Conceptual structure is shaped by competing pressures for simplicity and informativeness

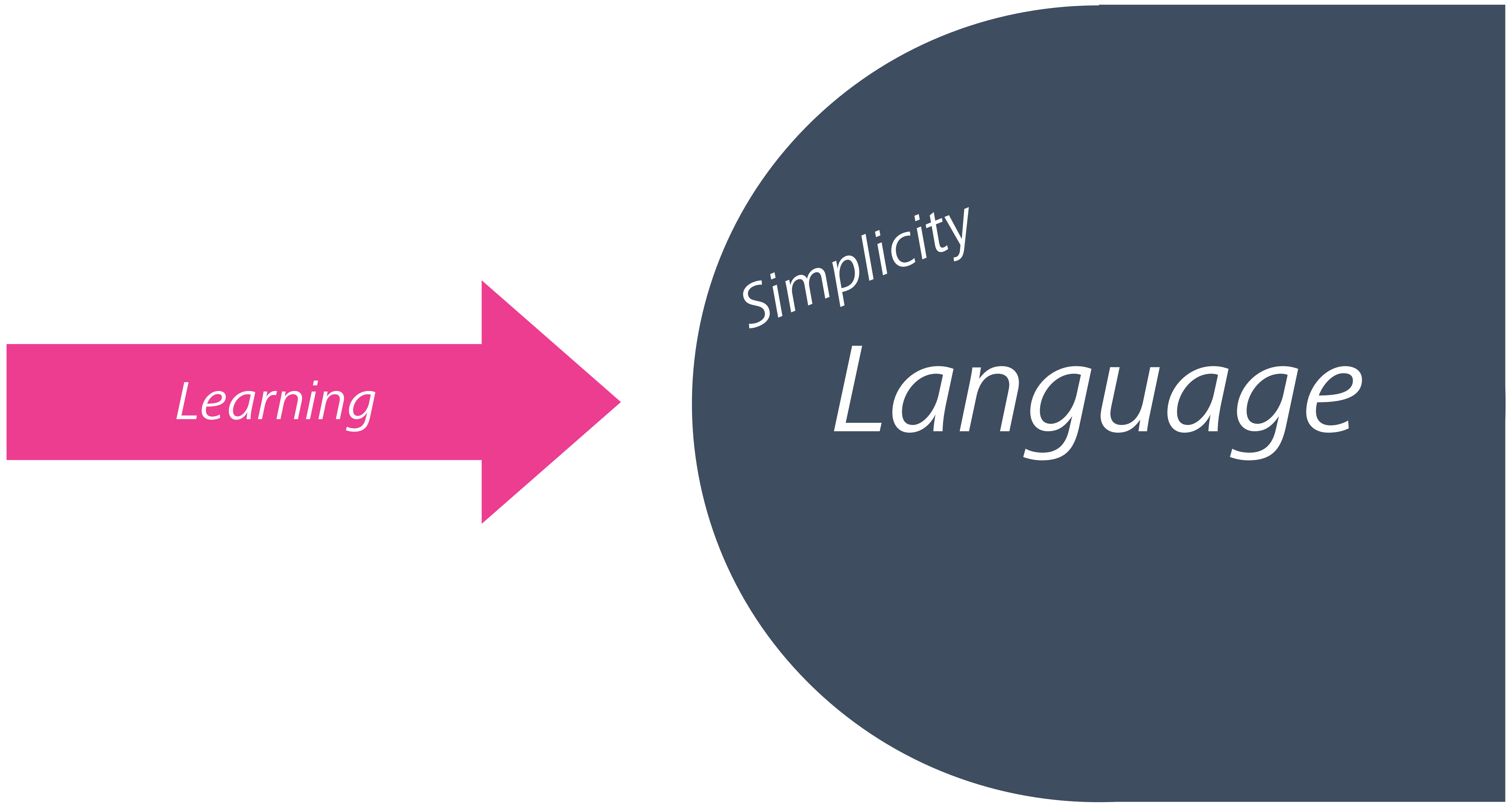
Jon W. Carr, Kenny Smith, Jennifer Culbertson, & Simon Kirby

Centre for Language Evolution
School of Philosophy, Psychology and Language Sciences
University of Edinburgh

Pressures shaping language

Language

Pressures shaping language



Pressures shaping language

Pressures shaping language

Kinship terms are simple and informative

Kinship Categories Across Languages Reflect General Communicative Principles

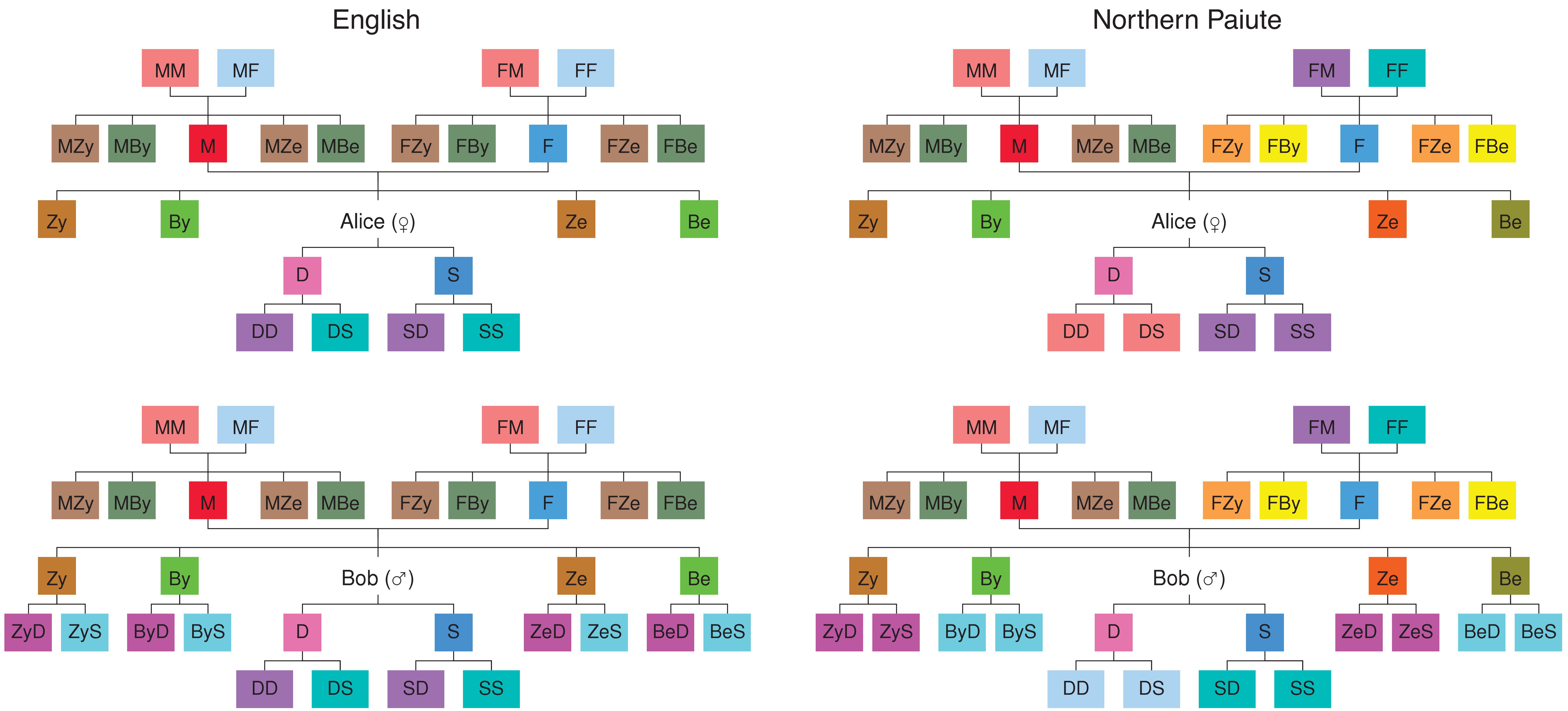
Charles Kemp[1]* and Terry Regier[2]

Languages vary in their systems of kinship categories, but the scope of possible variation appears to be constrained. Previous accounts of kin classification have often emphasized constraints that are specific to the domain of kinship and are not derived from general principles. Here, we propose an account that is founded on two domain-general principles: Good systems of categories are simple, and they enable informative communication. We show computationally that kin classification systems in the world's languages achieve a near-optimal trade-off between these two competing principles. We also show that our account explains several specific constraints on kin classification proposed previously. Because the principles of simplicity and informativeness are also relevant to other semantic domains, the trade-off between them may provide a domain-general foundation for variation in category systems across languages.

