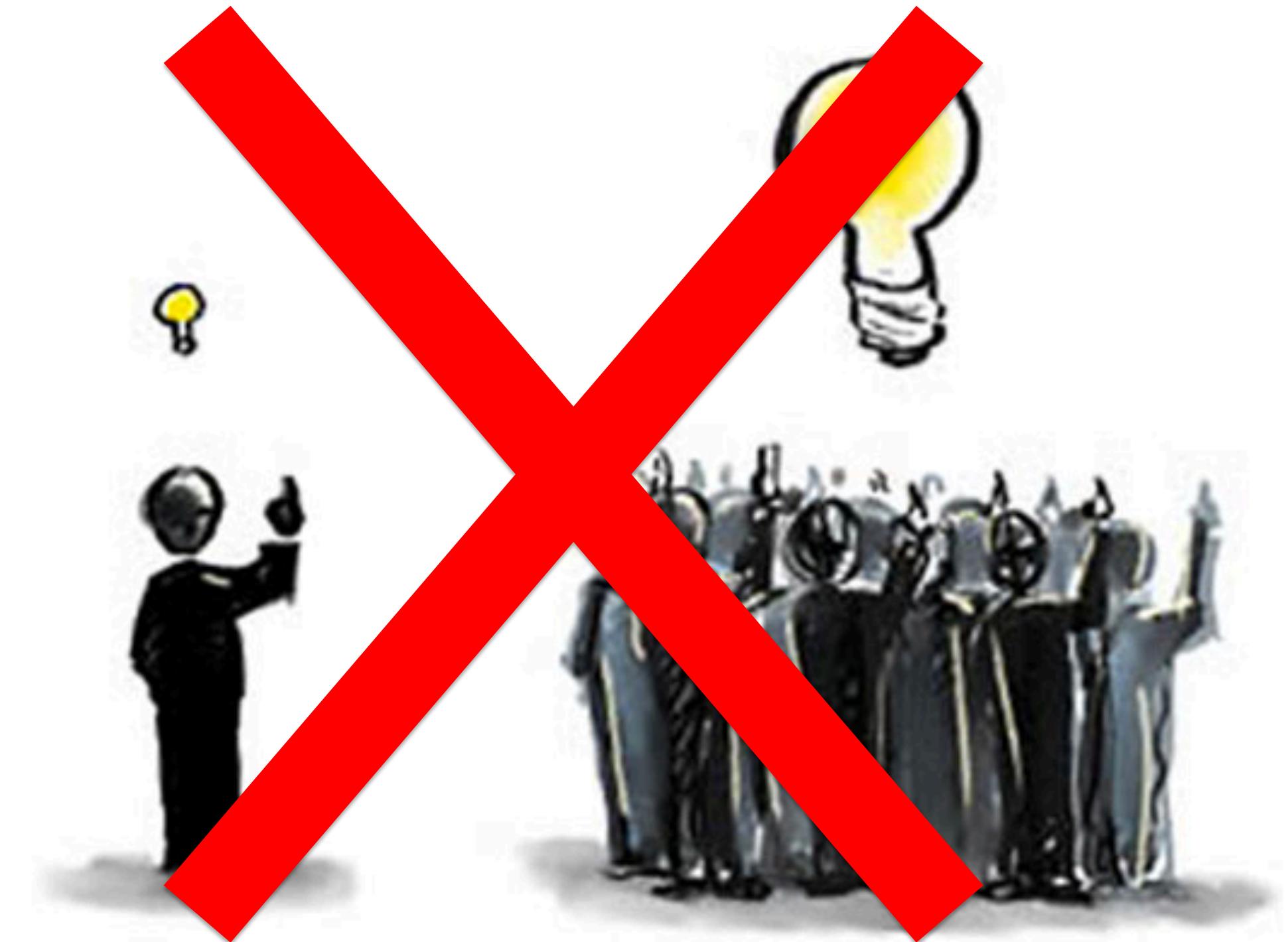
On social media, false news
beats the truth

(IMAGE) PETER BESHAI; (DATA) SOROUSH
VOSOUGHI, DEB ROY, AND SINAN ARAL



Collective Failure





**When does the wisdom of the crowds turn
into the madness of the mobs?**



Little consensus on what matters



Little consensus on what matters

- Individual level attributes
- Ability level
- Social perceptiveness
- Cognitive style diversity
- Diversity
- Skill diversity
- Cognitive style diversity
- Identity diversity
- Functional diversity
- Personality traits

- Team size
- Incentives / nature of the task
- Communication networks
- Distribution of information
- Aggregation mechanism
- ...

Evidence for a Collective Intelligence Factor in the Performance of Human Groups

Anita Williams Woolley^{1,*}, Christopher F. Chabris^{2,3}, Alex Pentland^{3,4}, Nada Hashmi^{3,5}, Thomas W. Malone^{3,5}

* See all authors and affiliations

Science 29 Oct 2010:

Groups of diverse problem solvers can outperform groups of high-ability problem solvers

Lu Hong and Scott E. Page

PNAS November 16, 2004. 101 (46) 16385-16389; <https://doi.org/10.1073/pnas.0403723101>

Optimally Interacting Minds

Bahador Bahrami^{1,2,3,*}, Karsten Olsen³, Peter E. Latham⁴, Andreas Roepstorff³, Geraint Rees^{1,2}, Chris D. Frith^{2,3}

* See all authors and affiliations

Science 27 Aug 2010:



Little consensus on what matters

- Individual level attributes
 - Team size
 - Incentives / nature of the task
 - **Communication networks**
 - Distribution of information
 - Aggregation mechanism
 - ...
- Ability level
- Social perceptiveness
- Cognitive style diversity
- Diversity
- Skill diversity
- Cognitive style diversity
- Identity diversity
- Functional diversity
- Personality traits

Evidence for a Collective Intelligence Factor in the Performance of Human Groups

Anita Williams Woolley^{1,*}, Christopher F. Chabris^{2,3}, Alex Pentland^{3,4}, Nada Hashmi^{3,5}, Thomas W. Malone^{3,5}

* See all authors and affiliations

Science 29 Oct 2010:

Groups of diverse problem solvers can outperform groups of high-ability problem solvers

Lu Hong and Scott E. Page

PNAS November 16, 2004. 101 (46) 16385-16389; <https://doi.org/10.1073/pnas.0403723101>

Optimally Interacting Minds

Bahador Bahrami^{1,2,3,*}, Karsten Olsen³, Peter E. Latham⁴, Andreas Roepstorff³, Geraint Rees^{1,2}, Chris D. Frith^{2,3}

* See all authors and affiliations

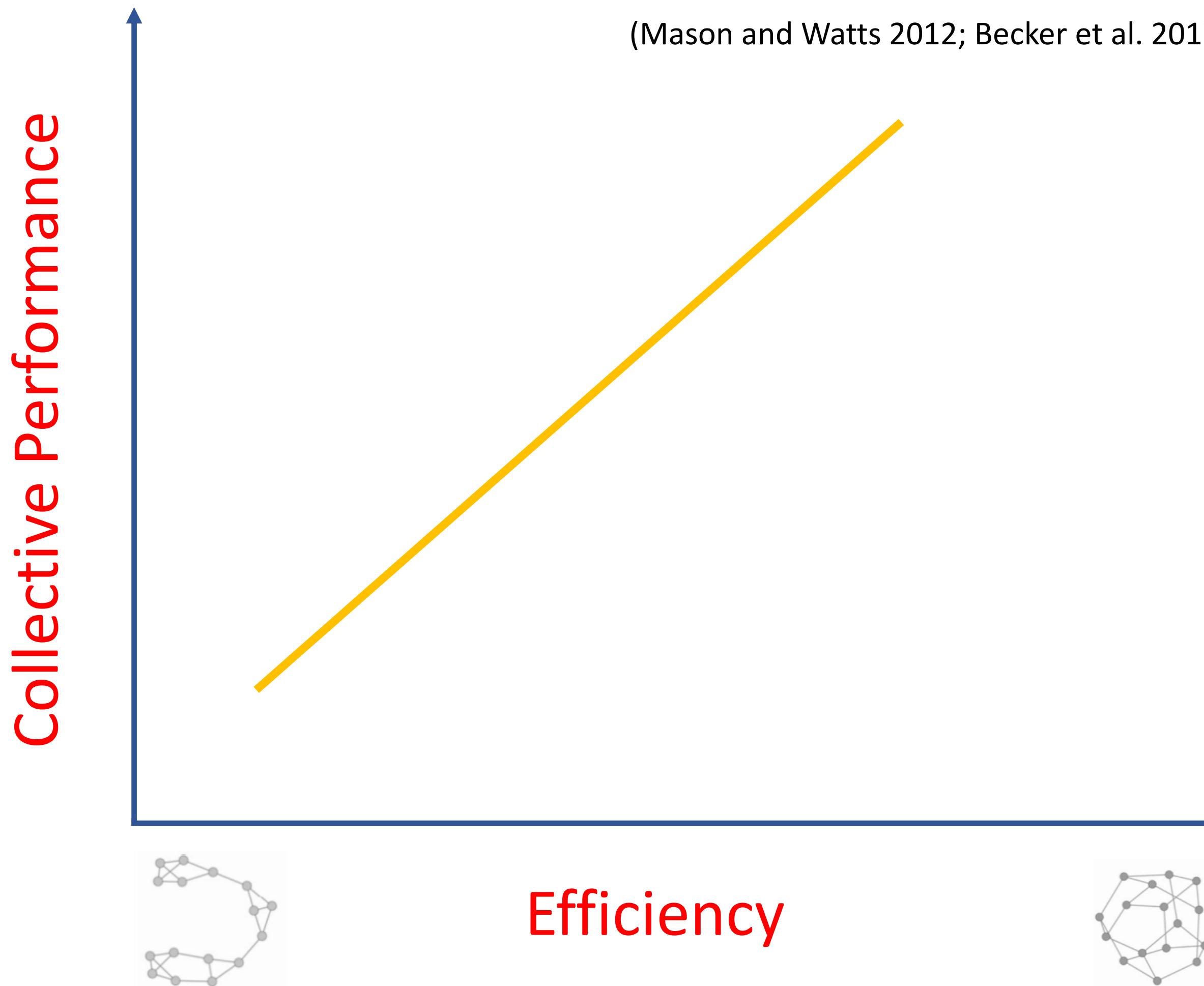
Science 27 Aug 2010:



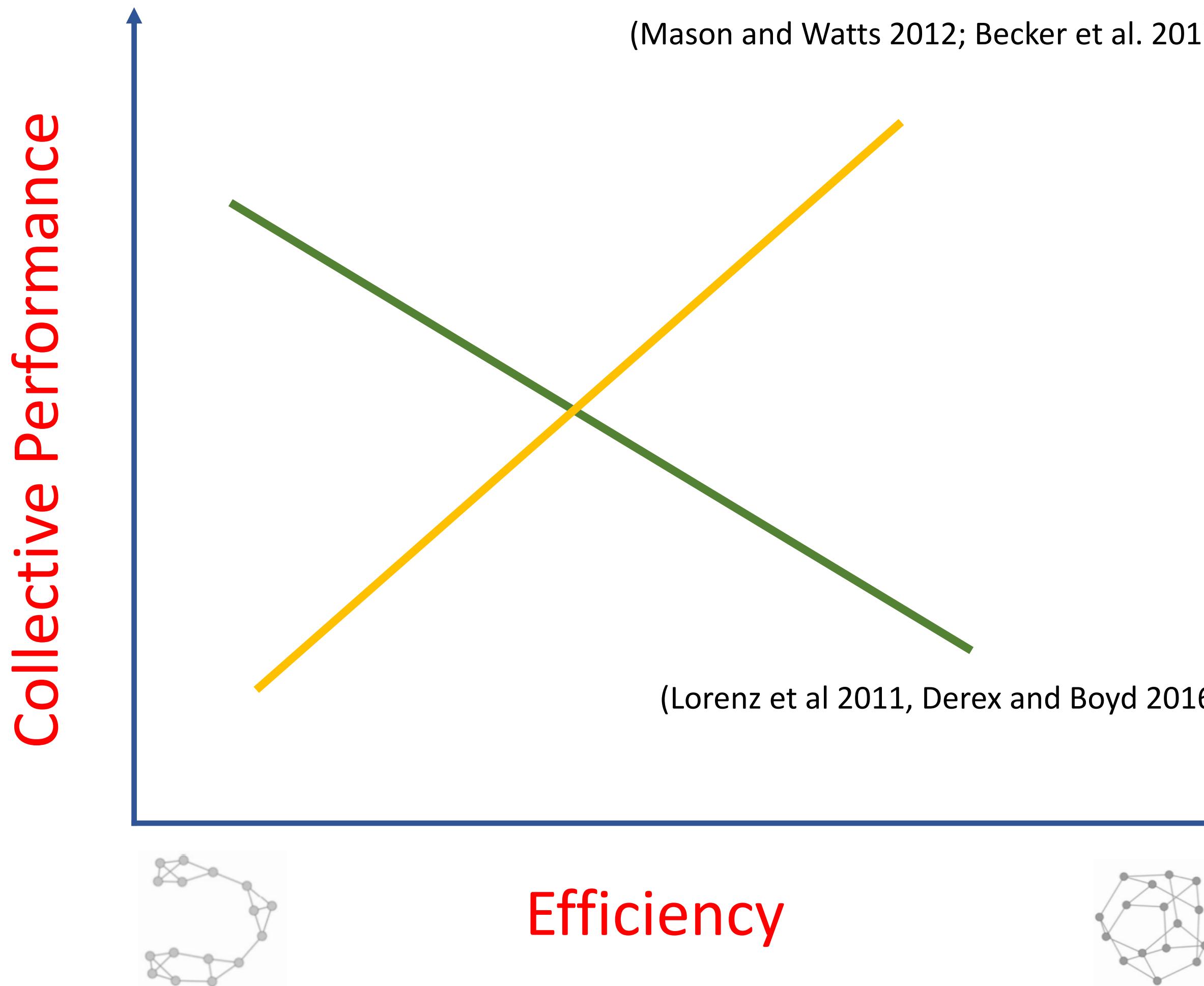
Communication Network Structure?



Communication Network Structure?



Communication Network Structure?



Disclaimer: These papers hard to compare due to different operationalizations

High-throughput behavioral science using
virtual labs

Abdullah M. Almaatouq
amaatouq@mit.edu

Little consensus on what matters

- Individual level attributes
- Ability level
- Social perceptiveness
- Cognitive style diversity
- Diversity
- Skill diversity
- Cognitive style diversity
- Identity diversity
- Functional diversity
- Personality traits

- Team size
- Incentives / nature of the task
- Communication networks
- Distribution of information
- Aggregation mechanism
- ...

Evidence for a Collective Intelligence Factor in the Performance of Human Groups

Anita Williams Woolley^{1,*}, Christopher F. Chabris^{2,3}, Alex Pentland^{3,4}, Nada Hashmi^{3,5}, Thomas W. Malone^{3,5}

* See all authors and affiliations

Science 29 Oct 2010:

Groups of diverse problem solvers can outperform groups of high-ability problem solvers

Lu Hong and Scott E. Page

PNAS November 16, 2004. 101 (46) 16385-16389; <https://doi.org/10.1073/pnas.0403723101>

Optimally Interacting Minds

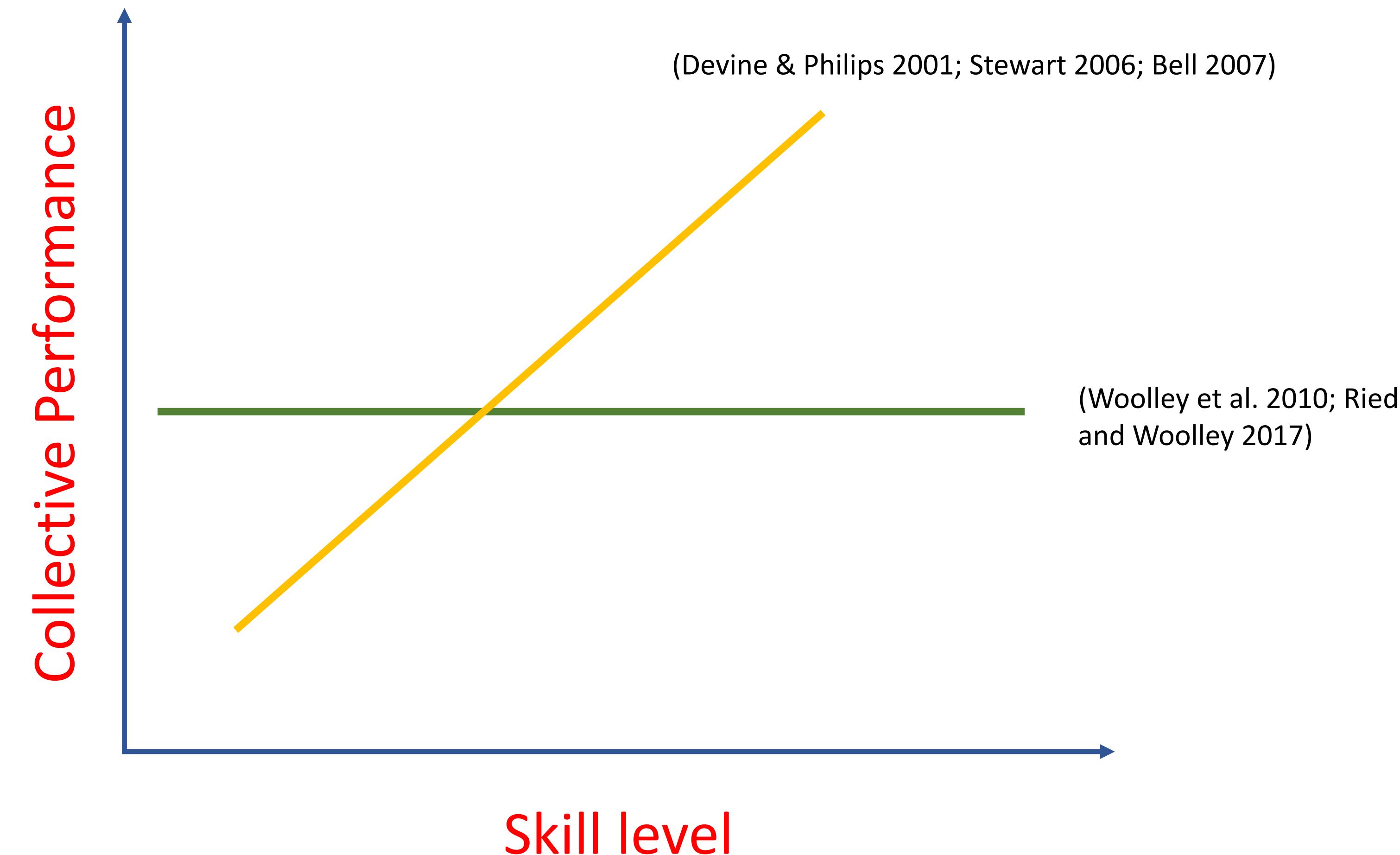
Bahador Bahrami^{1,2,3,*}, Karsten Olsen³, Peter E. Latham⁴, Andreas Roepstorff³, Geraint Rees^{1,2}, Chris D. Frith^{2,3}