Concepts and categories vary across cultures but may nevertheless be shaped by universal constraints (*1–4*). Cross-cultural studies have proposed universal constraints that help to explain how colors (*5*, *6*), plants, animals (*7*, *8*), and spatial relations (*9*, *10*) are organized into categories. Kinship has traditionally been a prominent domain for studies of this kind, and researchers have described many constraints that help to predict which of the many logically possible kin classification systems are encountered in practice (*11–15*). Typically these constraints are not derived from general principles, although it is often suggested that they are consistent with cognitive and functional considerations (*2*, *11–13*, *15*). Here, we show that major aspects of kin classification follow directly from two general principles: Categories tend to be simple, which minimizes cognitive load, and to be informative, which maximizes communicative efficiency. Principles like these have been discussed in other contexts by previous researchers (*16–19*). For example, Zipf suggested that word-frequency distributions achieve a trade-off between simplicity and communicative precision (*20*, *21*), Hawkins (*22*) has suggested that grammars are shaped by a trade-off between simplicity and communicative efficiency, and Rosch has suggested that category systems "provide maximum information with the least cognitive effort" [p. 190 of (*23*)].

Figure 1A shows a simple communication game that helps to illustrate how kin classification systems are shaped by the principles of simplicity and informativeness. The speaker has a specific relative in mind and utters the category label for that relative. Upon hearing this category label, the hearer must guess which relative the speaker had in mind.

Kinship terms are simple and informative



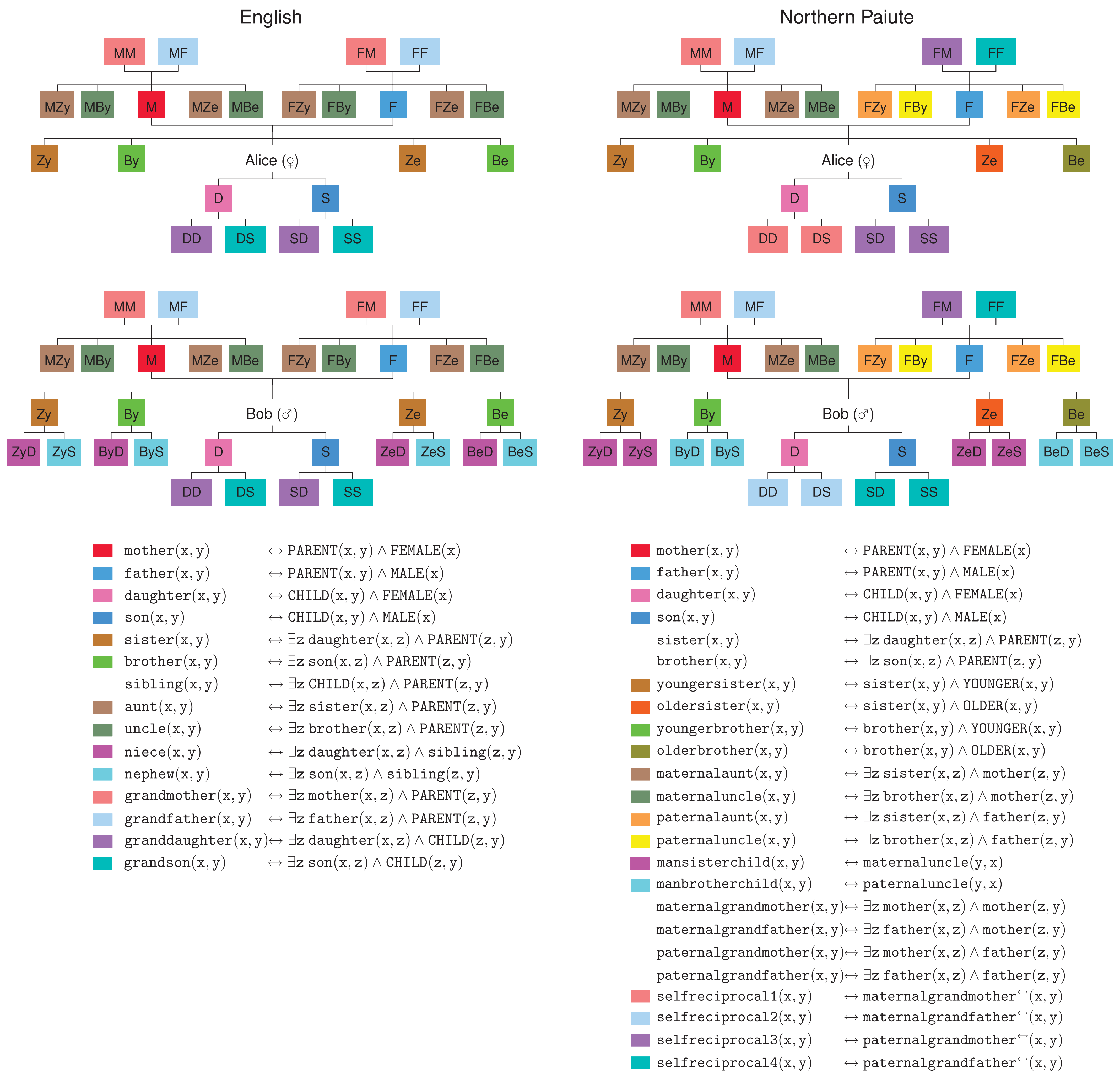
English:

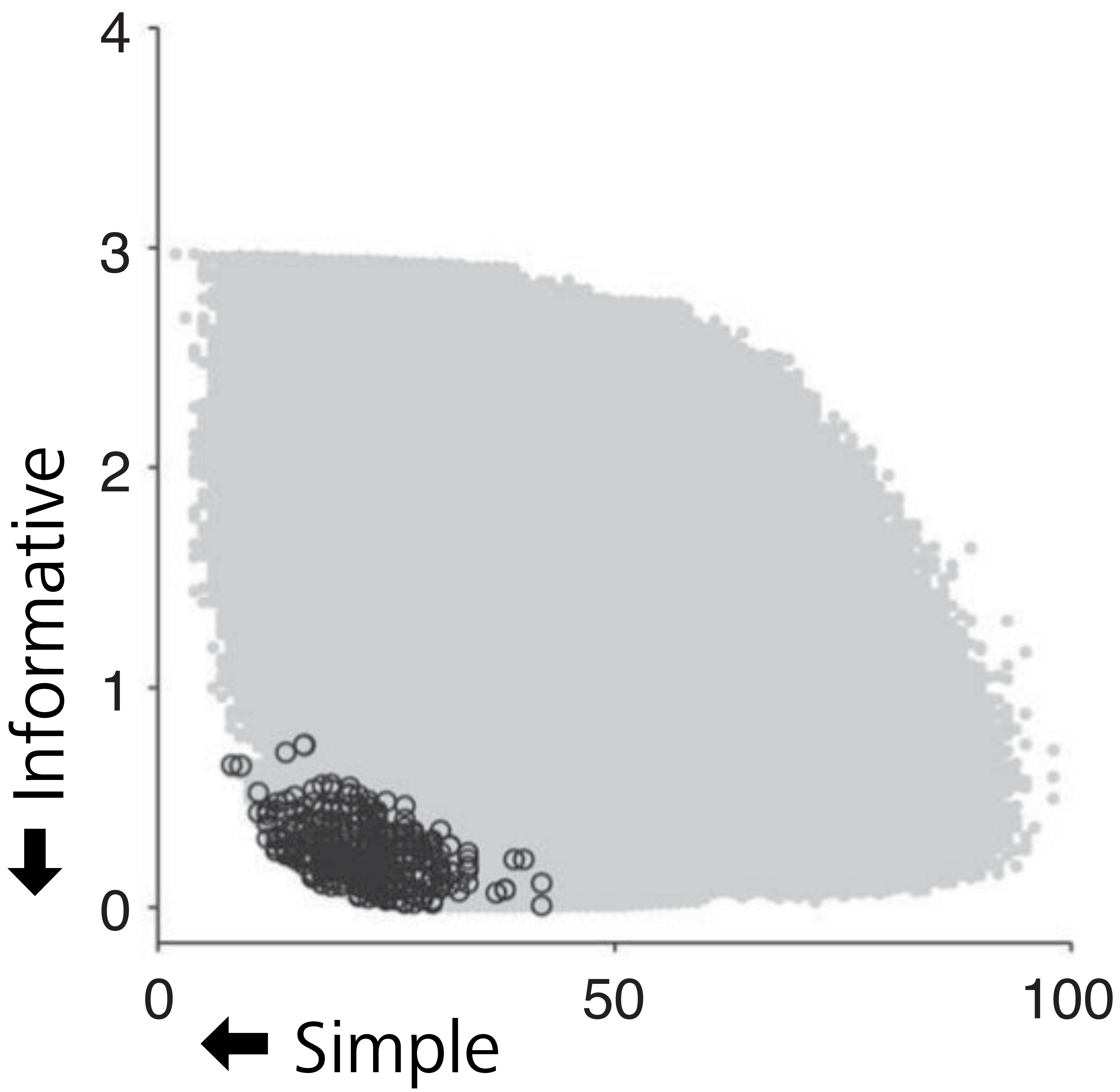
mother(x, y)	$\leftrightarrow$ PARENT(x, y) $\wedge$ FEMALE(x)
father(x, y)	$\leftrightarrow$ PARENT(x, y) $\wedge$ MALE(x)
daughter(x, y)	$\leftrightarrow$ CHILD(x, y) $\wedge$ FEMALE(x)
son(x, y)	$\leftrightarrow$ CHILD(x, y) $\wedge$ MALE(x)
sister(x, y)	$\leftrightarrow$ $\exists$z daughter(x, z) $\wedge$ PARENT(z, y)

Northern Paiute:

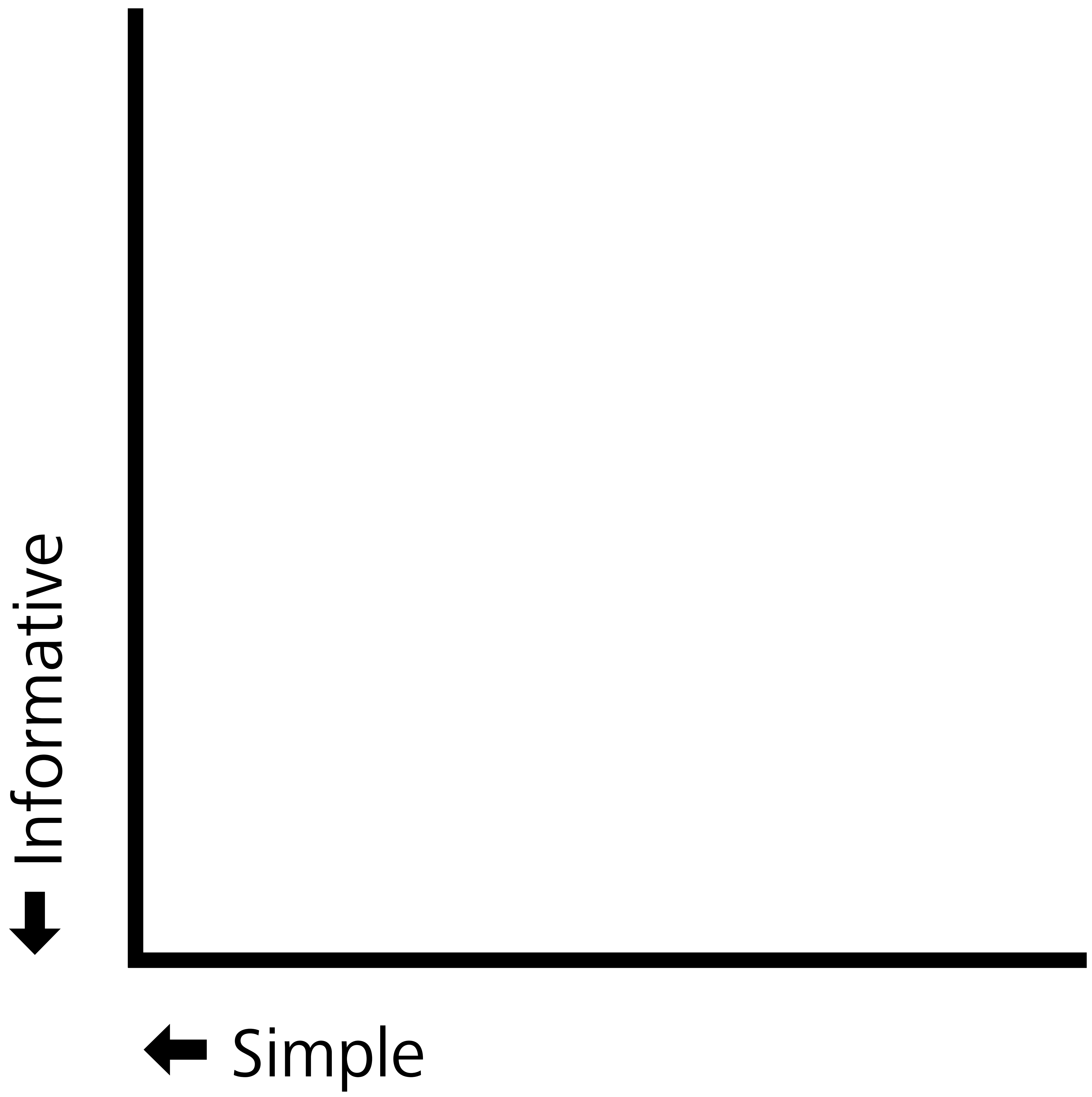
mother(x, y)	$\leftrightarrow$ PARENT(x, y) $\wedge$ FEMALE(x)
father(x, y)	$\leftrightarrow$ PARENT(x, y) $\wedge$ MALE(x)
daughter(x, y)	$\leftrightarrow$ CHILD(x, y) $\wedge$ FEMALE(x)
son(x, y)	$\leftrightarrow$ CHILD(x, y) $\wedge$ MALE(x)
sister(x, y)	$\leftrightarrow$ $\exists$z daughter(x, z) $\wedge$ PARENT(z, y)