* See all authors and affiliations

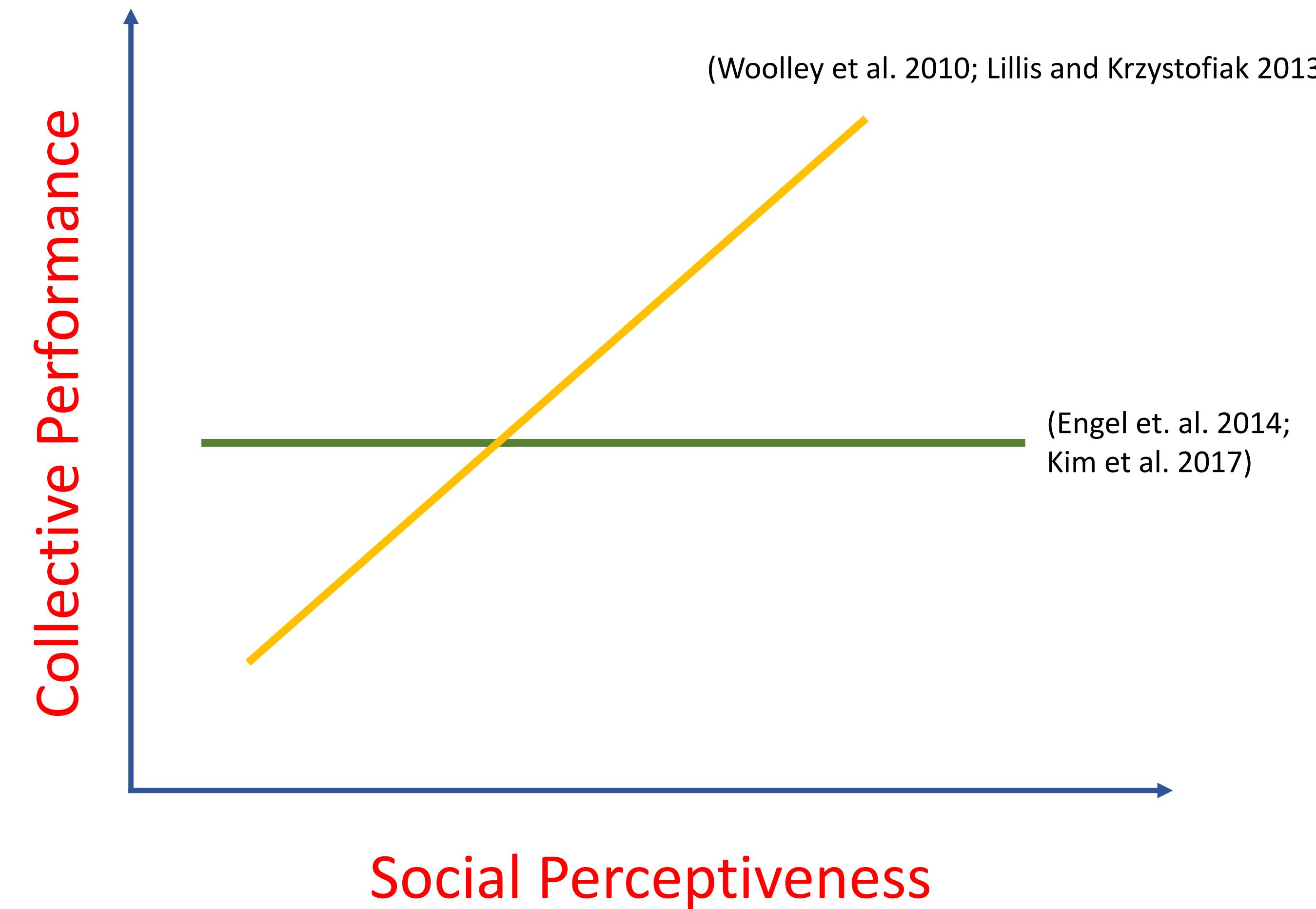
Science 27 Aug 2010:



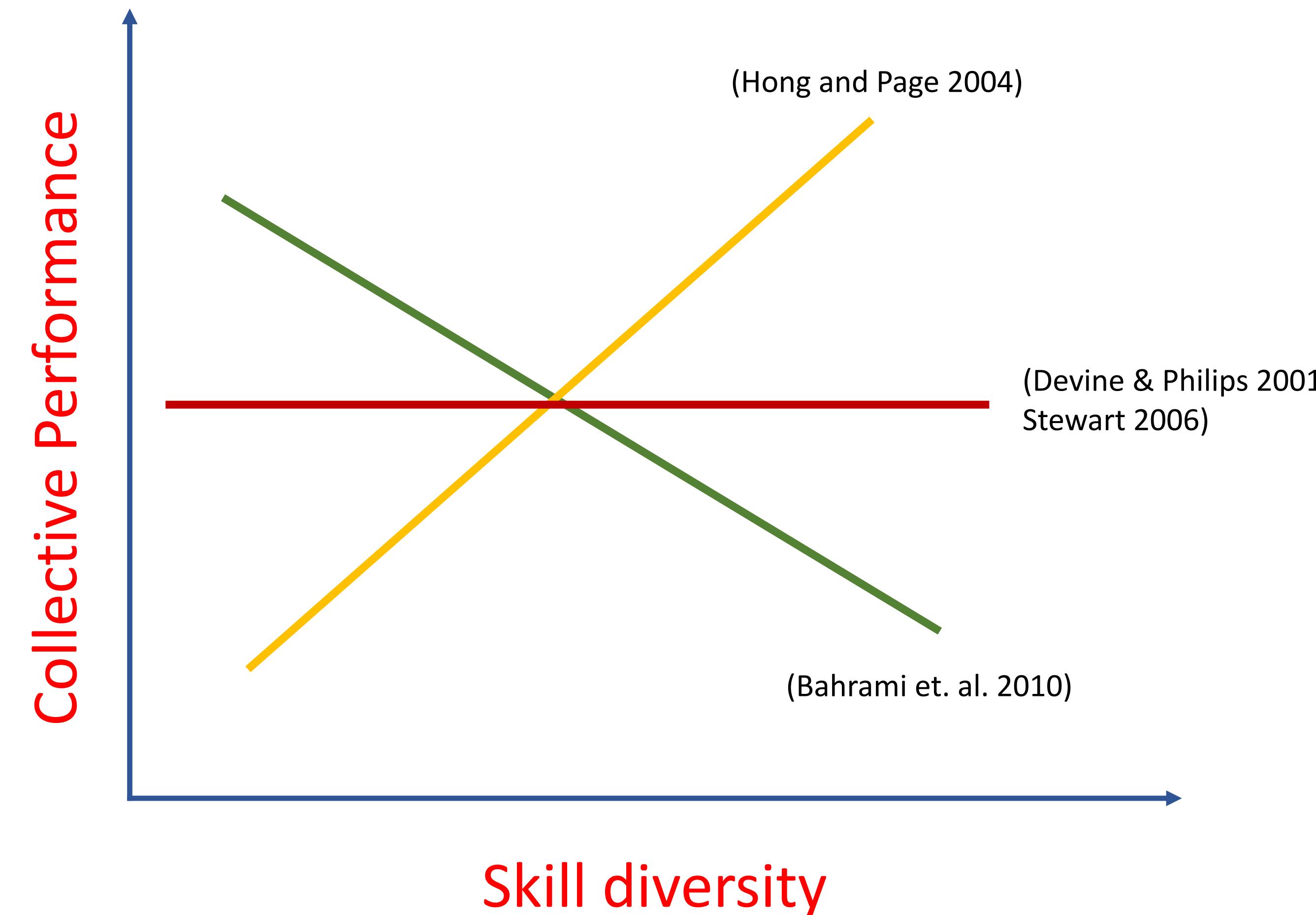
Individual Ability?



Social Perceptiveness (aka emotional intelligence)?

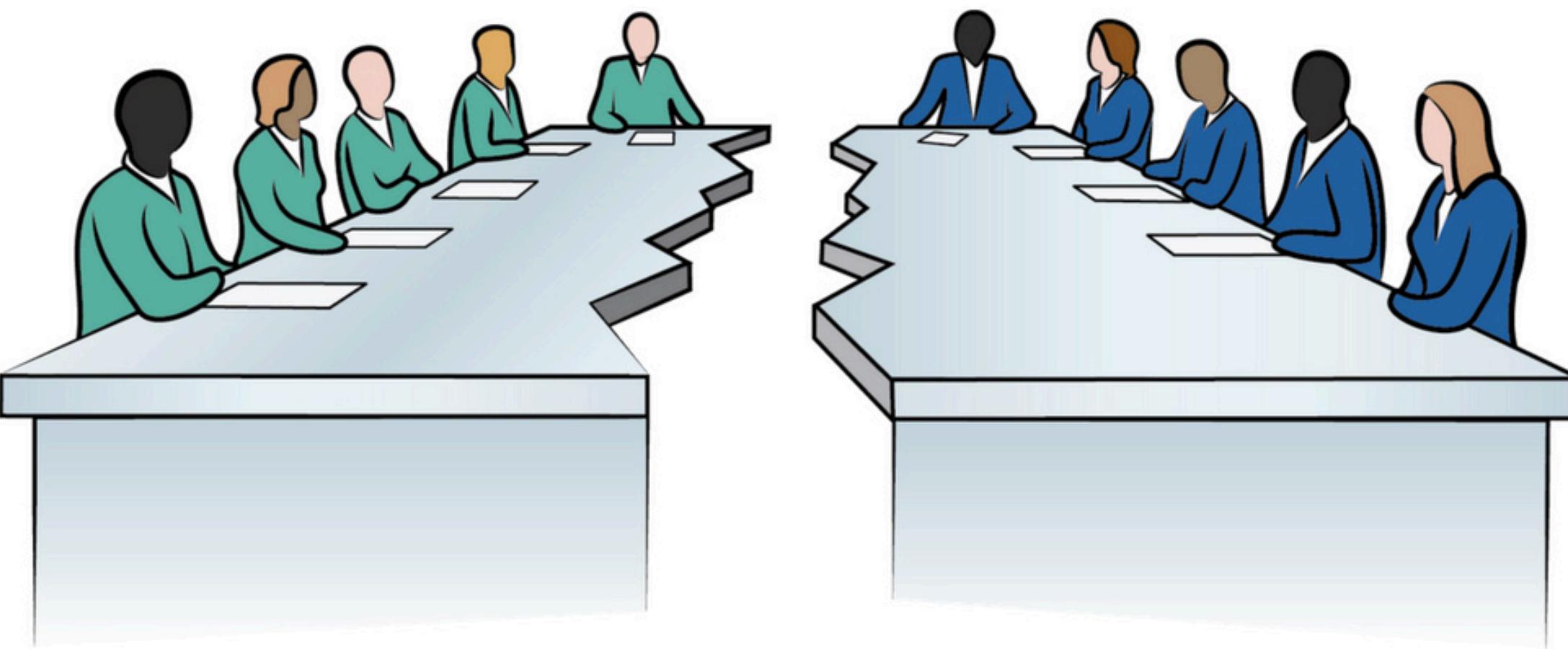


Skill Diversity?



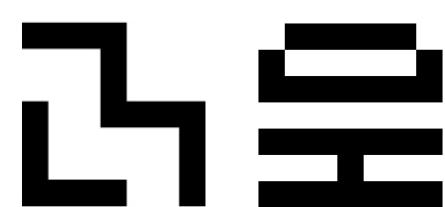
“Incoherence problem”





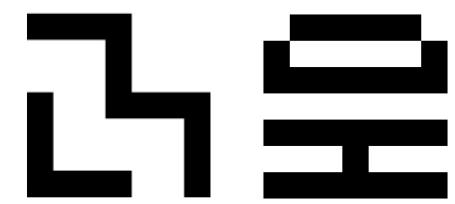
The Choice of Parameters...

The Environment



T1

T2



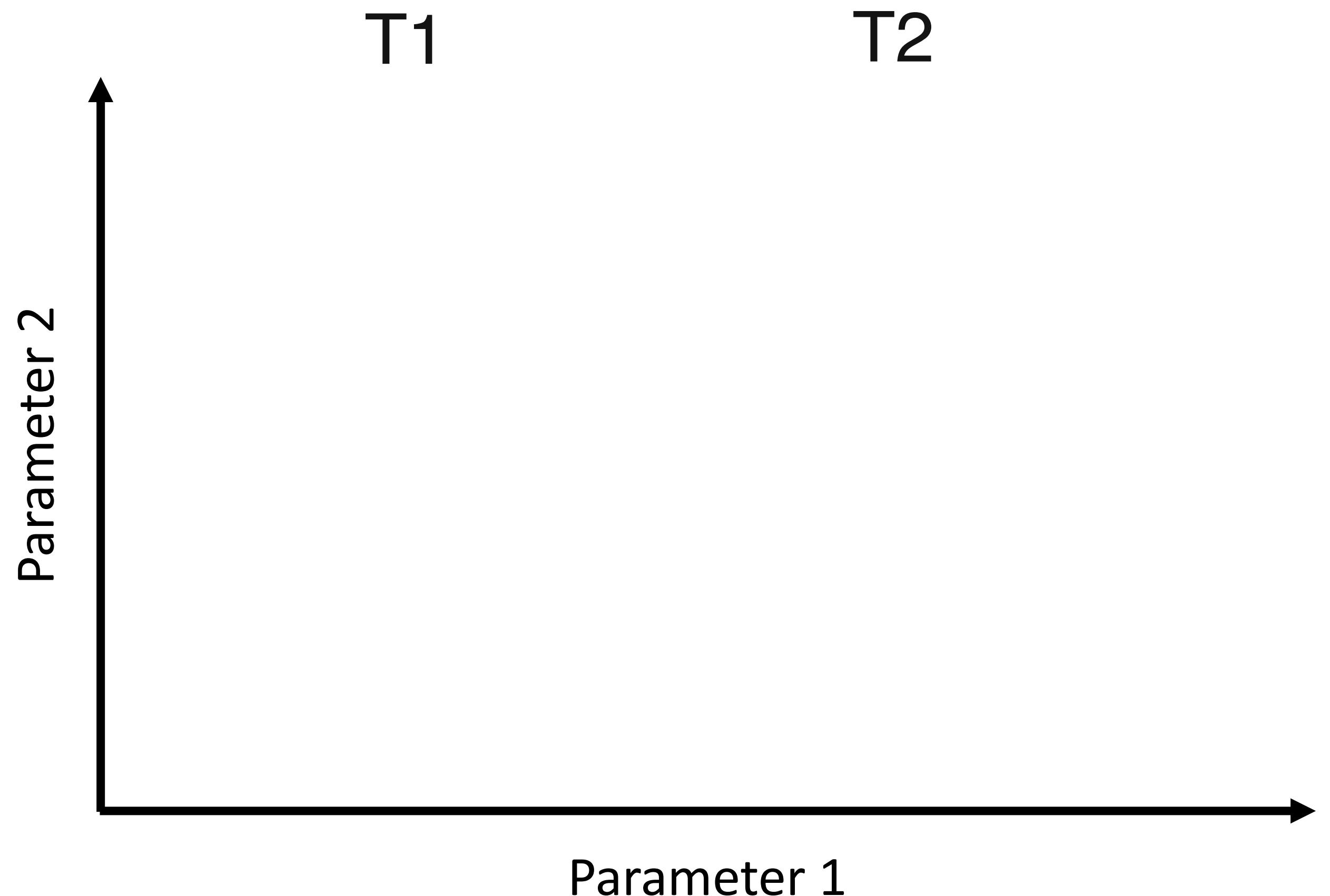
T1

T2

Parameters:

- Difficulty of the task
- Number of choices
- Proportion of experts
- Saliency of social signal
- Payoff
- Group size

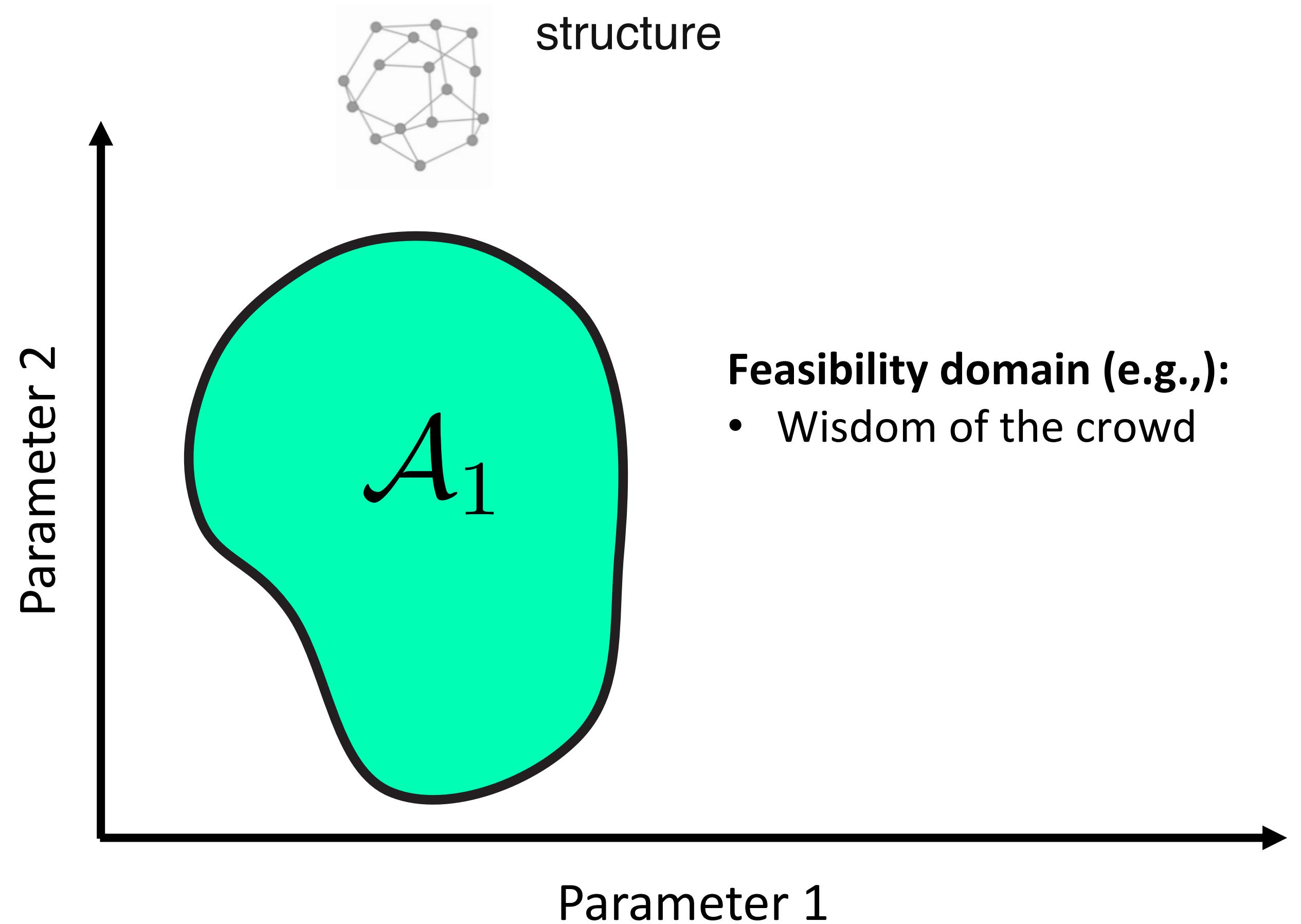




Parameters:

- Difficulty of the task
- Number of choices
- Proportion of experts
- Saliency of social signal
- Payoff
- Group size

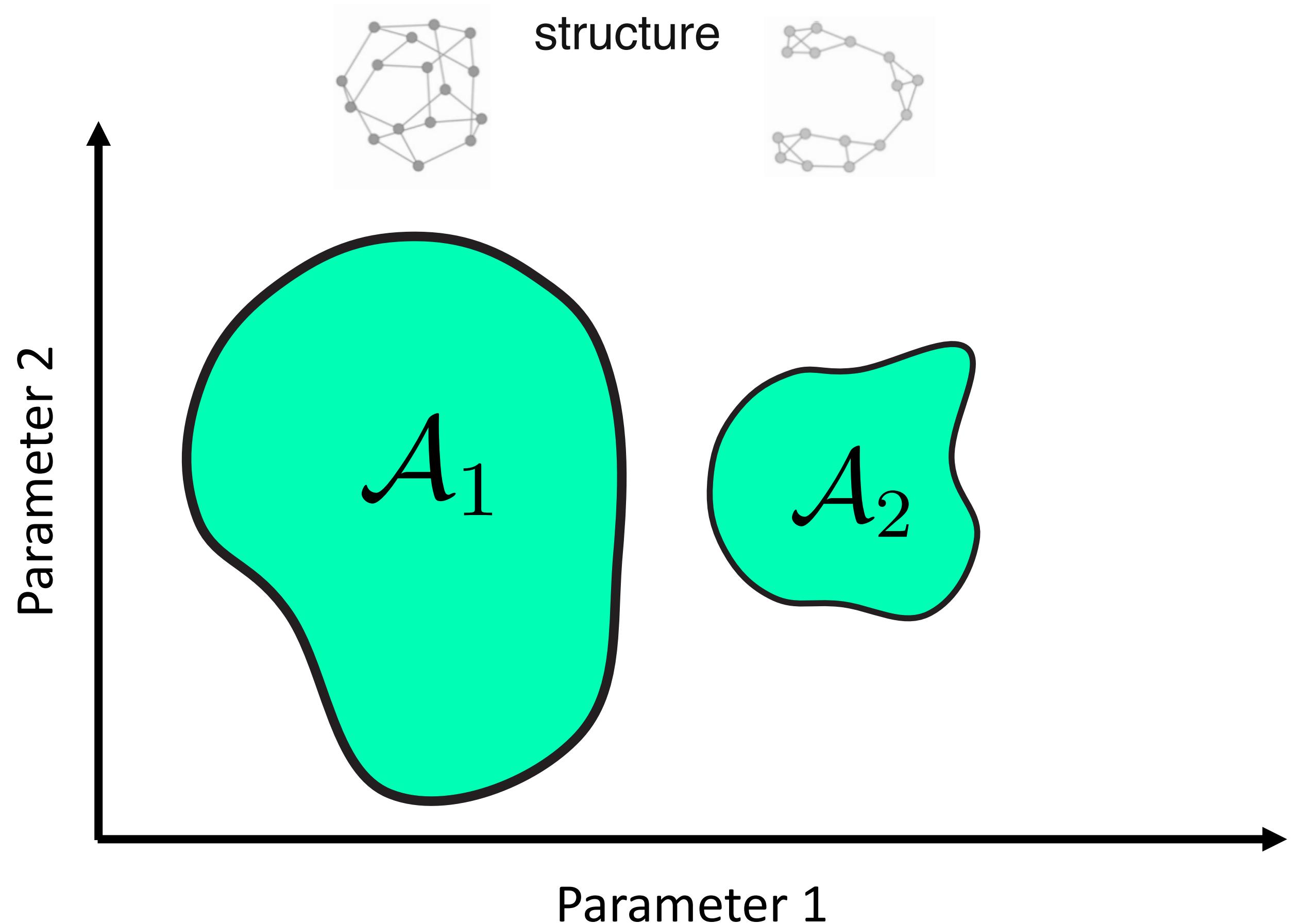




Parameters:

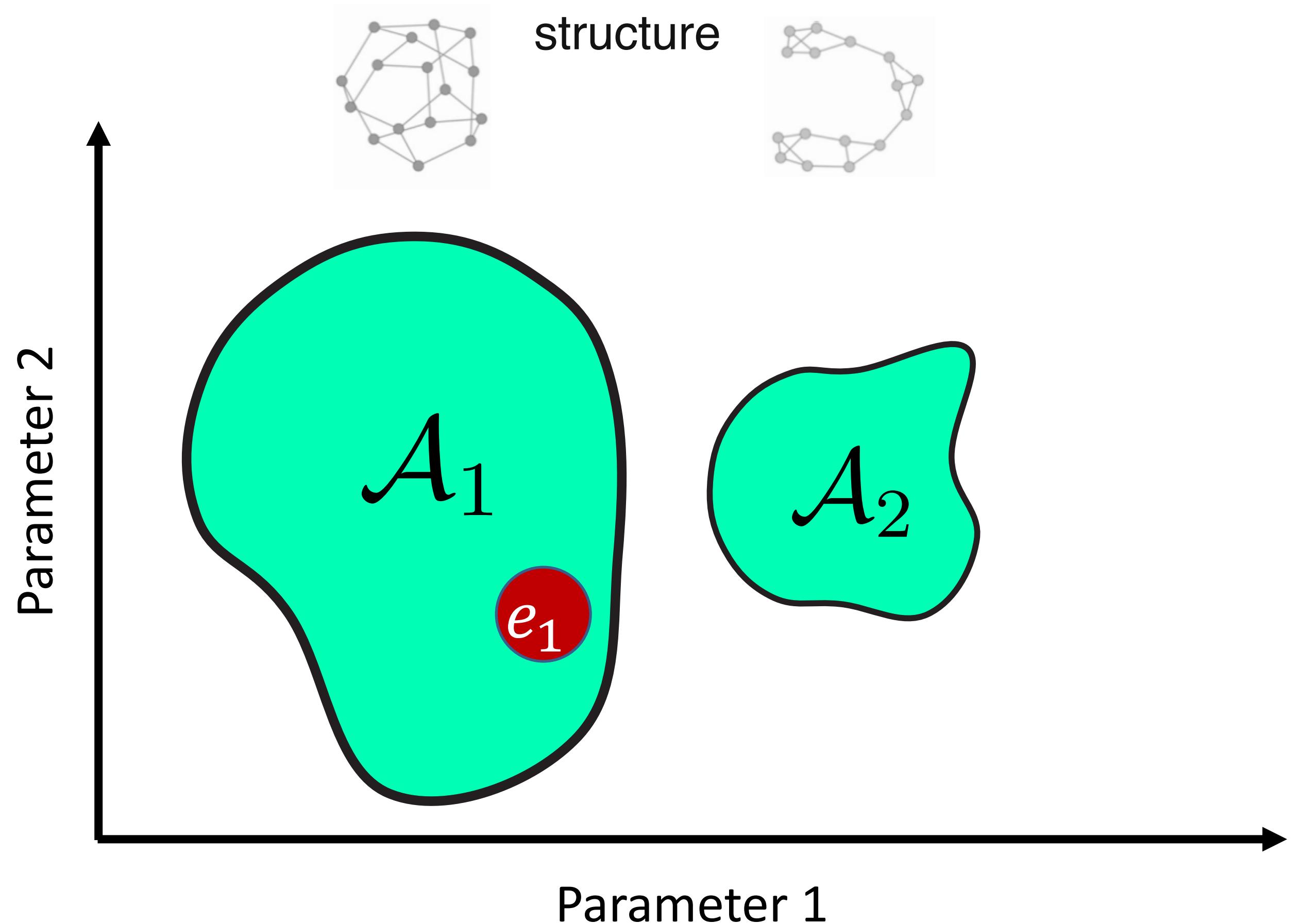
- Difficulty of the task
- Number of choices
- Proportion of experts
- Saliency of social signal
- Payoff
- Group size





Parameters:

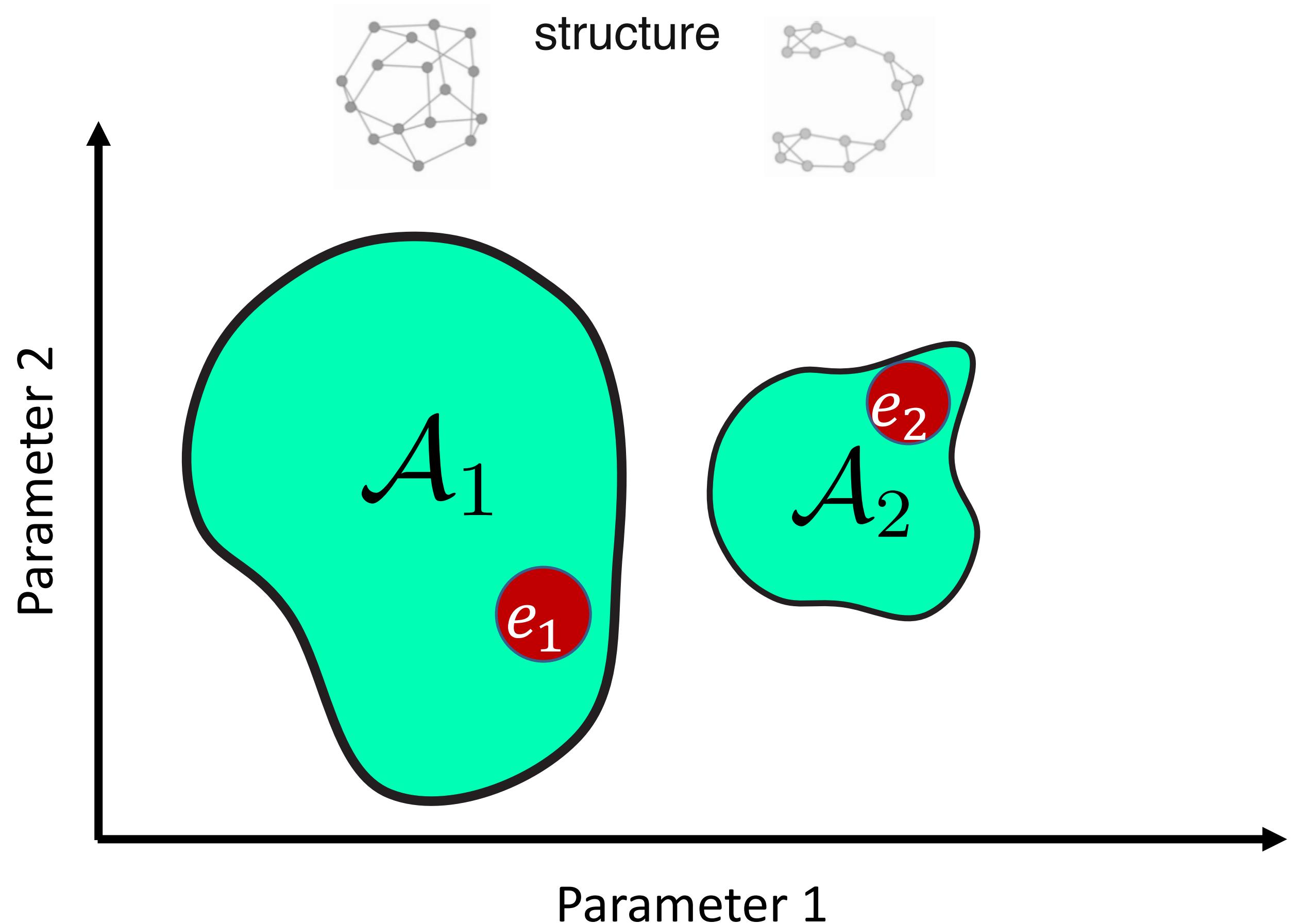
- Difficulty of the task
- Number of choices
- Proportion of experts
- Saliency of social signal
- Payoff
- Group size



Parameters:

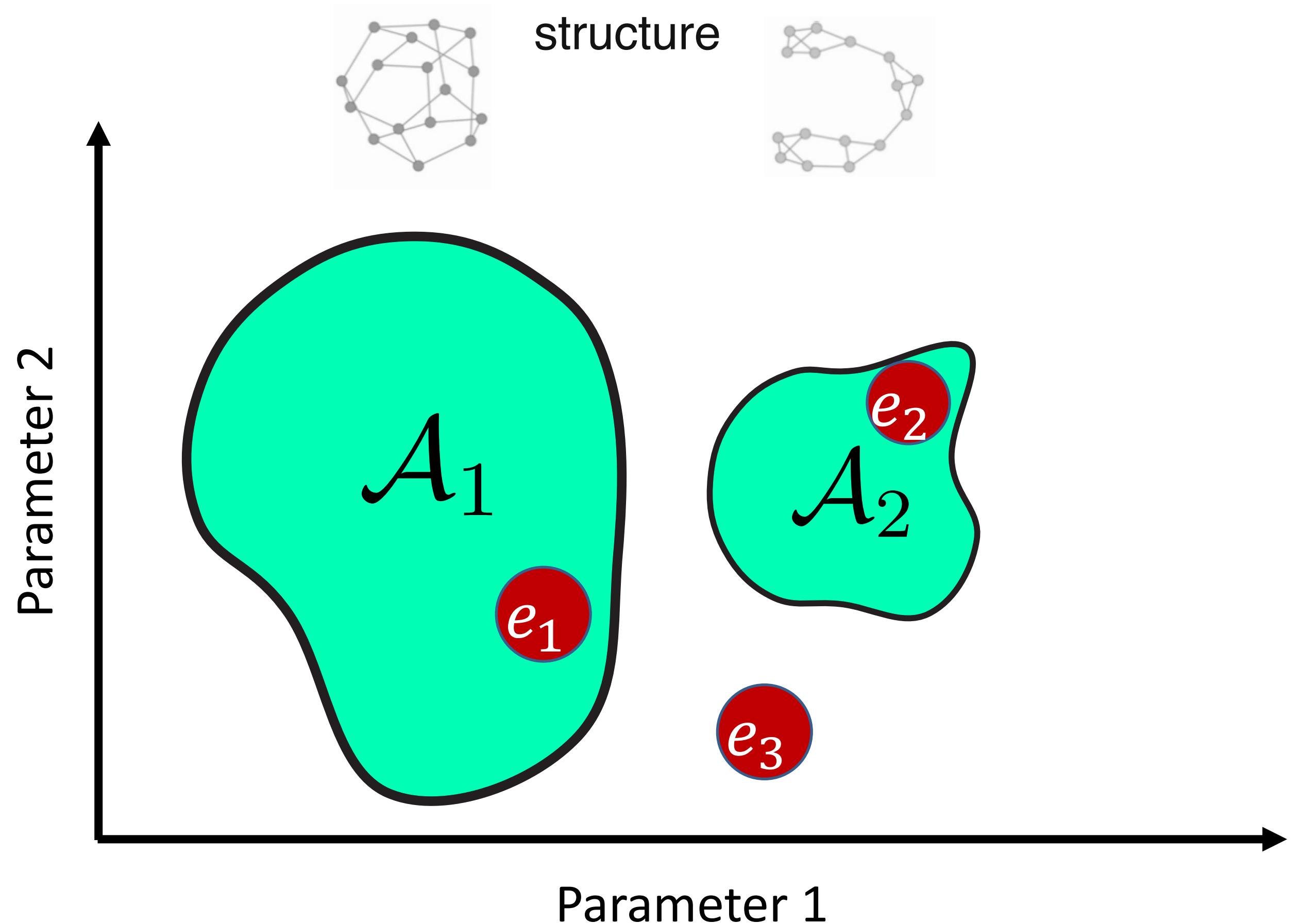
- Difficulty of the task
- Number of choices
- Proportion of experts
- Saliency of social signal
- Payoff
- Group size