Kinship terms are simple and informative

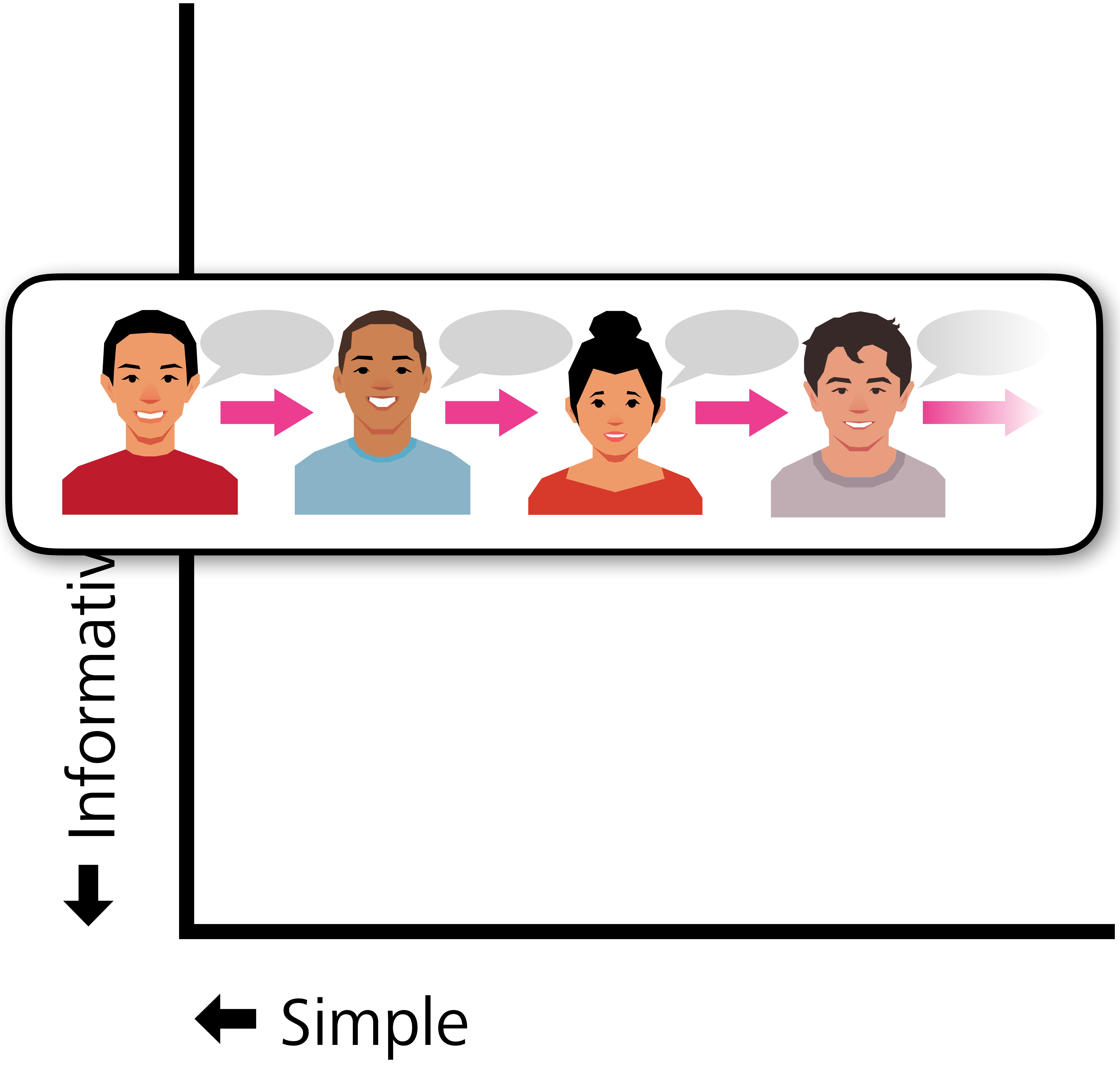




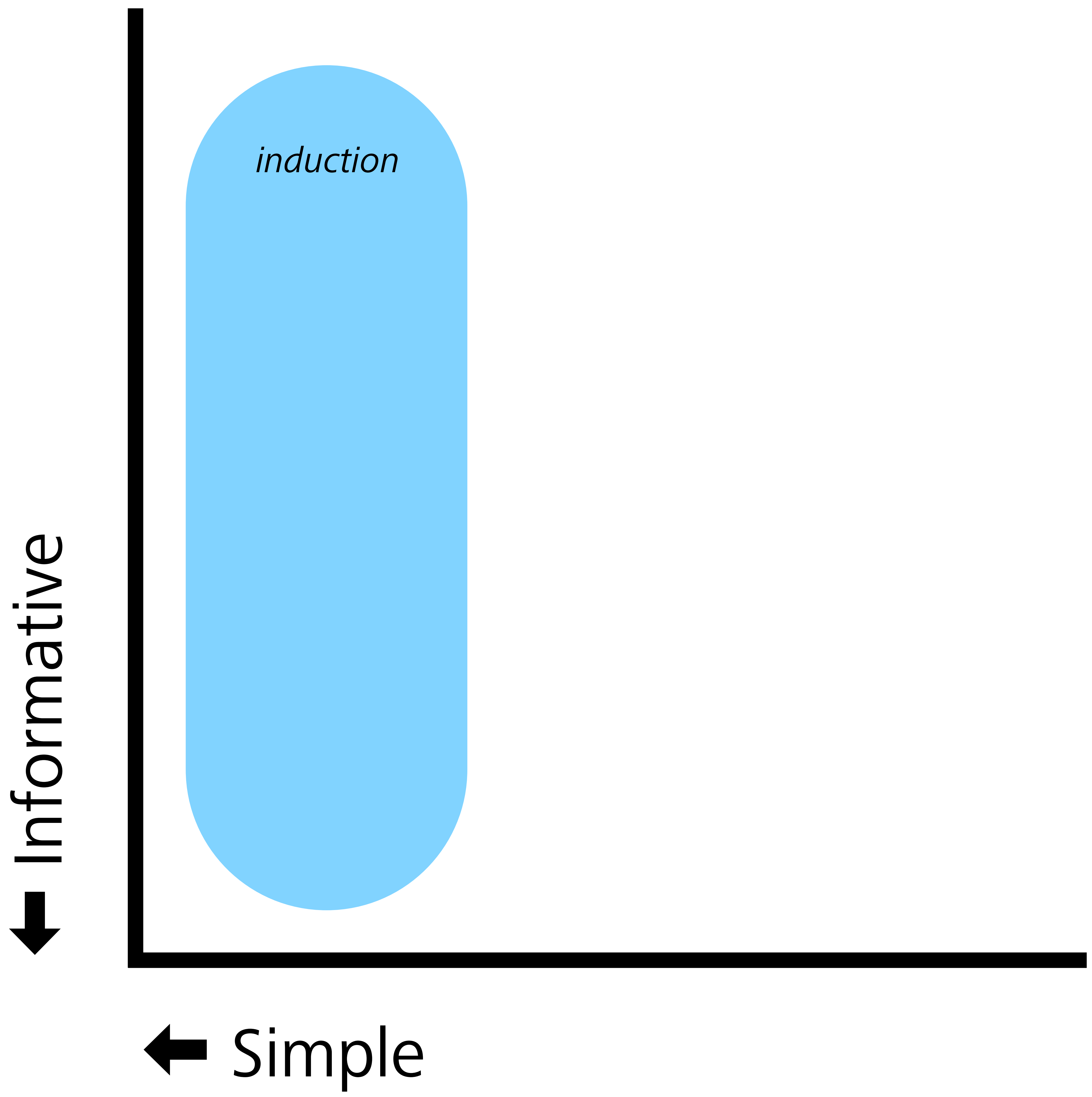
Induction & interaction: Pressures for simplicity and informativeness



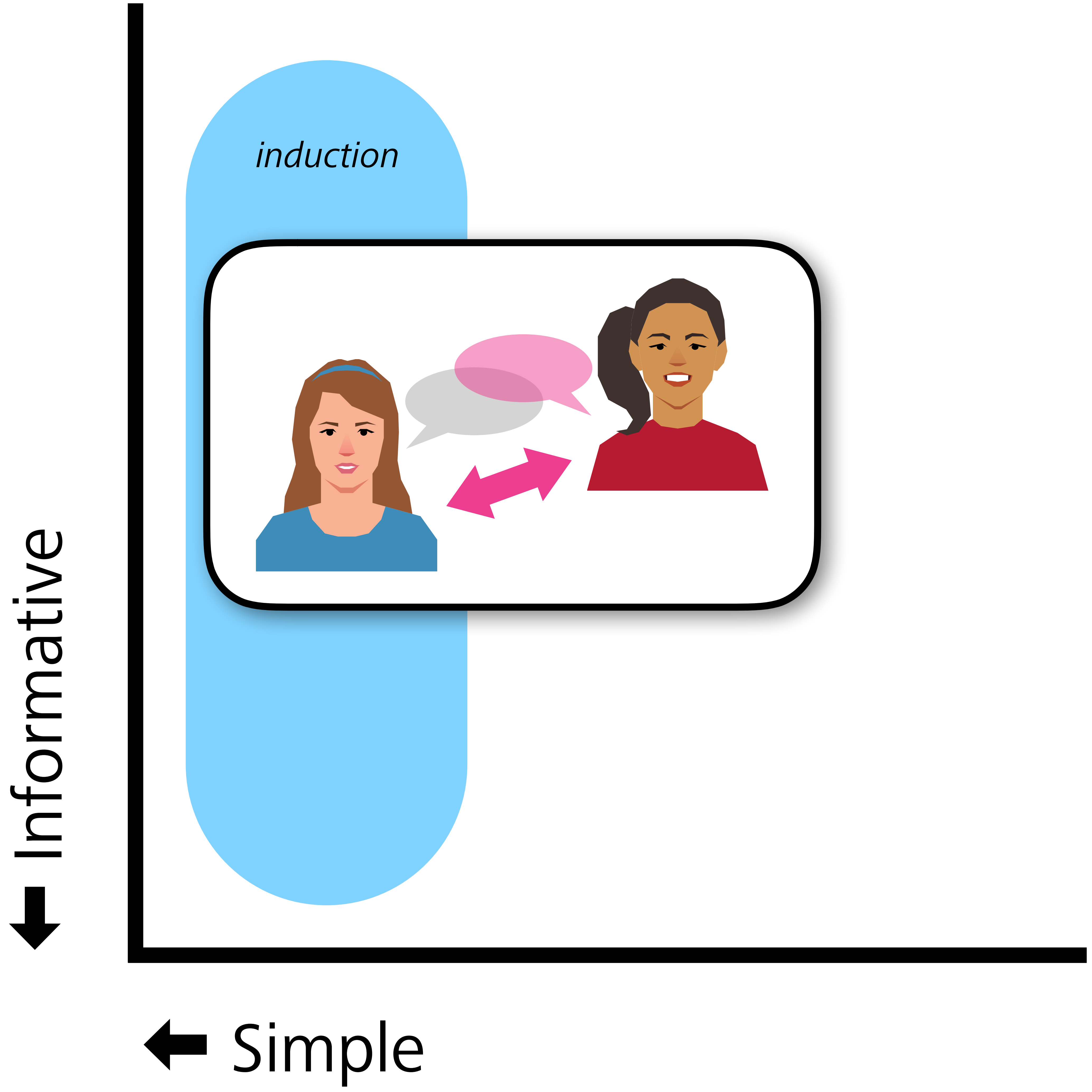
Induction & interaction: Pressures for simplicity and informativeness



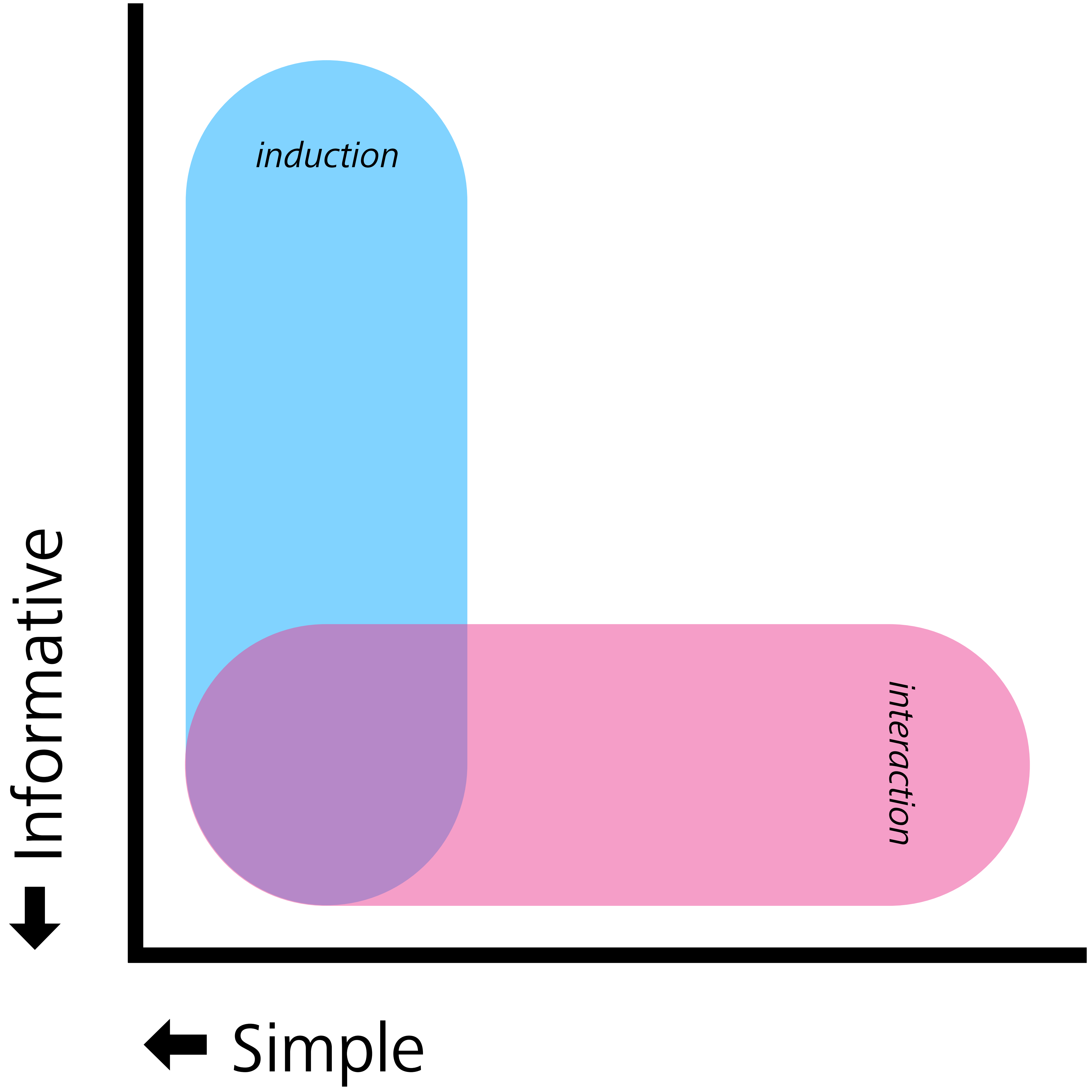
Induction & interaction: Pressures for simplicity and informativeness



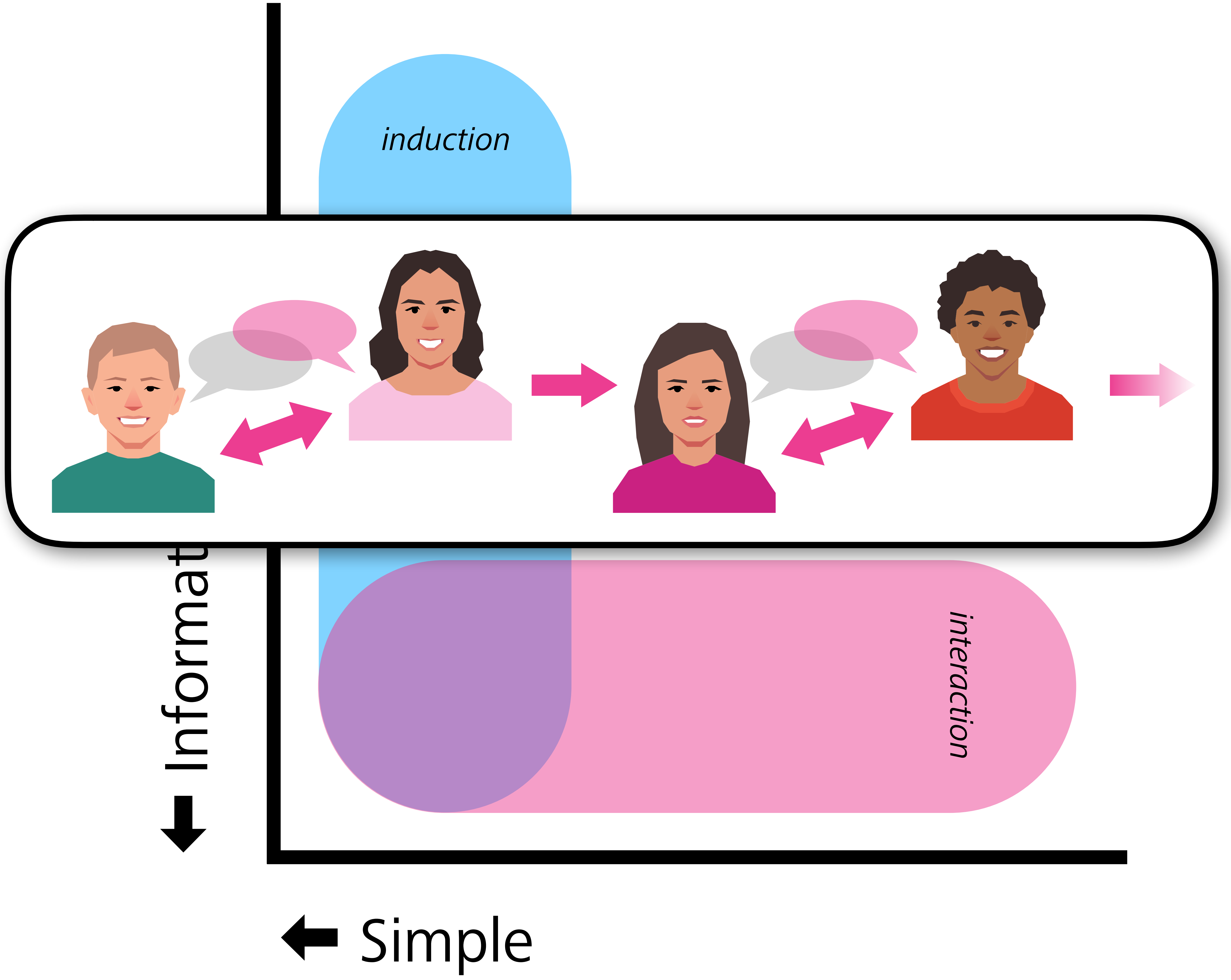
Induction & interaction: Pressures for simplicity and informativeness



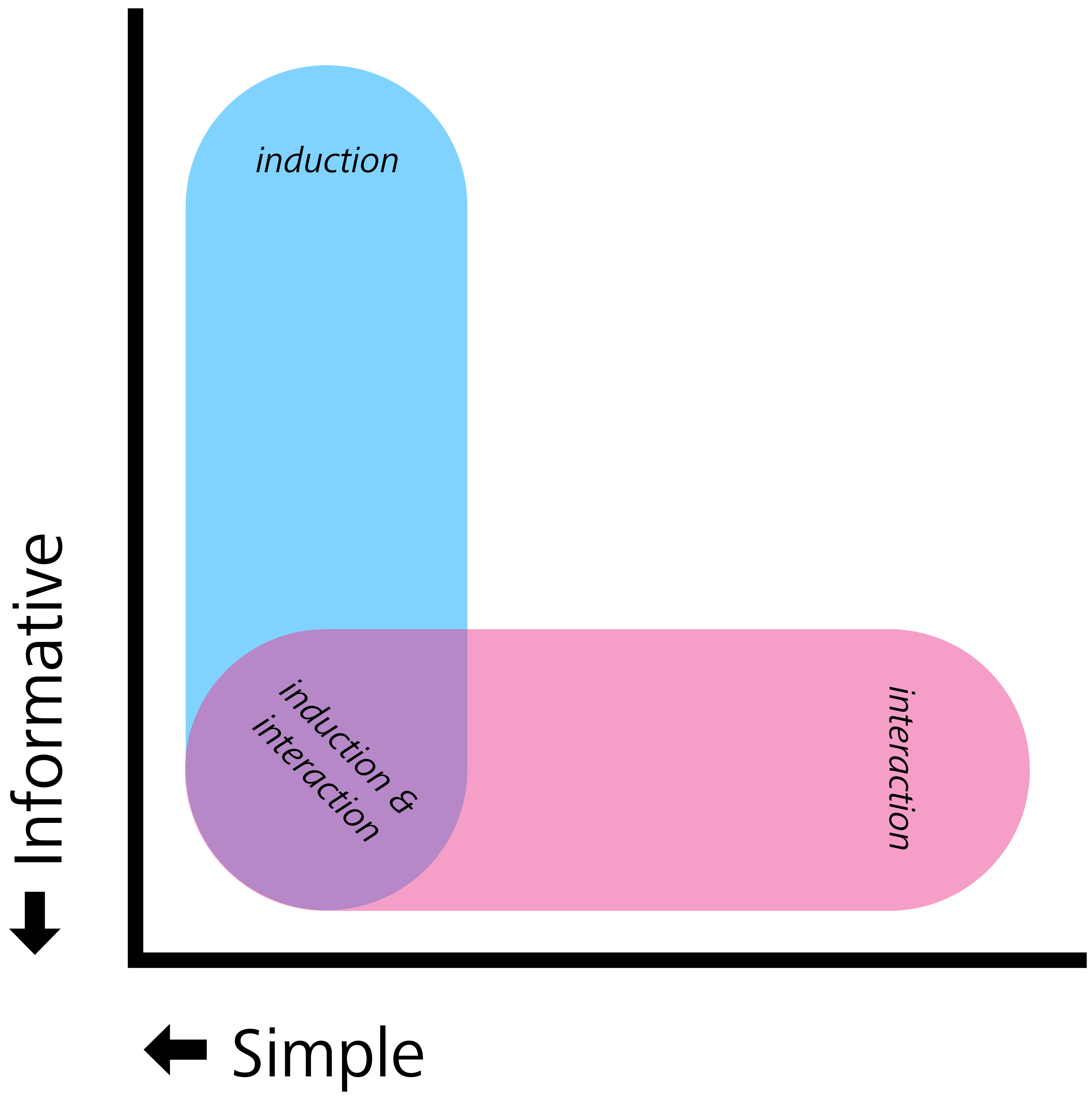
Induction & interaction: Pressures for simplicity and informativeness