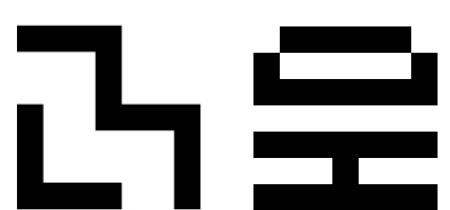
Parameters:

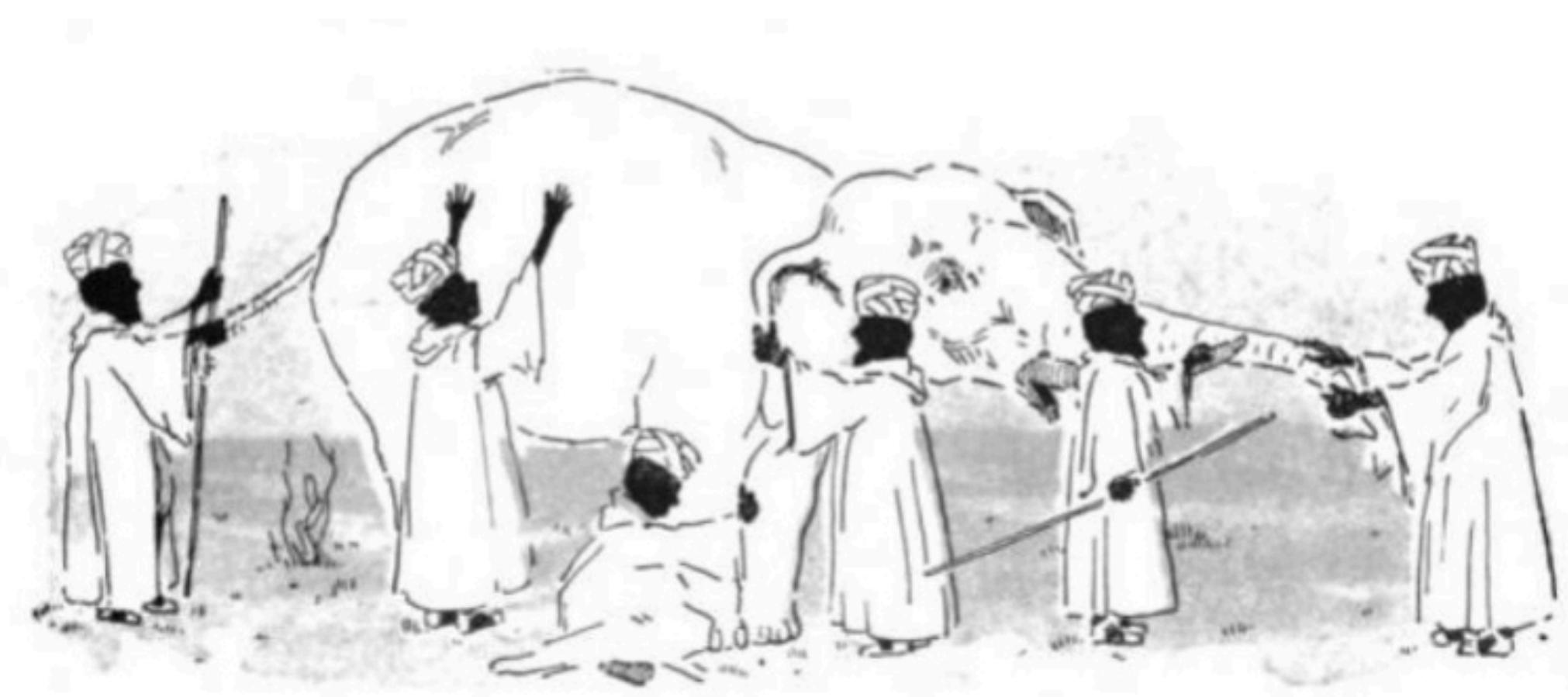
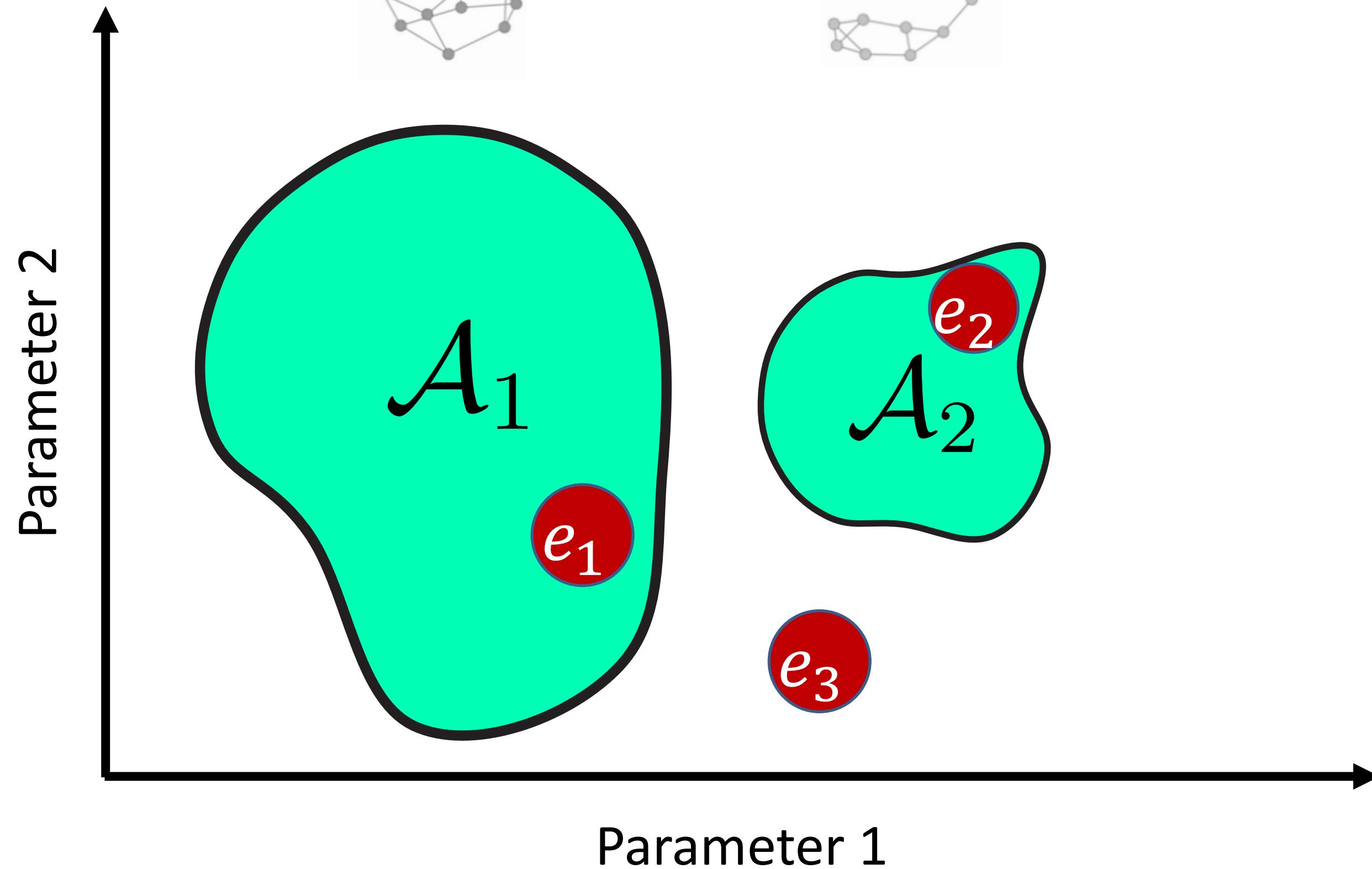
- Difficulty of the task
- Number of choices
- Proportion of experts
- Saliency of social signal
- Payoff
- Group size



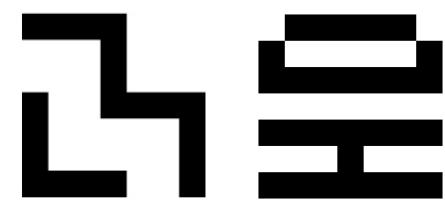
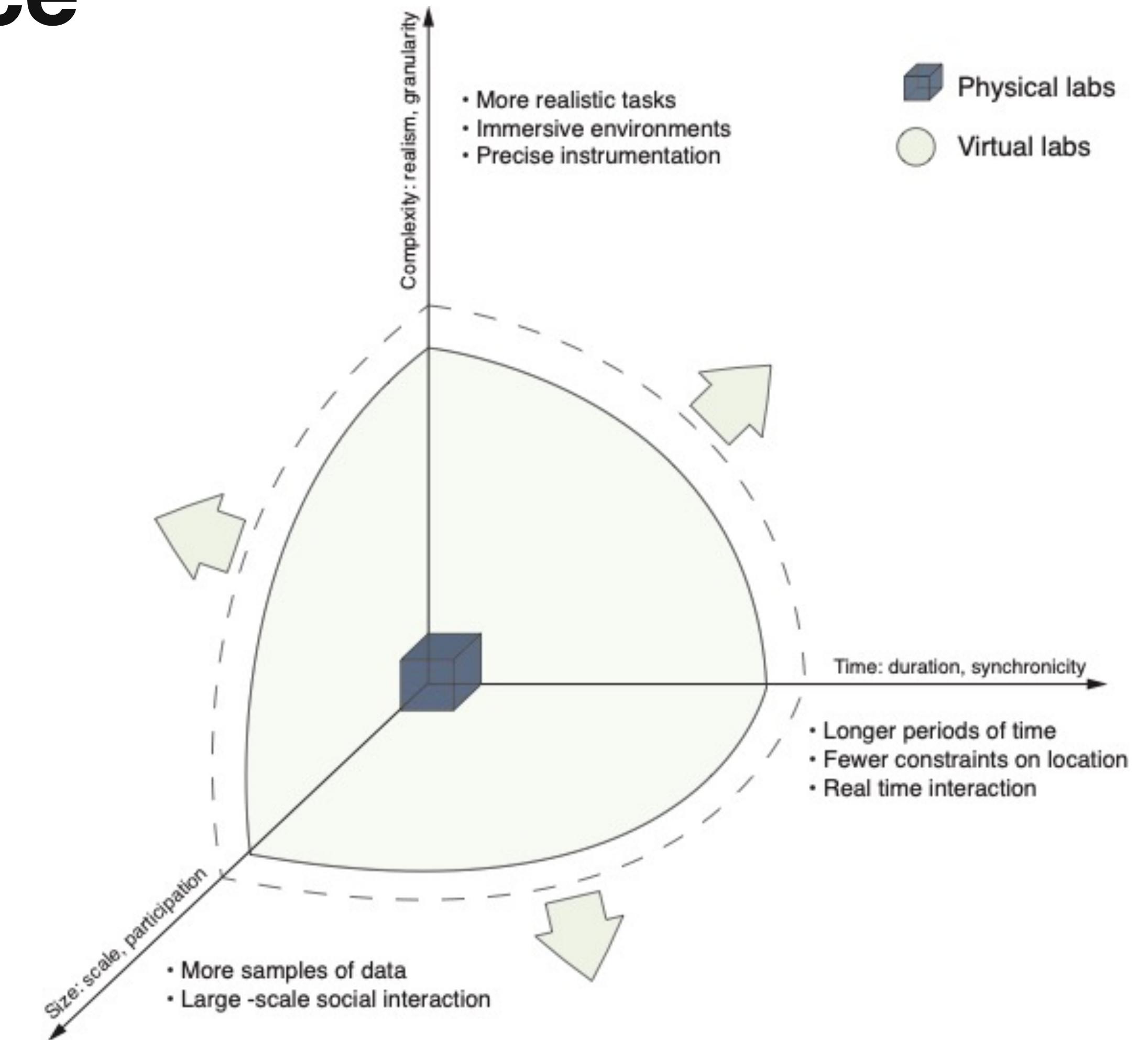
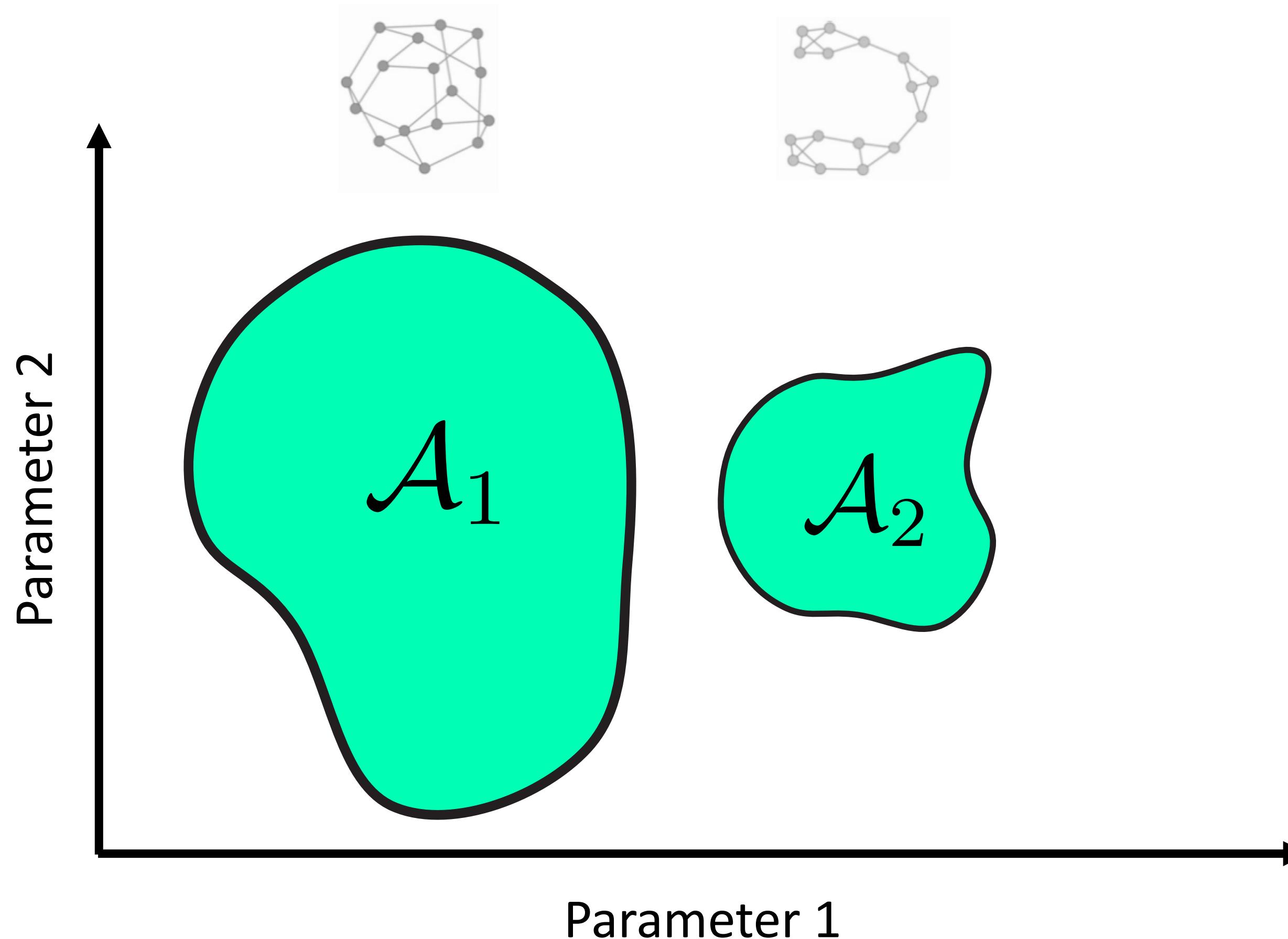
Parameters:

- Difficulty of the task
- Number of choices
- Proportion of experts
- Saliency of social signal
- Payoff
- Group size

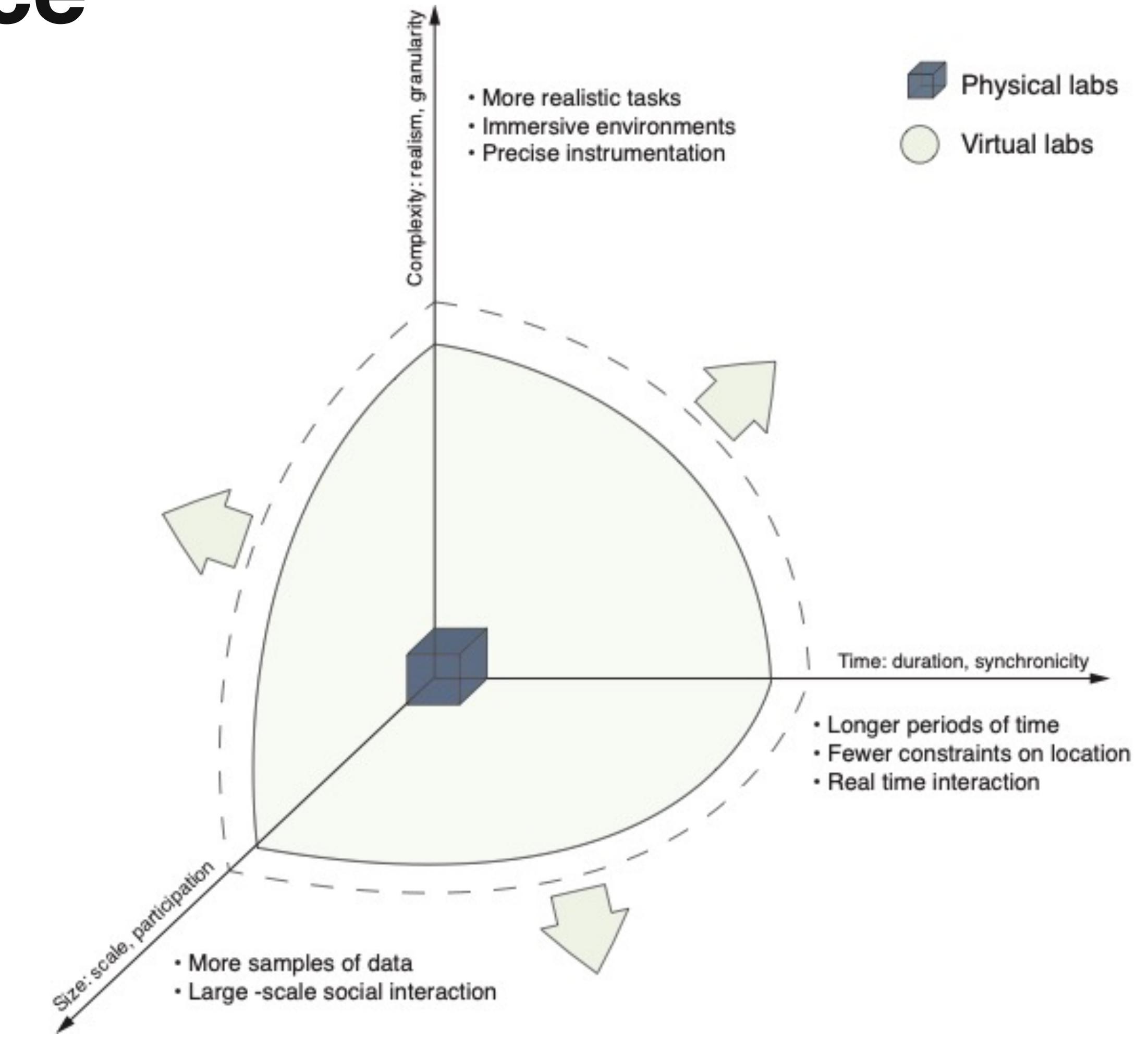
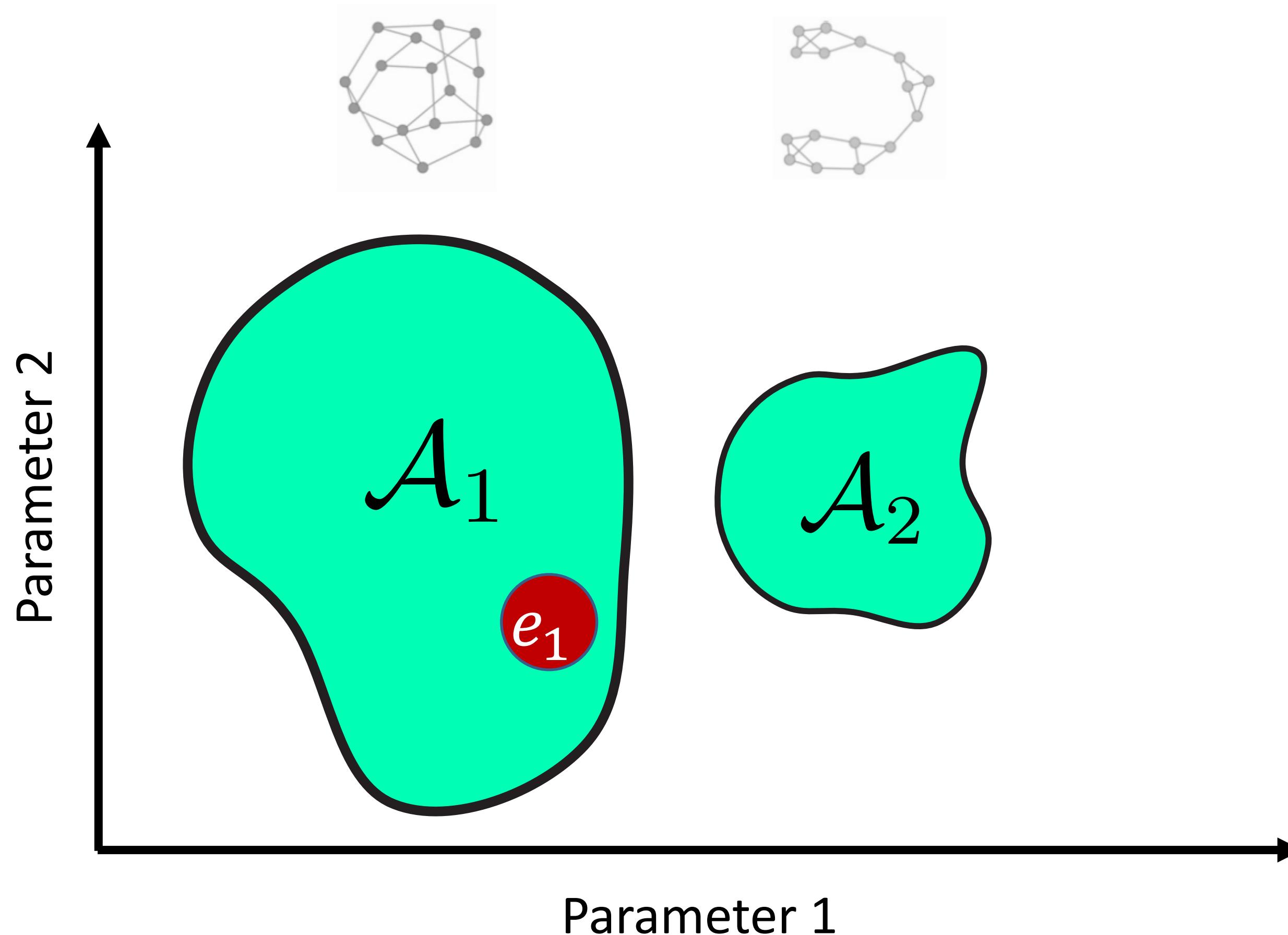




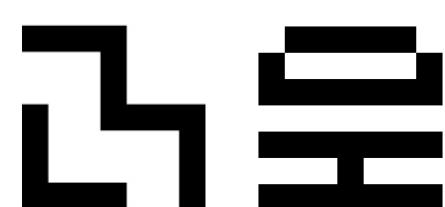
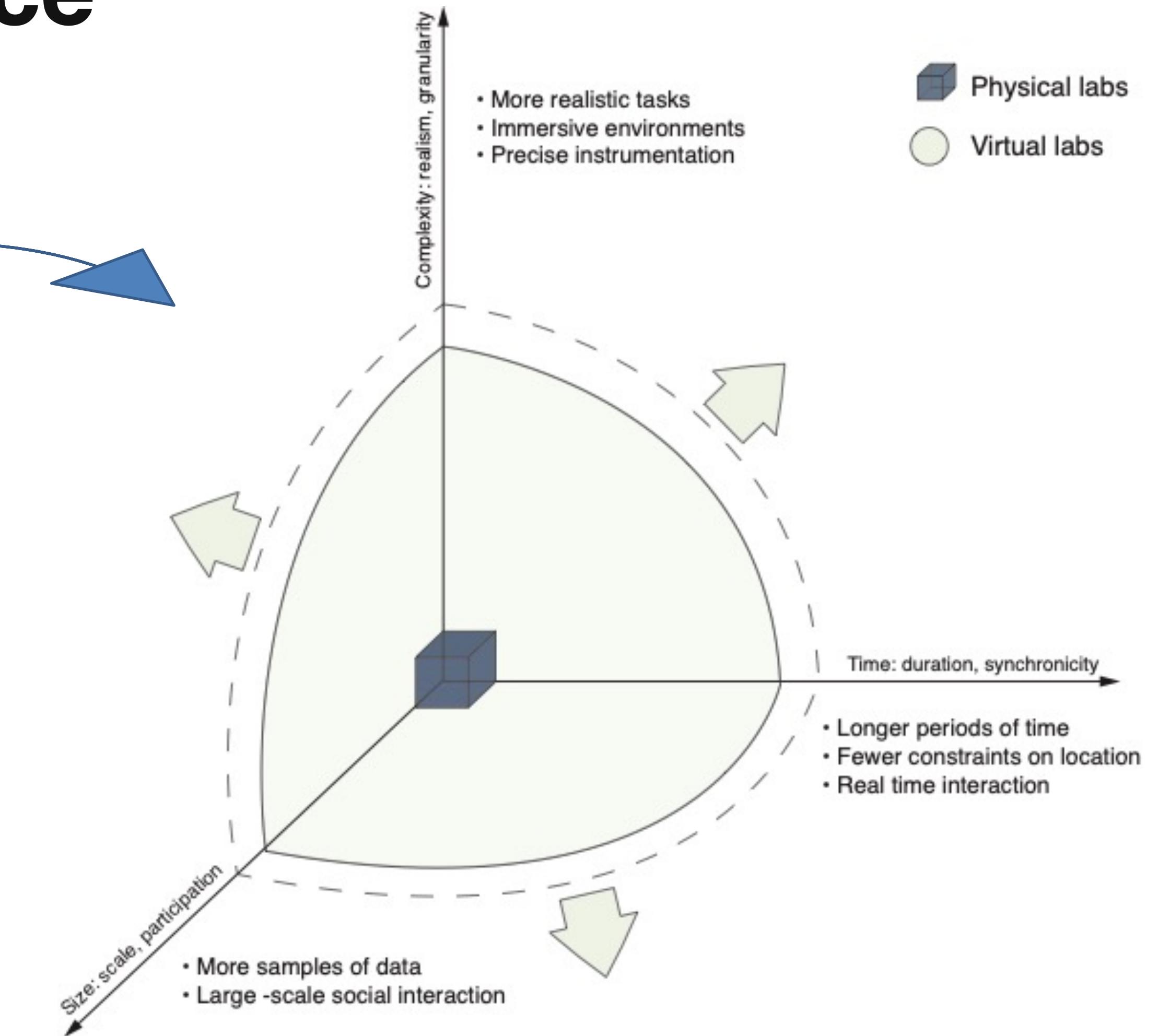
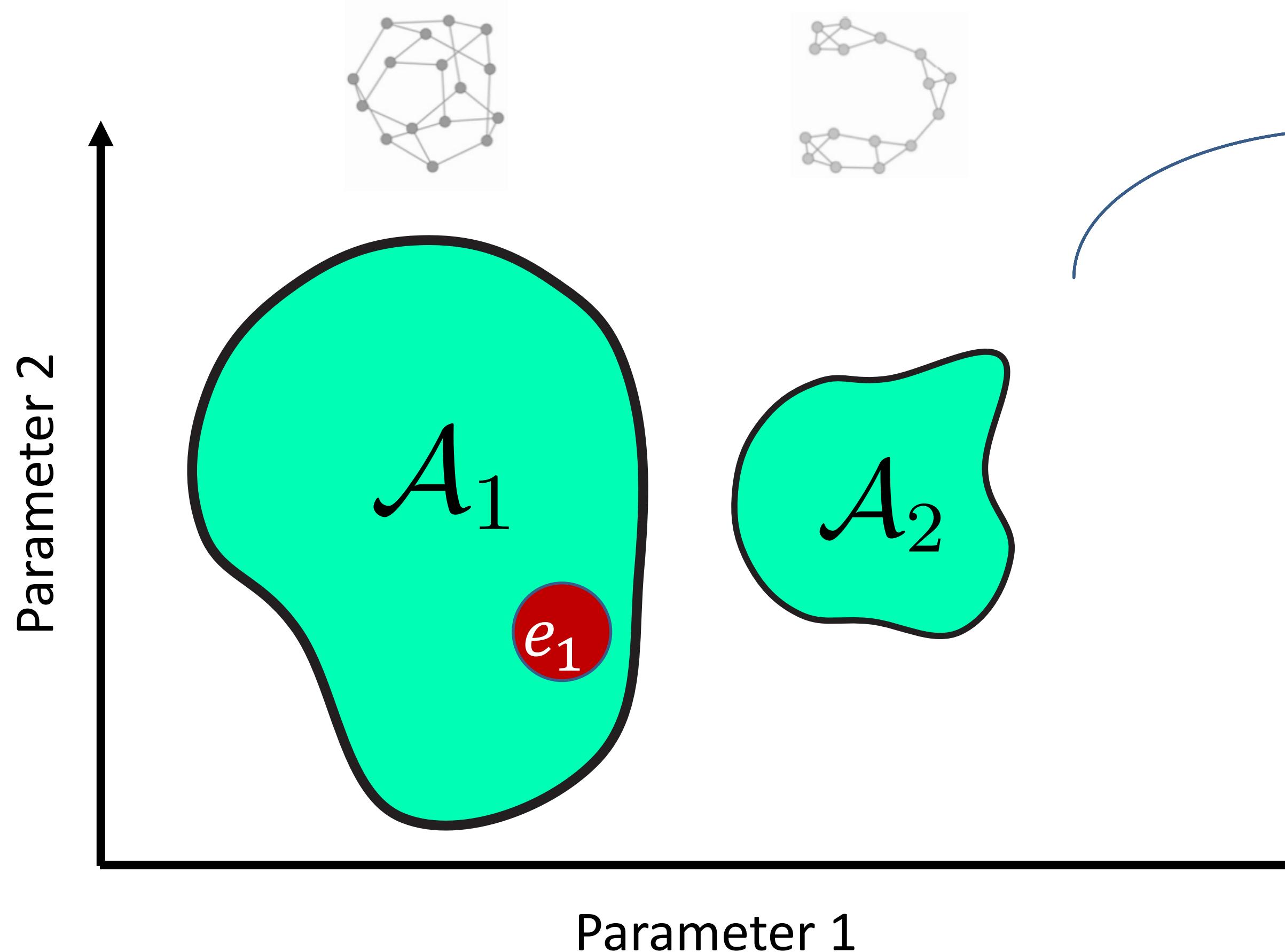
High-throughput social science



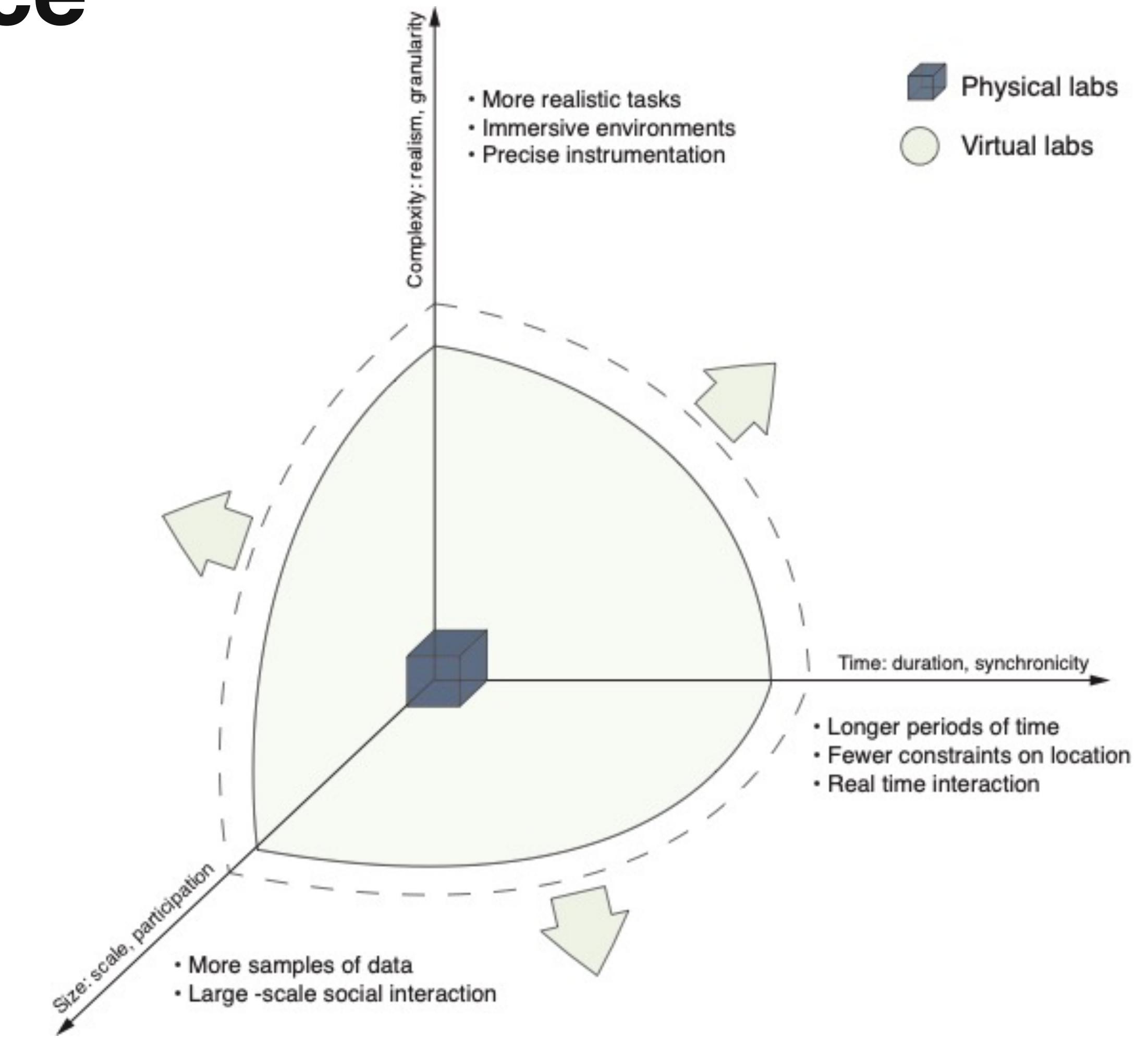
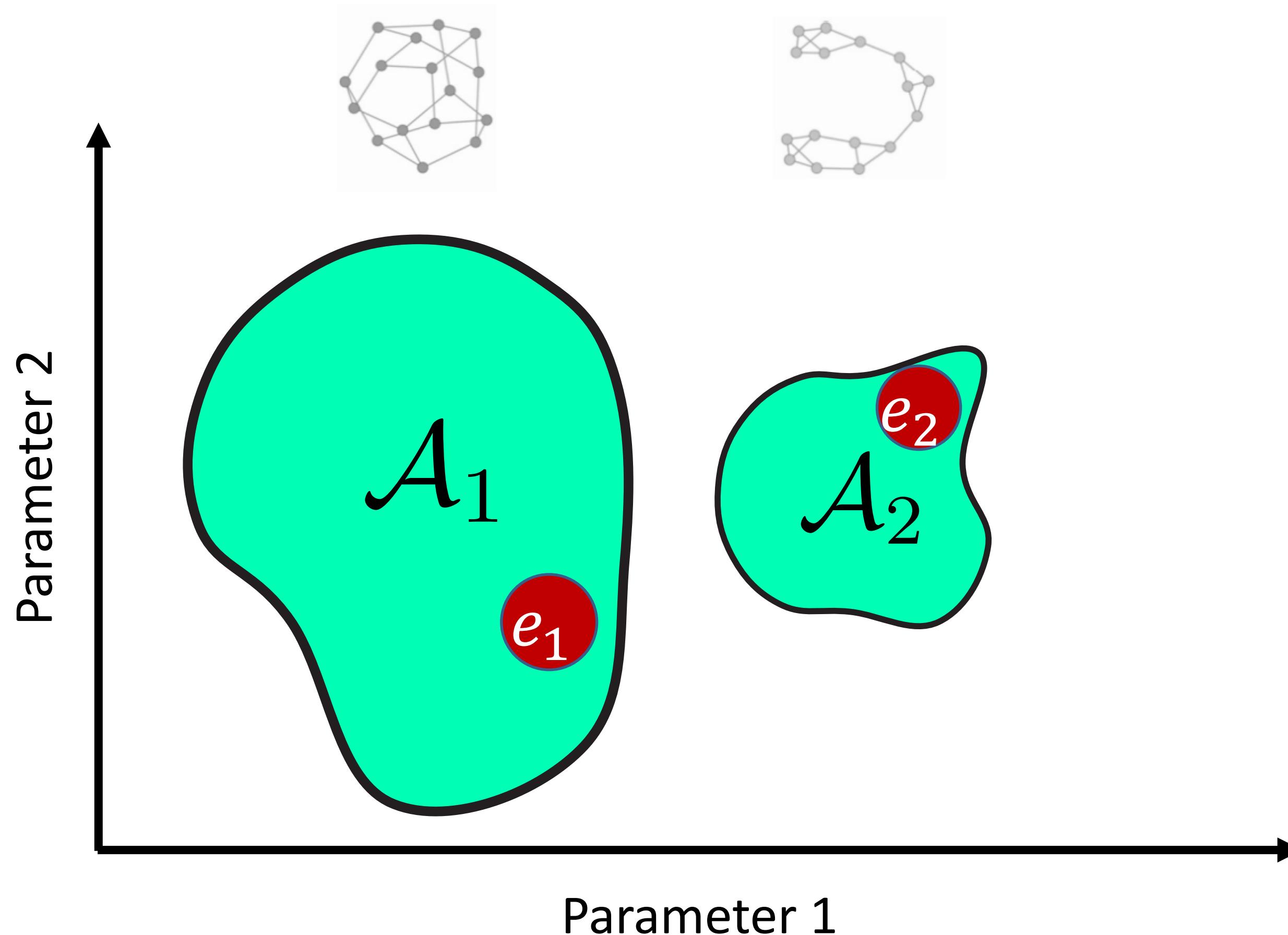
High-throughput social science