Induction & interaction: Pressures for simplicity and informativeness



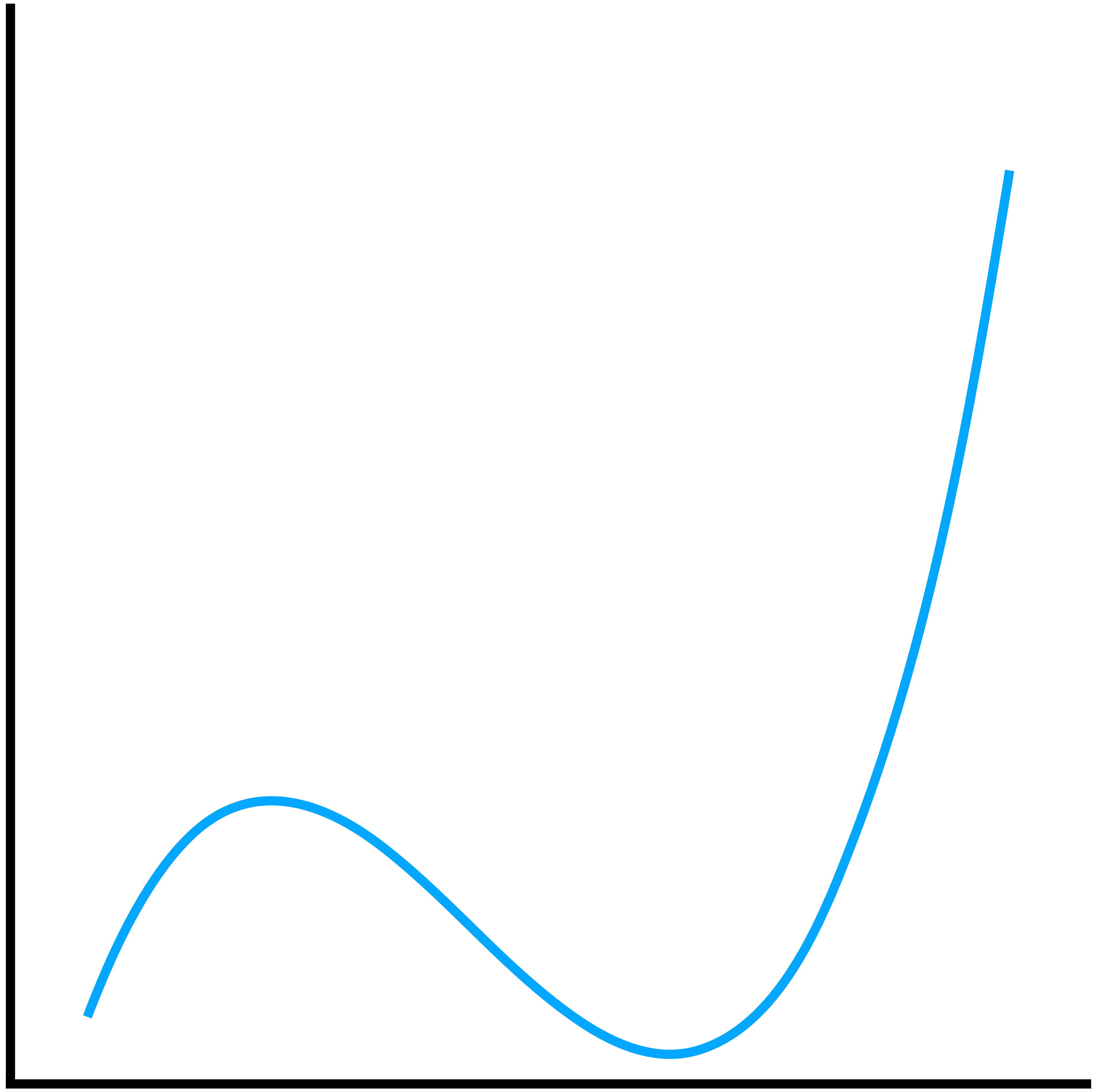
Induction & interaction: Pressures for simplicity and informativeness