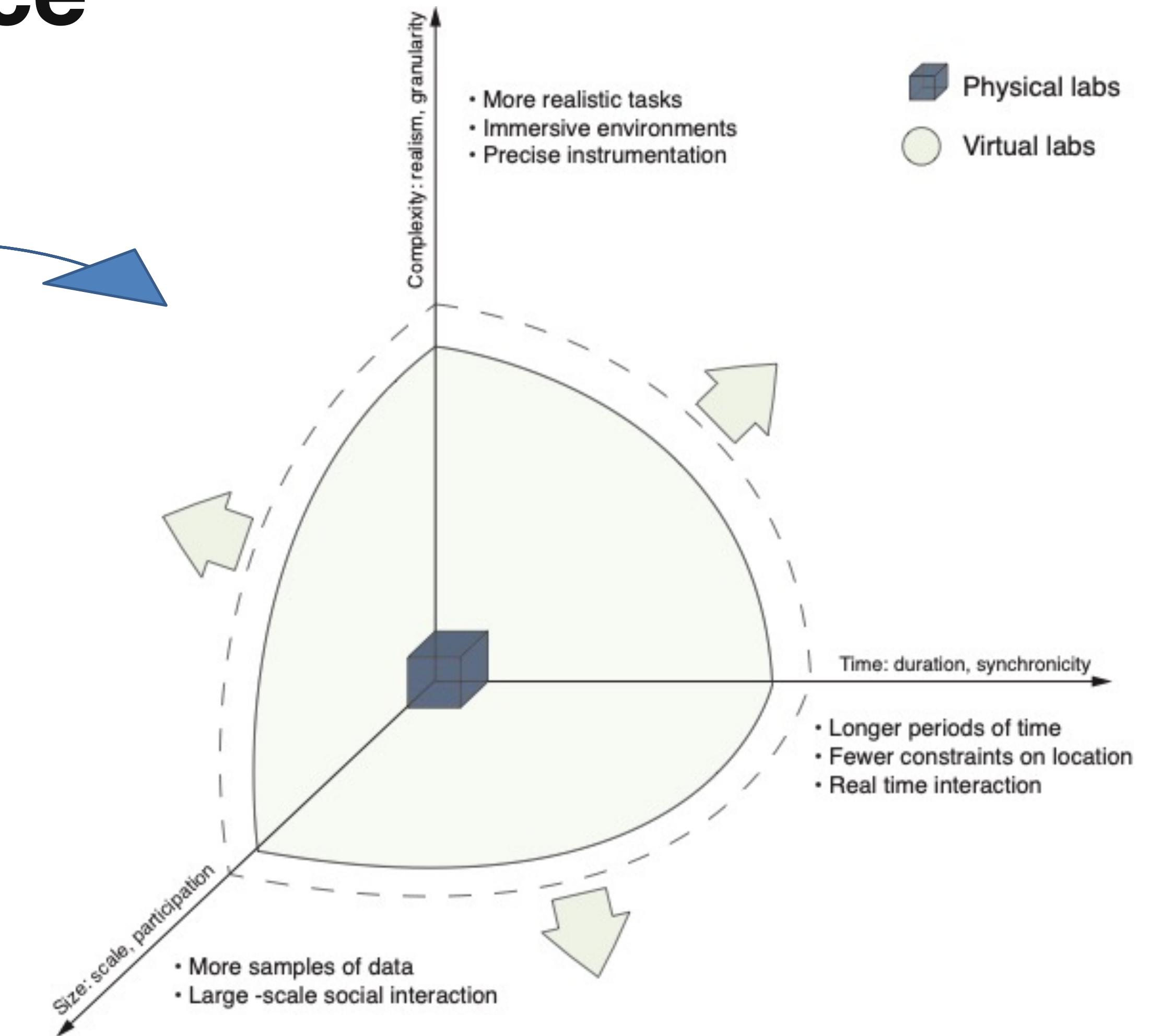
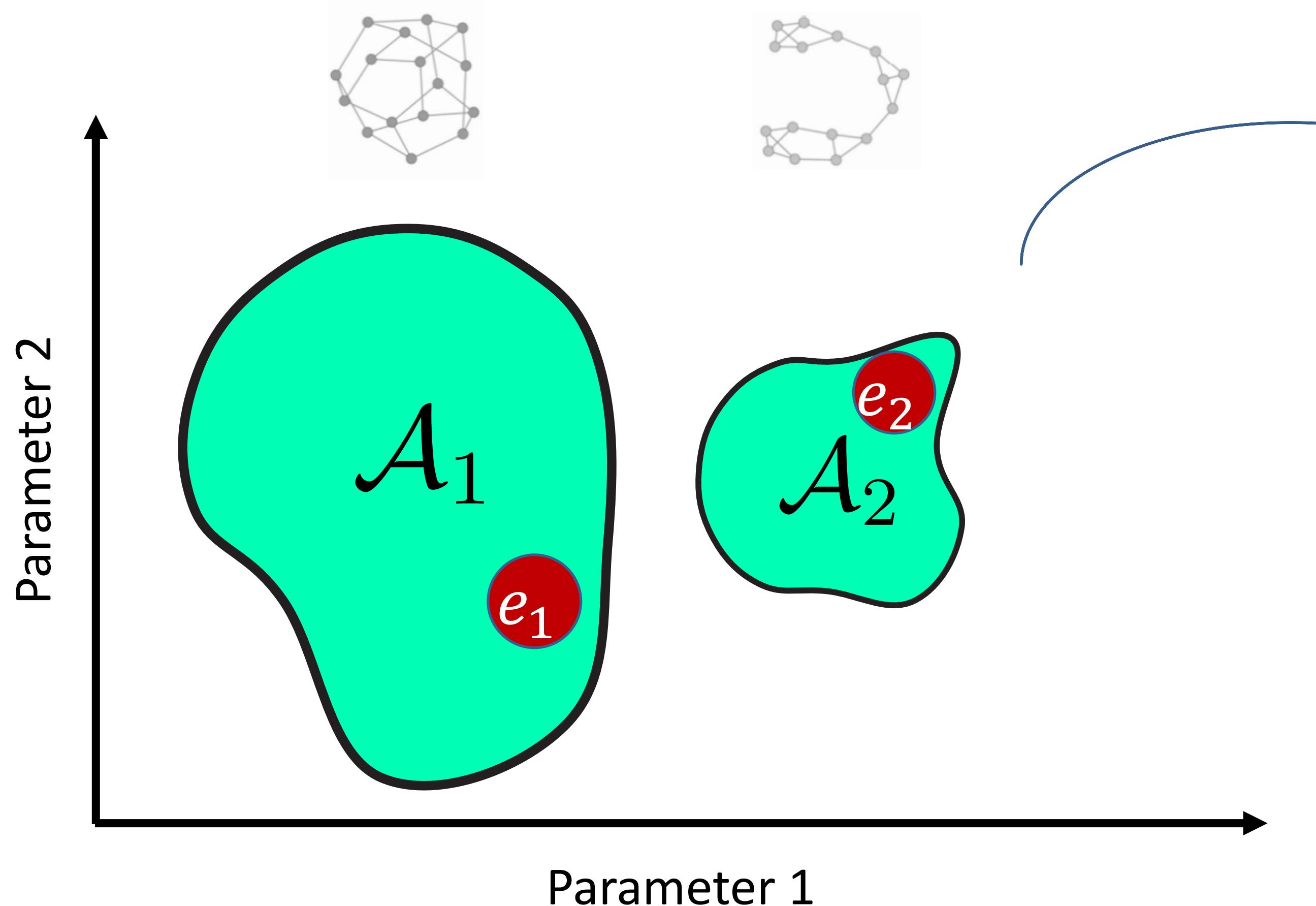
High-throughput social science



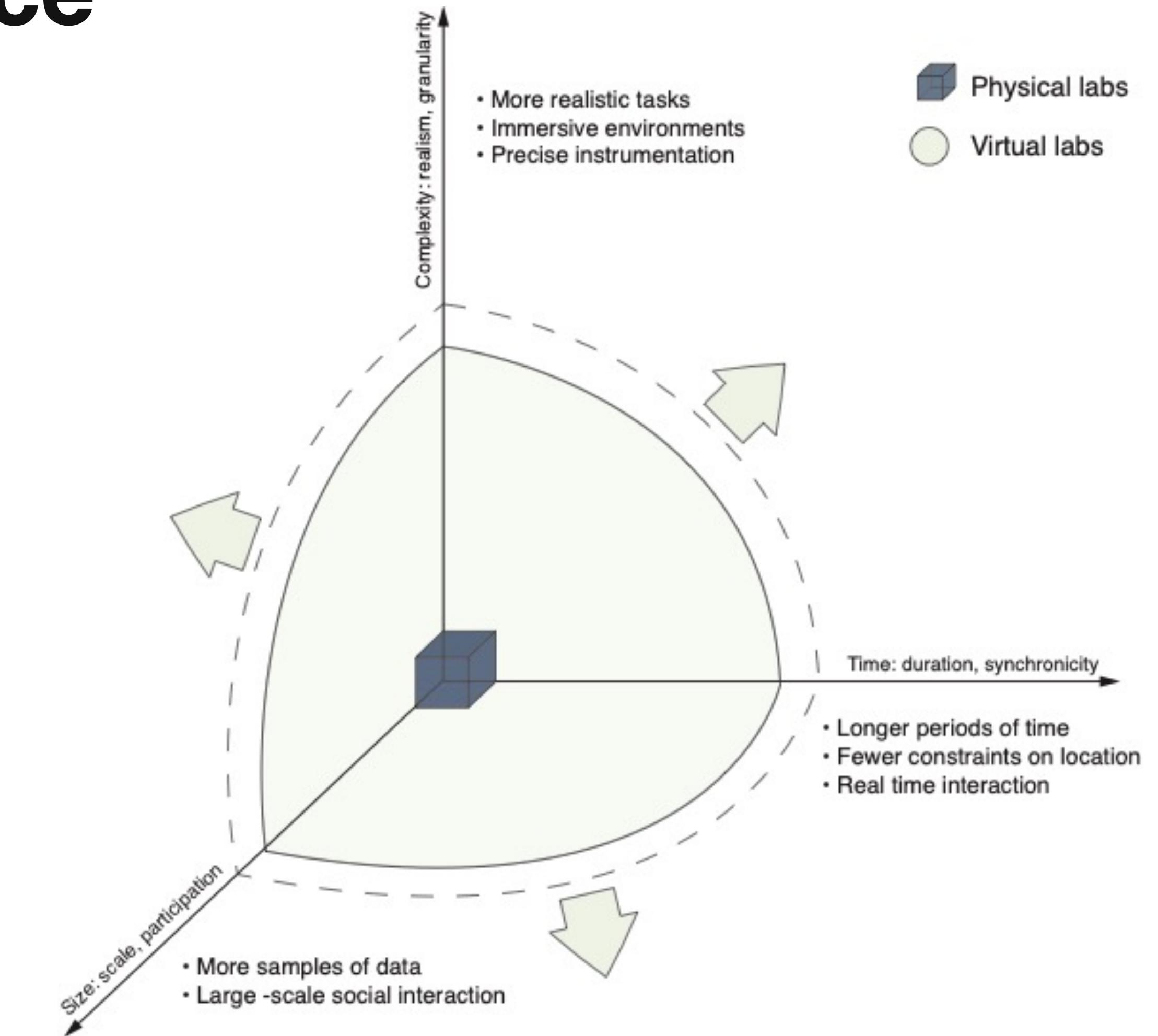
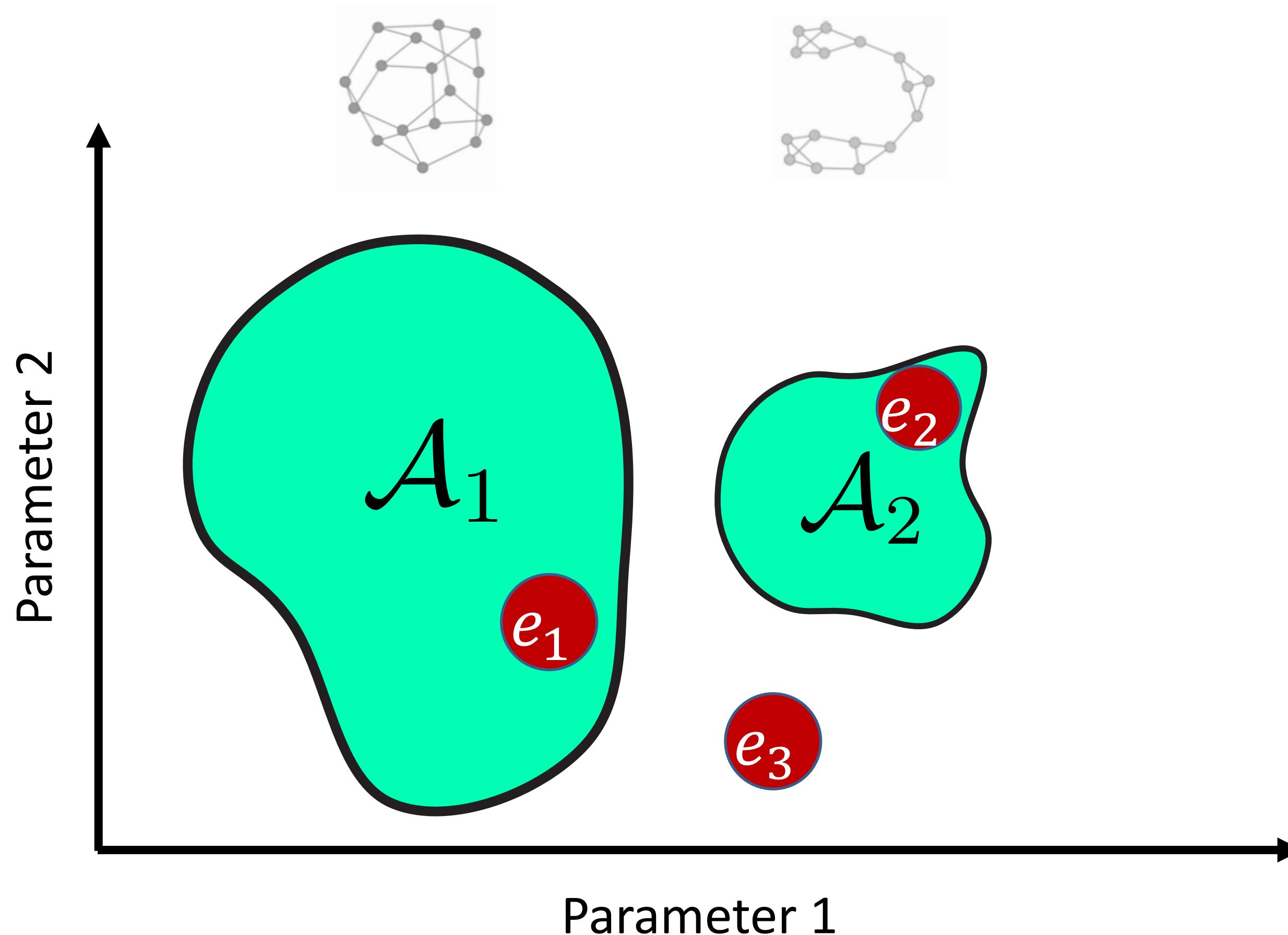
High-throughput social science



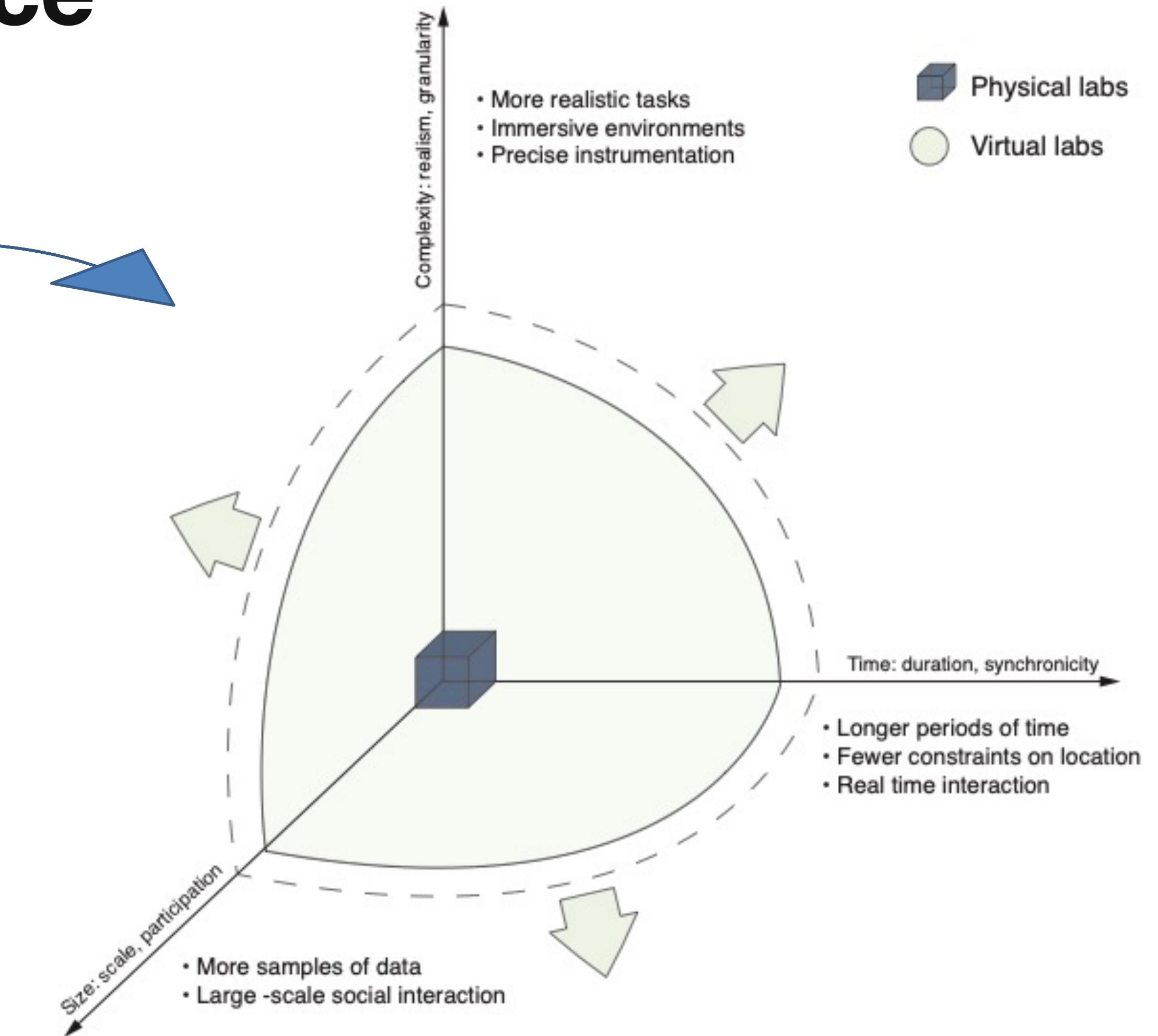
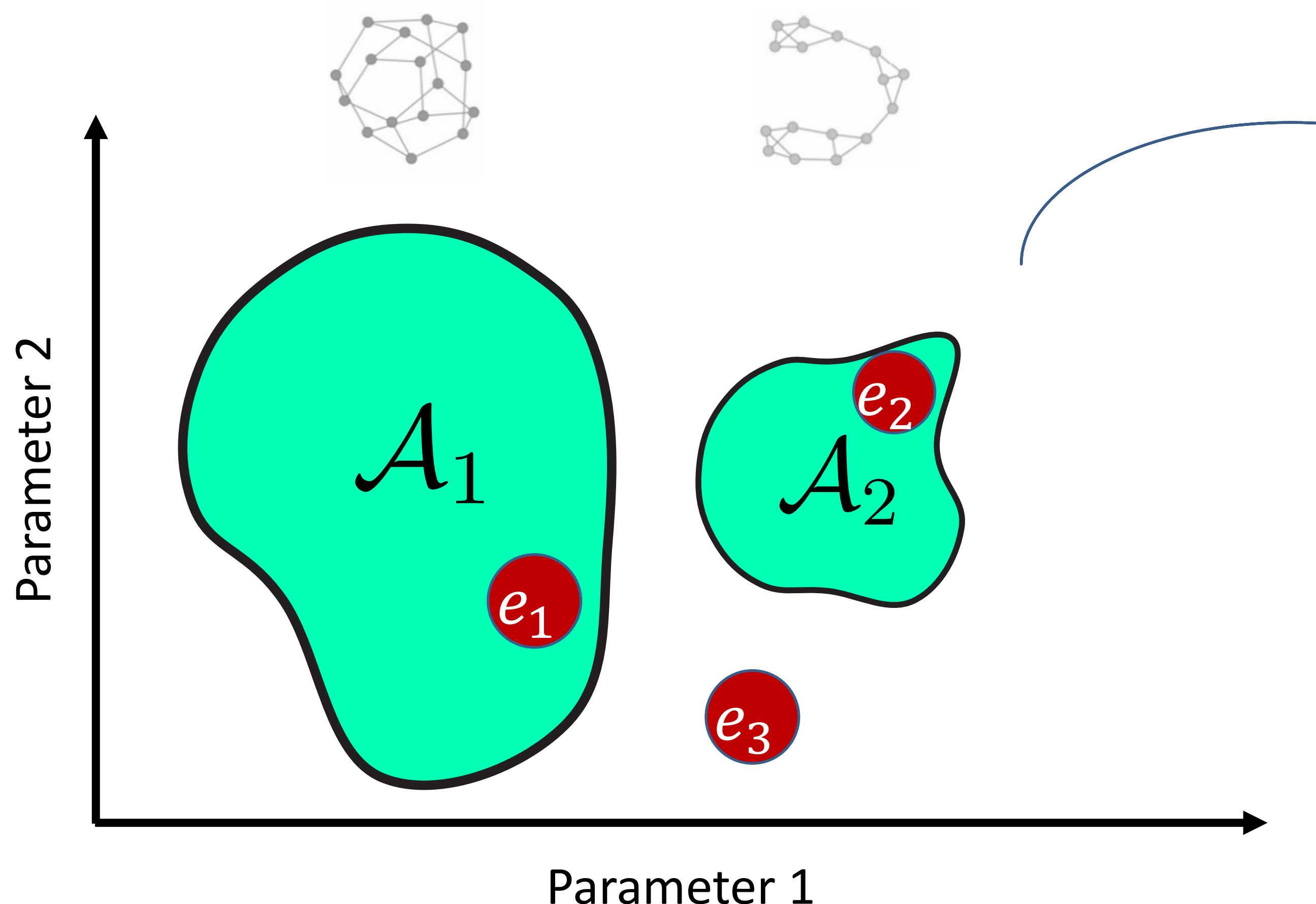
High-throughput social science



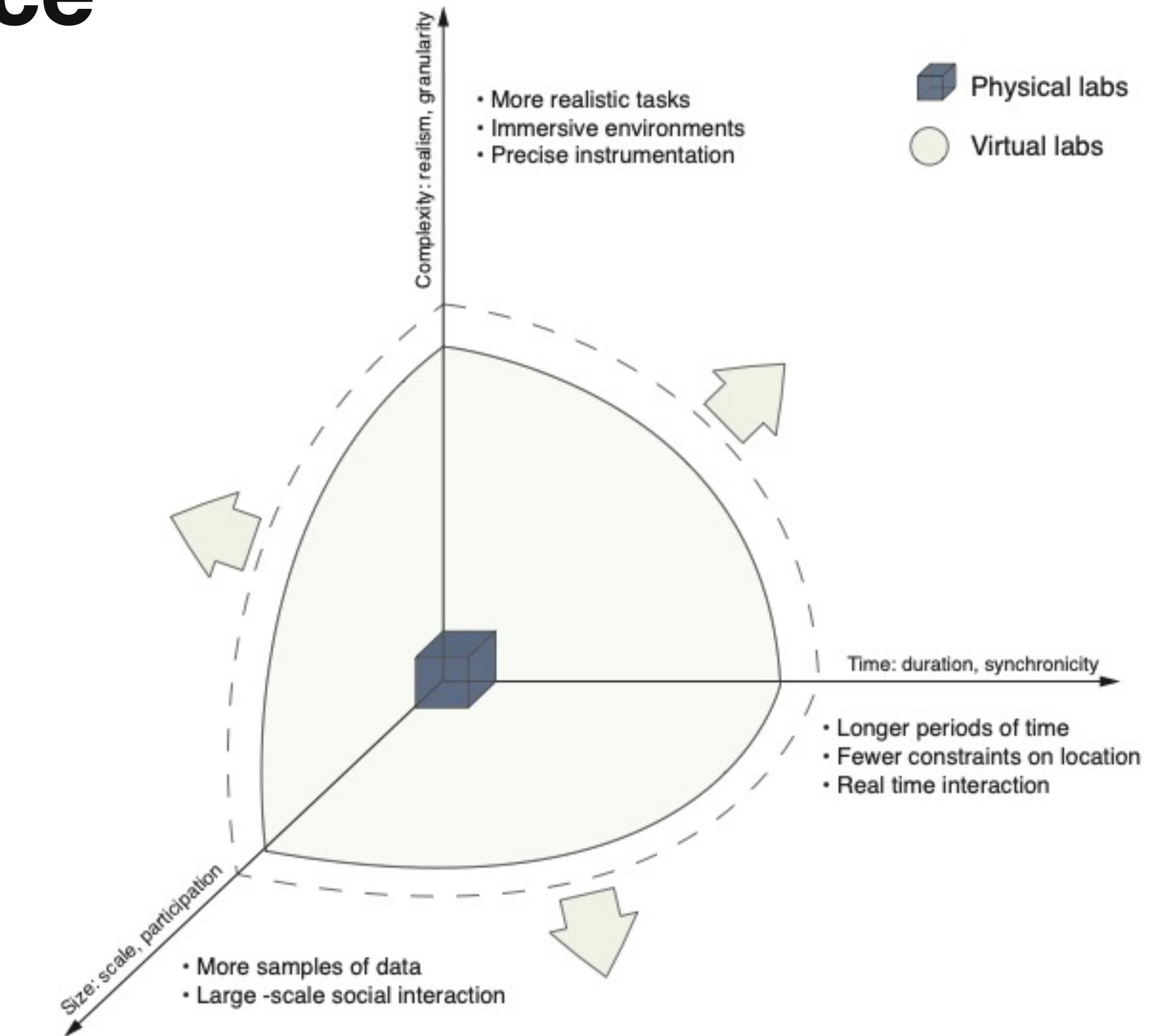
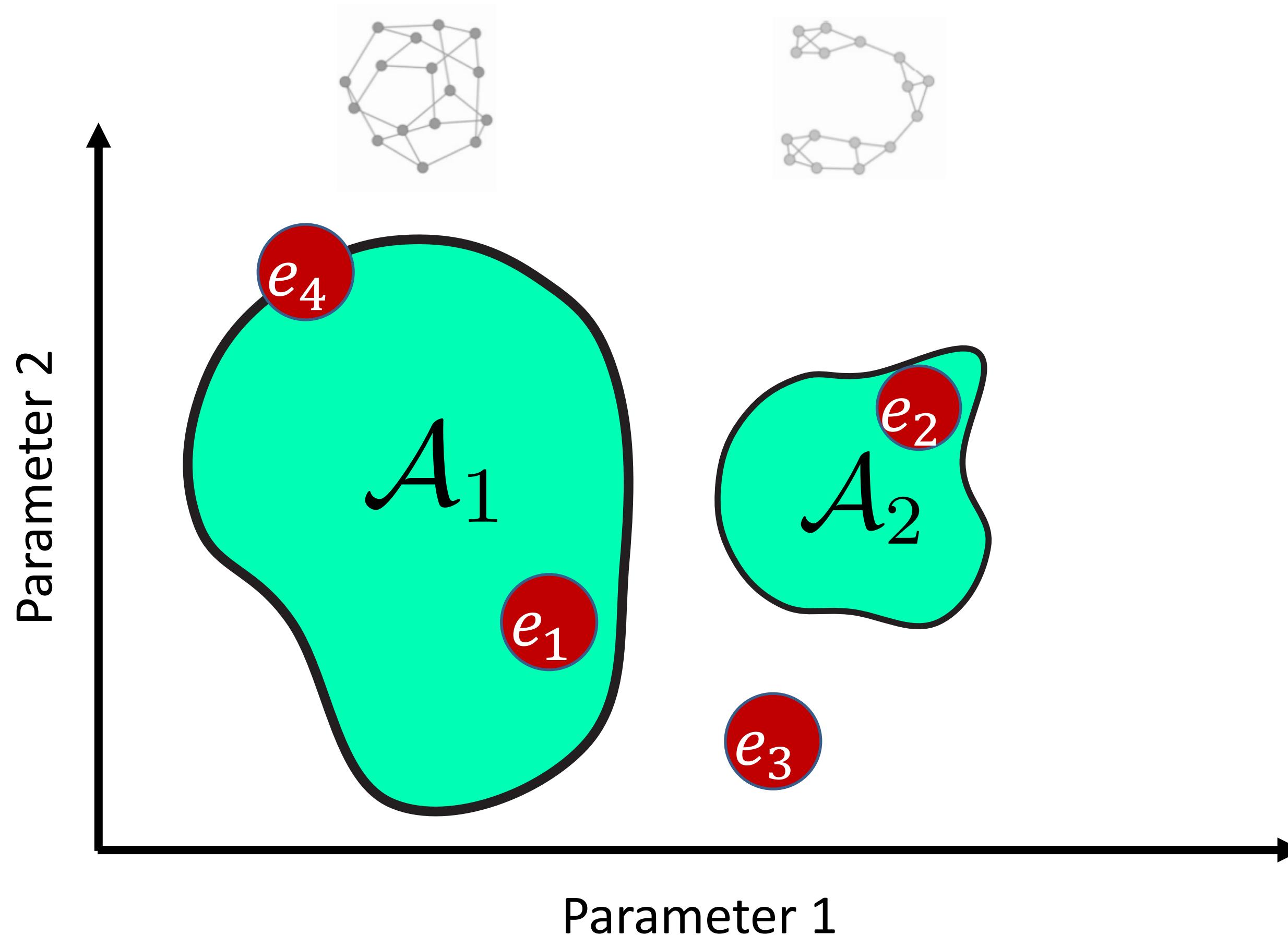
High-throughput social science