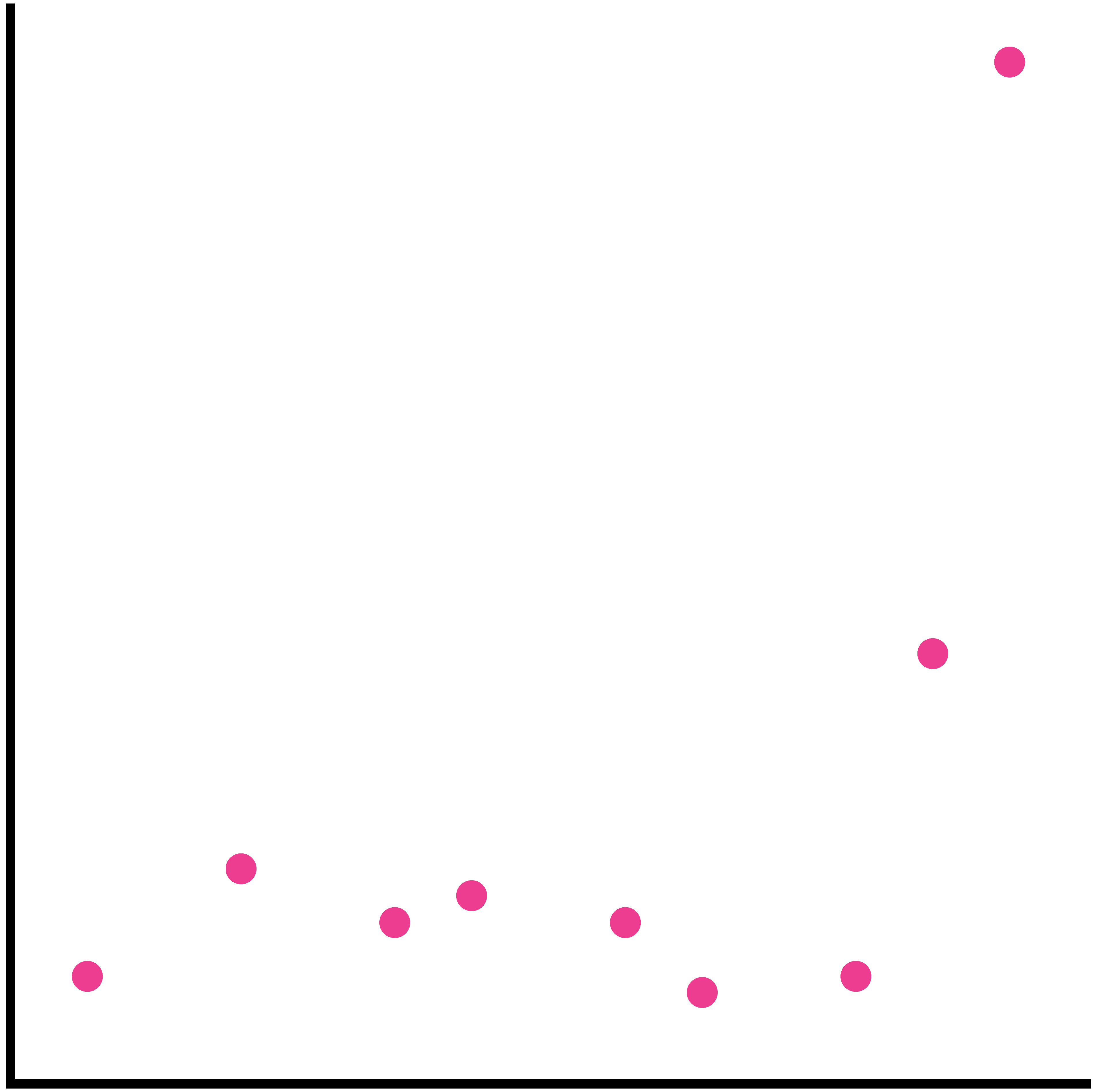
Induction

as the pressure for simplicity

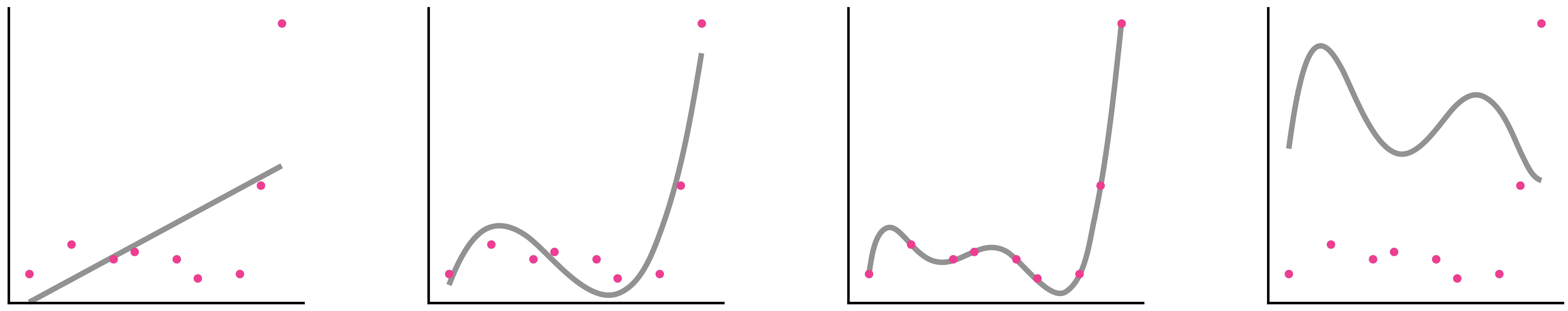
Induction and the simplicity prior



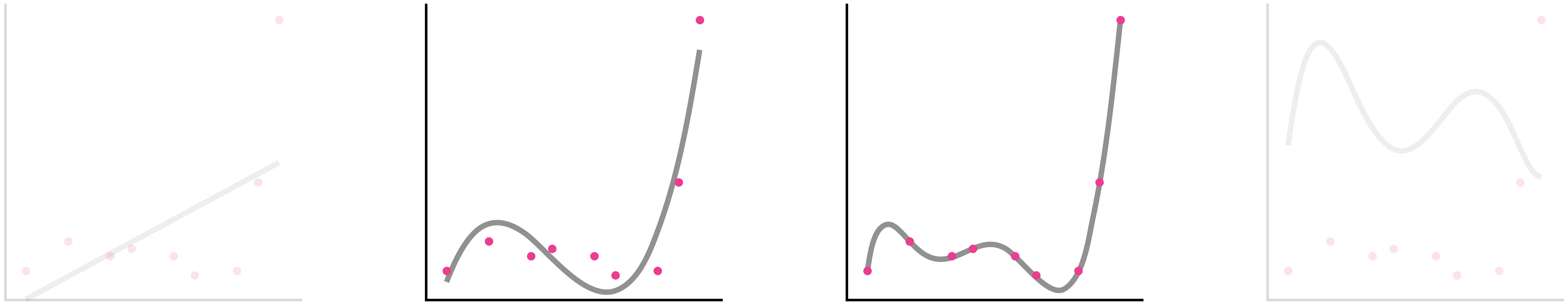
Induction and the simplicity prior



Induction and the simplicity prior

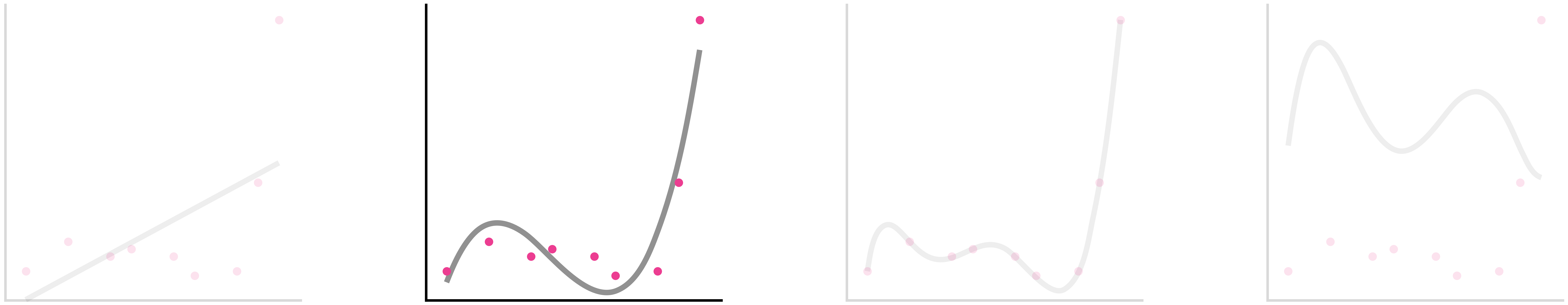


Induction and the simplicity prior



Principle of multiple explanations

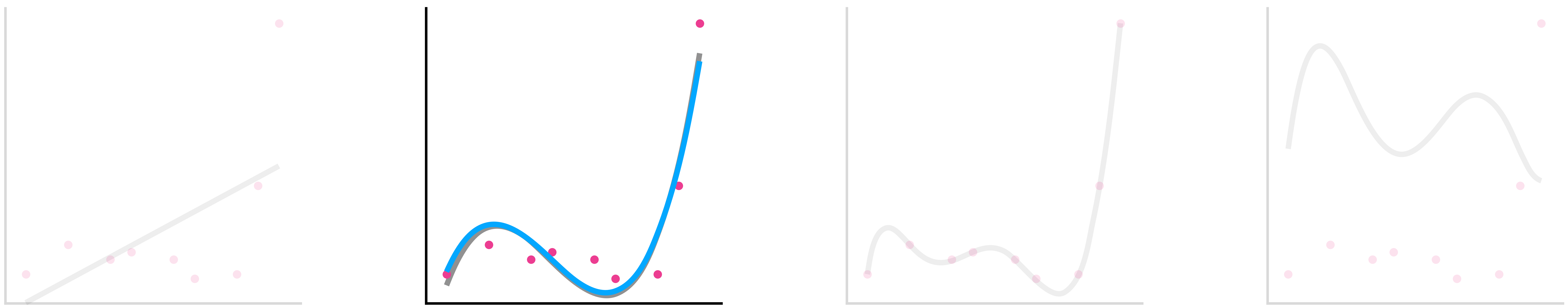
Induction and the simplicity prior



Occam's razor