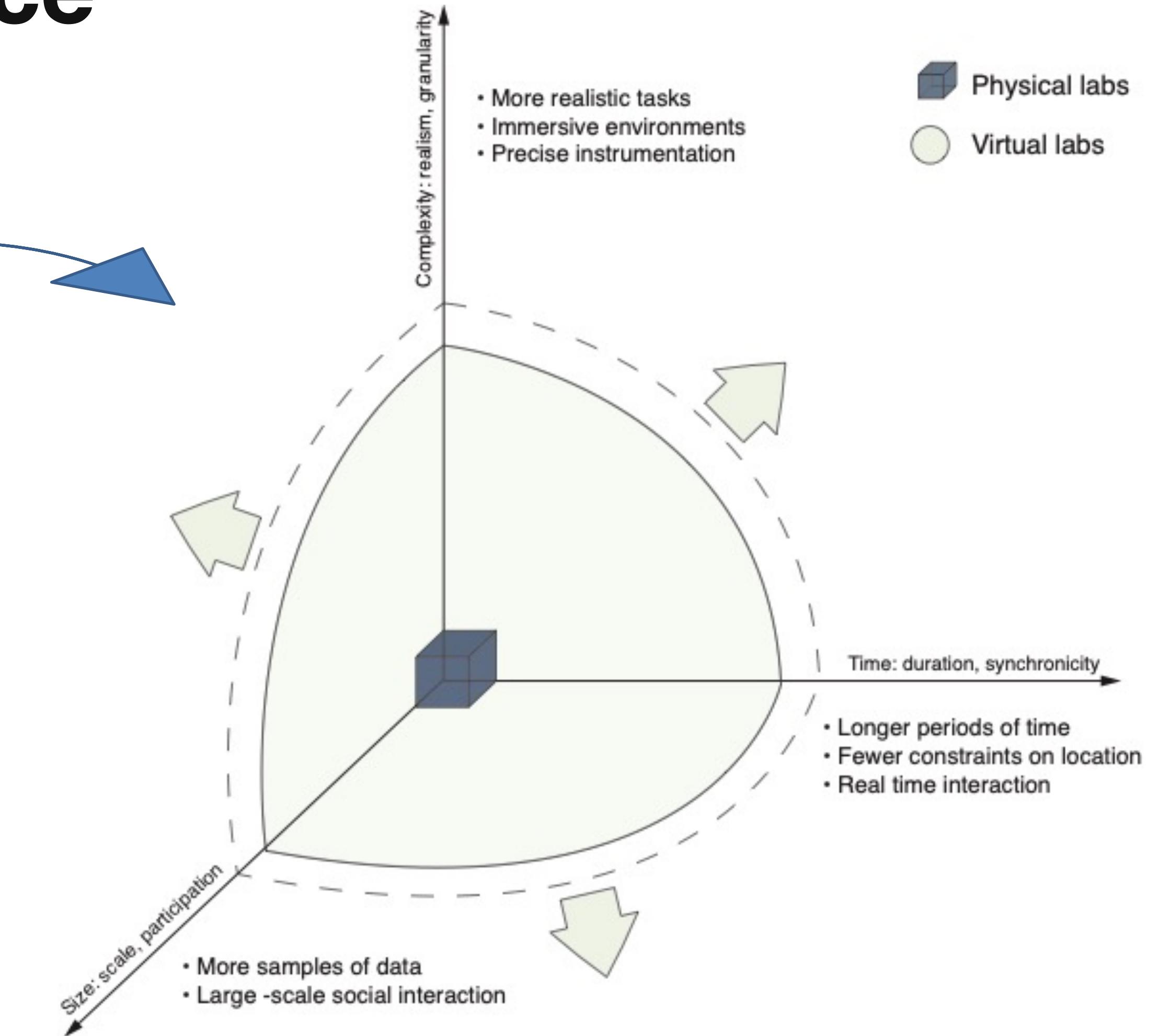
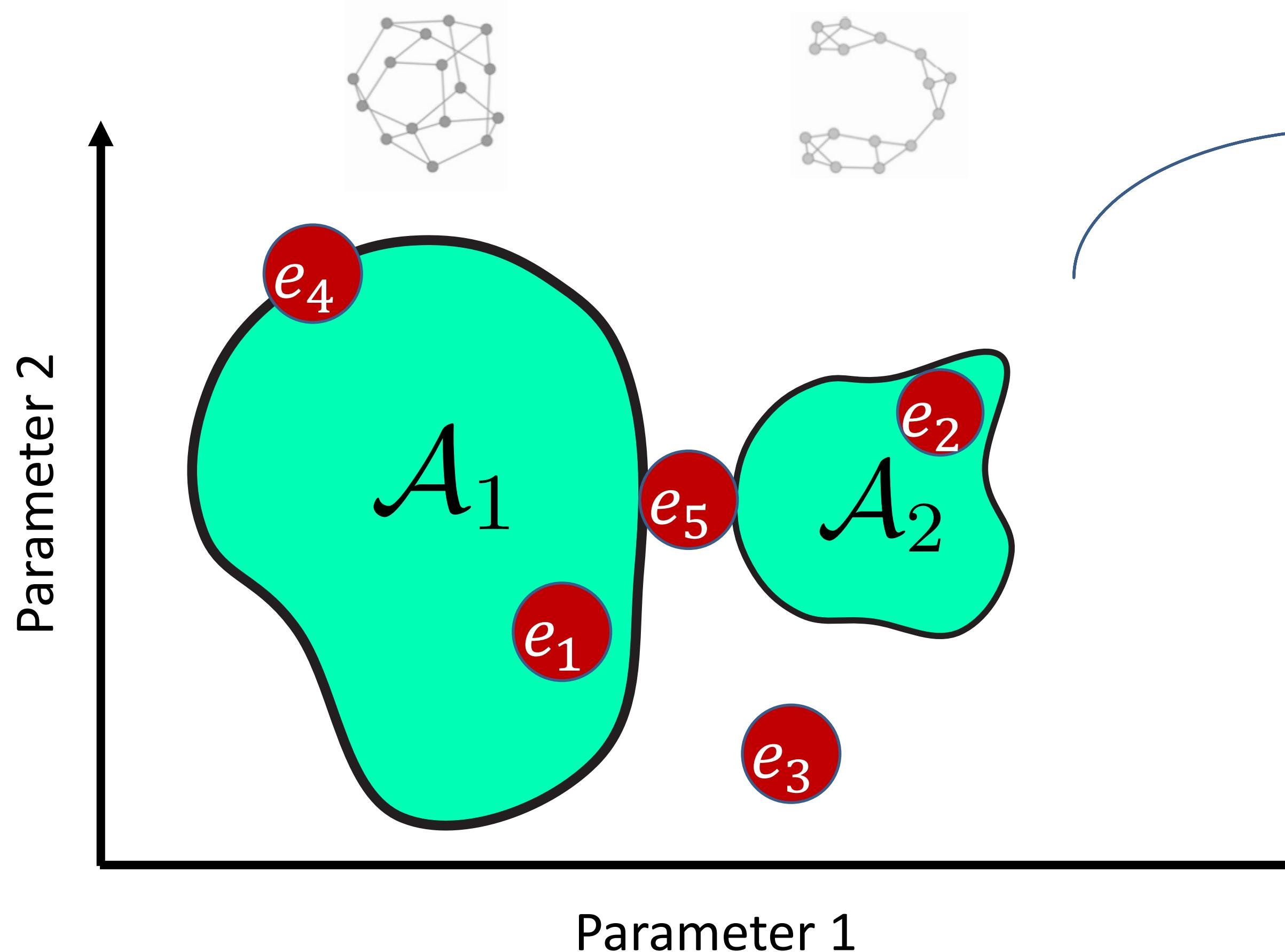
High-throughput social science



High-throughput social science



High-throughput social science



Empirica

Easy Multiplayer Interactive Experiments in the Browser

GETTING STARTED



Be Productive

Empirica will handle all the tedious logistics: from crossing your independent variables to form treatments, randomization, synchronization, data collection, and managing players. You get straight to what is unique about your research and what interests you, whatever that may be.



Be Expansive & Realistic

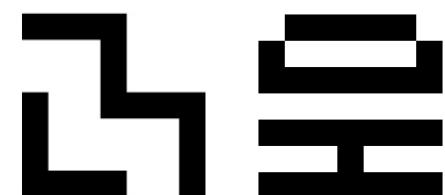
Study more complex tasks and set up interactions that happen over any period of time, from seconds to months, or among any number of people, from a single player to groups of any size. Don't be constrained by the limitations of standard behavioral experiments.



Simple or Complex

It is easy to implement simple A/B tests with independent players in Empirica. But it's just as easy to implement group experiments with real-time or asynchronous interactions in a factorial or within-subjects design, or designs involving multiple types of units and conditional logic.

<https://empirica.ly/>



High-throughput behavioral science using virtual labs

Abdullah M. Almaatouq
amaatouq@mit.edu

playerCount

The Number of players participating in the given game.

Name	Value
solo	1
small group	6
group	12

botsCount

The Number of bots that should participate in a given game

Name	Value
none	0
1	1
3	3
6	6

altersCount

The Number of connections for each player

Name	Value
no social	0
small social	3
8	8

stageDuration

The maximum time in seconds for each stage before moving to the next

Name	Value
30s	30

rewiring

Can the player change their alters on each round

Name	Value
false	false
true	true

feedbackRate

how frequent the feedback is (1 = every round; 0 = never)

Name	Value
never	0
always	1

feedbackNoise

The level of noise added to performance of the alters

Name	Value
none	0

shockRate

The rate at which we change difficulties for the players

Name	Value
rare	0.1

nRounds

This is the number of rounds for the game

Name	Value
1	1
full game	20

selfFeedback

whether to show feedback about one's performance

Name	Value
false	false
true	true

peersFeedback

whether to show feedback about one's peers

Name	Value
false	false
true	true

difficultyTypes

how to distribute difficulties across players

Name	Value
all easy	easy
equal mix	equal_mix
all hard	hard
all medium	medium
random	random

communicationMode

how can participants communicate

Name	Value
chatBox	chatBox
estimates	estimates
realTimeVideo	realTimeVideo

incentivesLevel

the reward is based on what

Name	Value
global	global
individual	individual
local	local

explanationLevel

the type of explanation provided by the bot

Name	Value
aLot	aLot
little	little
no	no

New Factor



High-throughput behavioral science using virtual labs

Abdullah M. Almaatouq
amaatouq@mit.edu

Name (optional)

playerCount *

solo (1) small_group (6) group (12)

botsCount

none (0) 1 3 6

altersCount *

no_social (0) small_social (3) 8

stageDuration *

30s (30)

rewiring *

false true

feedbackRate *

never (0) always (1)

feedbackNoise

none (0)

shockRate *

rare (0.1)

nRounds *

1 full_game (20)

selfFeedback

false true

peersFeedback

false true

difficultyTypes

all_easy (easy) equal_mix all_hard (hard)

all_medium (medium) random

basePay

2

conversionRate *

0.022 0.0275 0.03 0.035

explanationLevel *

You decide what you want to vary
(treatment conditions based on research question)

But you have to ***explicitly*** recognize what other parameter choices you are making

New Lobby Configuration X

Name (optional)

Timeout Type ?
Lobby ▼

Timeout Duration in Seconds *
 ^ ▼

Timeout Type ?
Fail ▼

Create Lobby Configuration

There are 2 timeout types:

- **Lobby:** the timeout start when the first player reaches the lobby and runs out for all the players whether they have even reached the lobby or not.
- **Individual:** the timeout is started for each player as they reach the room. Some players might time out before all players are in the lobby, they might continue waiting for another timeout period. They might also leave the game and a new player can replace them. The lobby itself never times out.

The Timeout Strategy determines what to do in case people are waiting in the lobby for longer than the timeout duration. Available strategies:

- **Ignore:** start the game anyway
- **Fail:** take the player to the exit survey
- **Bots:** fill the missing players slots with bots from timeoutBots.



 **New Batch**

Assignment Method
Complete

Treatments

Treatment	Lobby Configuration
T1	default 3  Remove
T2	default 3  Remove

Add a new treatment... 

Game Count
6

Create Batch

Assignment methods:

- Simple randomization
- Complete randomization
- Block randomization
- No randomization
- Custom randomization



You .set() and .get() values. They will be reactive! (real-time synchronization)

```
1
2
3  onGameStart = (game, players) => {};
4
5  onRoundStart = (game, round, players) => {};
6
7  onStageStart = (game, round, stage, players) => {};
8
9  onStageEnd = (game, round, stage, players) => {};
10
11 onRoundEnd = (game, round, players) => {};
12
13 onGameEnd = (game, players) => {};
```



Example experiments

Empirica

Round 3 > Response > **Interactive Response** > Round Outcome

Your Profile



Total score
10

Timer
159

The sings are great, but the video director and audio editor (Giovanna Nocetti) is out! Total amateur!

The other player:



Negative Positive

100% not sure **56%** 100%

Negative Positive

100% 50% **32%** not sure 50% 100%

✓ Submit



Example experiments

Empirica

Round 1 > Response > **observe C** > observe B > observe A > observe C > observe A > observe B > outcome

Your Profile

B



Total score
0

Timer
16

Drag the arrow to change your guess!

✓ Submit

You observe:





Example experiments

Empirica

New Player Reset current session About

Round practice > Training - Hard > Super Hard > Hard > Very Hard > Medium > Easy

Timer Score
08:34 0

This is practice round and the Score will not count

Constraints

X A and B must be neighbors.

- A and D can't live in the same room or be neighbors.
- A and G can't live in the same room or be neighbors.
- C and D must live in the same room.

X C and F must be neighbors.

- C and G can't live in the same room or be neighbors.
- D and G can't live in the same room.
- H and I can't live in the same room or be neighbors.

Payoff

Rooms	101	102	103	104	105	106
Student A	27	61	67	40	74	52
Student B	97	46	41	56	43	71
Student C	23	35	80	39	92	58
Student D	87	32	79	93	68	49
Student E	34	69	77	96	38	53
Student F	98	72	76	92	31	66
Student G	57	26	39	75	21	43
Student H	61	59	36	65	20	41
Student I	58	23	60	38	45	33

x Unsatisfied ✓ Satisfied

Pink (You) **Blue** **Green**

Total Score **0**

11:02:54 Blue moved B to Room 104.
11:02:55 Blue started moving C.
11:02:55 Blue moved C to Room 102.
11:03:38 Green started moving F.
11:03:38 Green moved F to Room 105.
11:03:39 Green started moving H.
11:03:39 Green moved H to Room 102.

Green Hi guys
 Green should I control A, B, and C while Pink controls D, E, F and Blue the rest?
 Blue Yes, I think that's a good idea
 You but there is a conflict between A and G
 You I think each should take the set of students that share conflict?

Enter chat message

Example experiments

Empirica

New Player Reset current session About

Round 1 > Response > Feedback > Question

Your Profile



Total score
0

Timer
24

Will there be a locally-transmitted case of the Zika virus in Singapore between 1 July 2018 and 31 August 2018?

HR Holton Recorder @HoltonRecorder

According to the official government, there are no currently active clusters of Zika in Singapore. Travel restrictions over disease concerns have been lifted in the region.

05:33 PM - Jul 04, 2018

31,888 people are talking about this

Alexandria Echo Press @echopress

New vaccines against Zika are expected to be widely available in South East Asia in the next month.

05:33 PM - Jul 04, 2018

31,888 people are talking about this

0% 100%

✓ Submit



Papers using Empirica

- Becker, Joshua, Ethan Porter, and Damon Centola. "The wisdom of partisan crowds." *Proceedings of the National Academy of Sciences* 116.22 (2019): 10717-10722.
- Feng, Dan, et al. "Exploring Improvisational Approaches to Social Knowledge Acquisition." *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2019.
- Almaatouq, Abdullah, et al. "Collective Performance Across Tasks of Varying Complexity" under review
- Almaatouq, Abdullah, et al. "The Wisdom of the Network: How Adaptive Networks Promote Collective Intelligence" under review



List of Platforms for Online experiments

Group-Behavior

- nodeGame <https://nodeGame.org>
- Wextor (pioneer) <https://www.wextor.eu>
- Otree (large base) <https://www.otree.org>
- Lioness (new) <https://lioness-lab.org>
- Breadboard (networks) <http://breadboard.yale.edu>
- TurkServer (discontinued) <https://turkserver.readthedocs.io>

Individual

- JSPsych (many plugins) <https://www.jspsych.org>
- PsiTurk (groups possible) <https://psiturk.org>



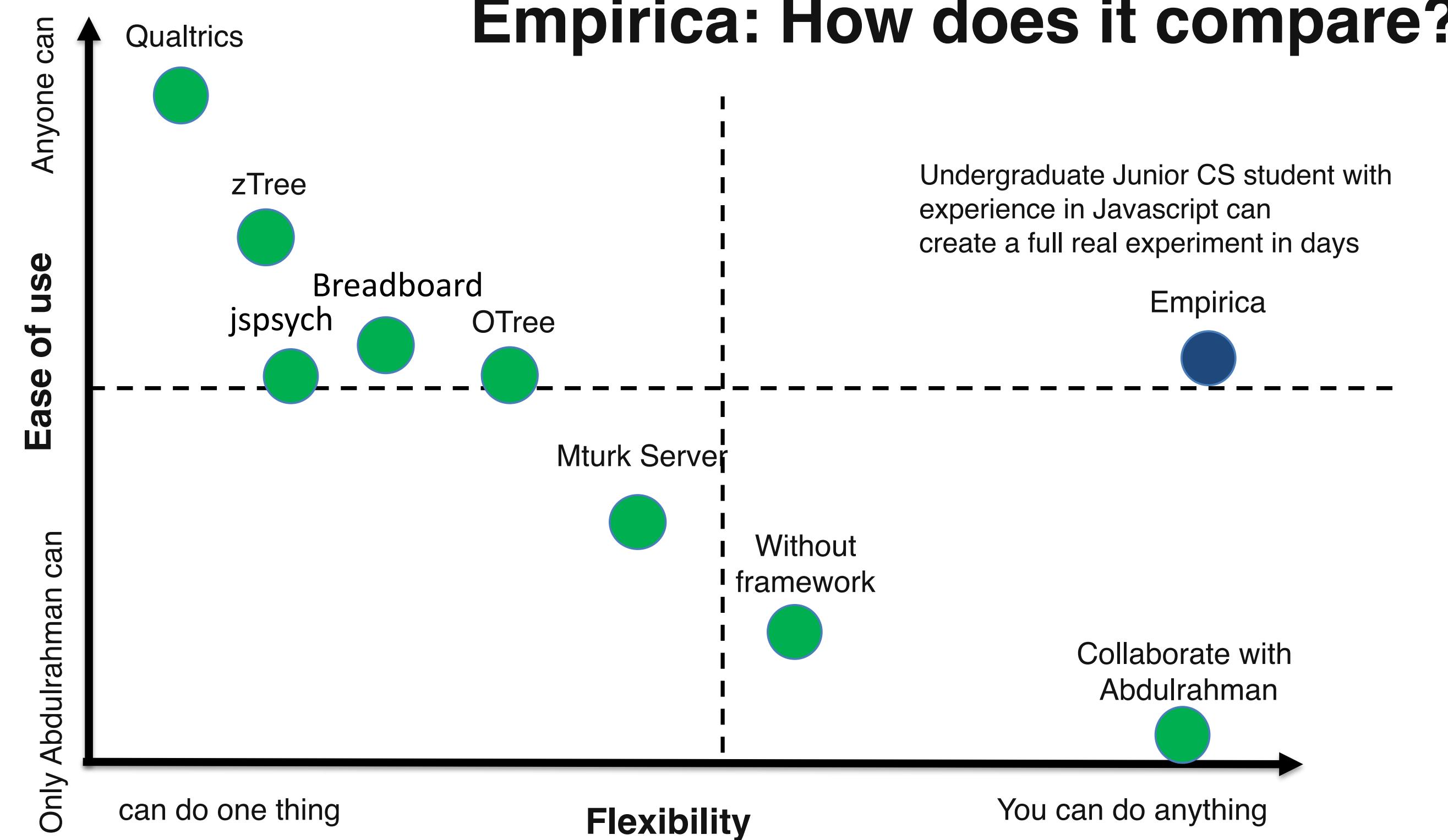
List of Platforms for Online experiments

Group-Behavior

- nodeGame <https://nodeGame.org>
- Wextor (pioneer) <https://www.wextor.eu>
- Otree (large base) <https://www.otree.org>
- Lioness (new) <https://lioness-lab.org>
- Breadboard (networks) <http://breadboard.yale.edu>
- TurkServer (discontinued) <https://turkserver.readthedocs.io>

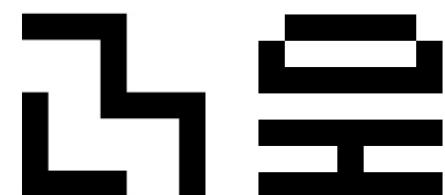
Individual

- JSPsych (many plugins) <https://www.jspsych.org>
- PsiTurk (groups possible) <https://psiturk.org>



What can you do to help?

- **Use the platform to develop your own experiments**
 - We can use your use case as a template for others
- **Help us develop the platform**
 - We are missing something that you would like to add or improve?
- **Mention it to others**
 - You know anyone interested to use the framework?
- **Financial support?**
 - We can benefit from having freelance developers / designers
 - Cost of hosting the platform examples etc.



Questions/Discussion