Principle of multiple explanations

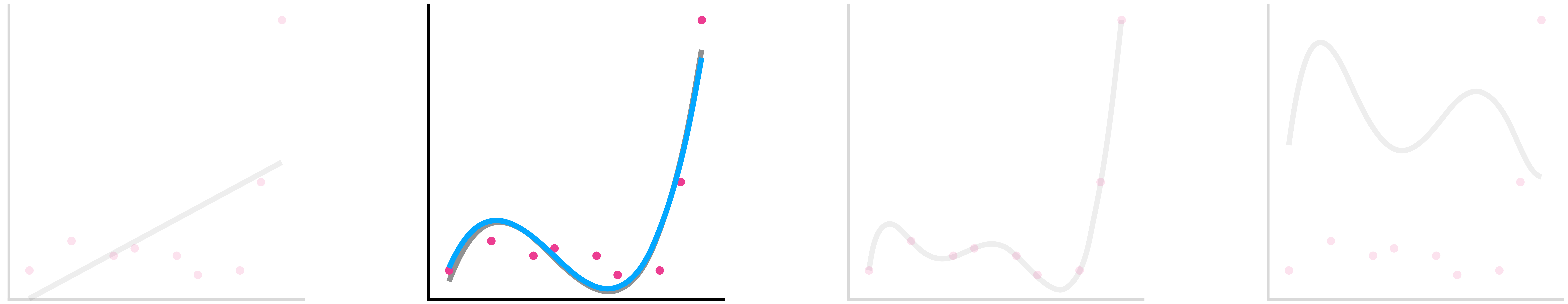
Induction and the simplicity prior



Occam’s razor

Principle of multiple explanations

Induction and the simplicity prior



Occam's razor

Principle of multiple explanations

$$P(H|D) \propto P(H)P(D|H)$$

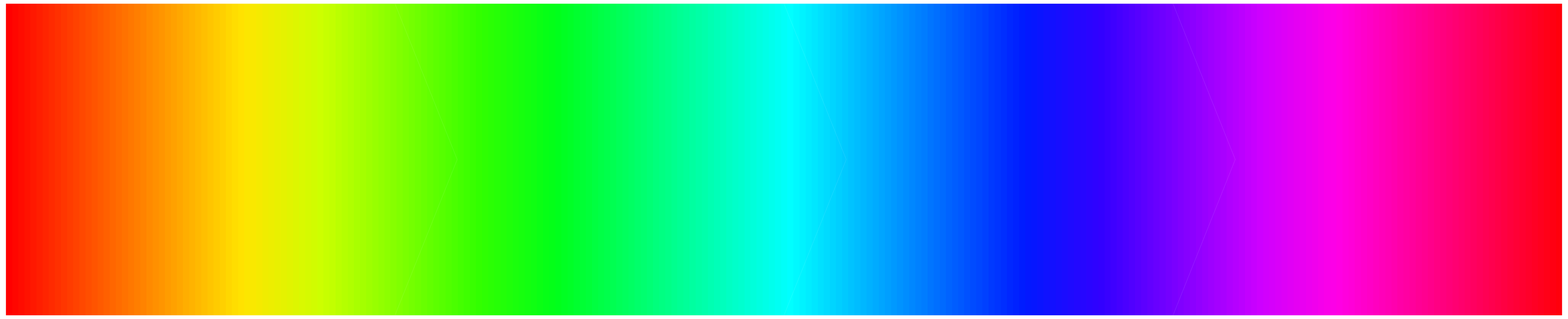
Interaction

as the pressure for informativeness

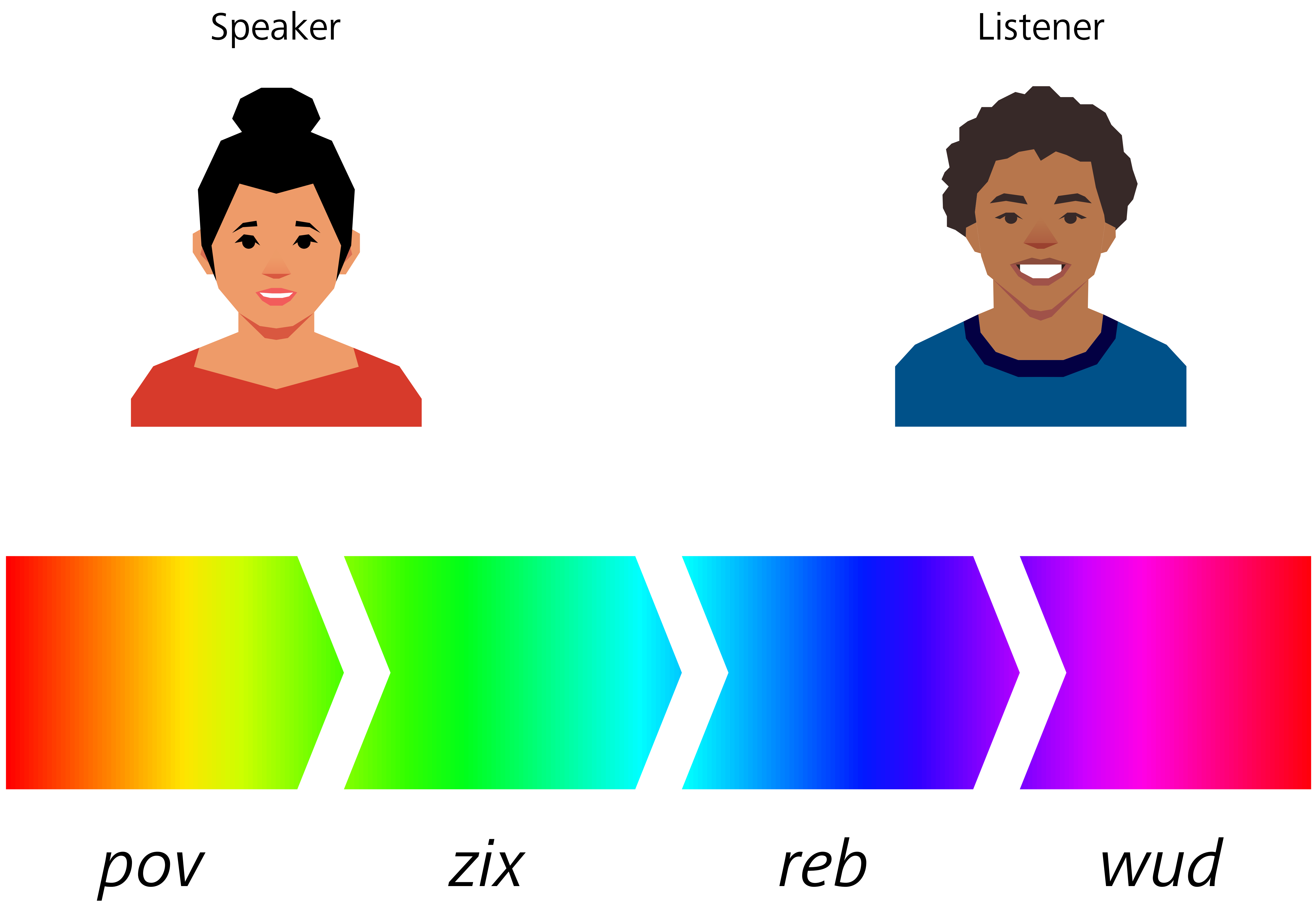
Regier et al.'s informativeness model

Speaker

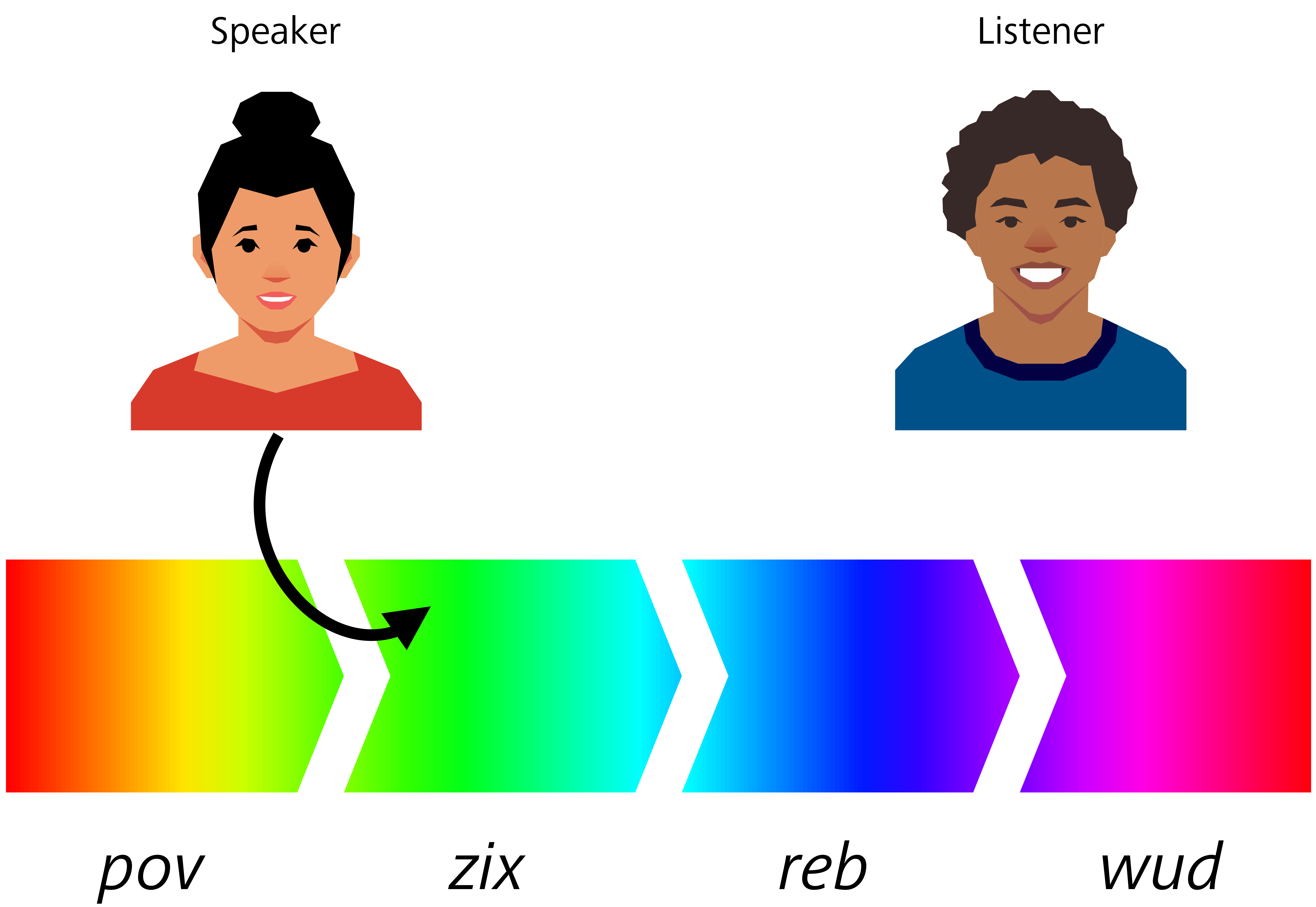
Listener



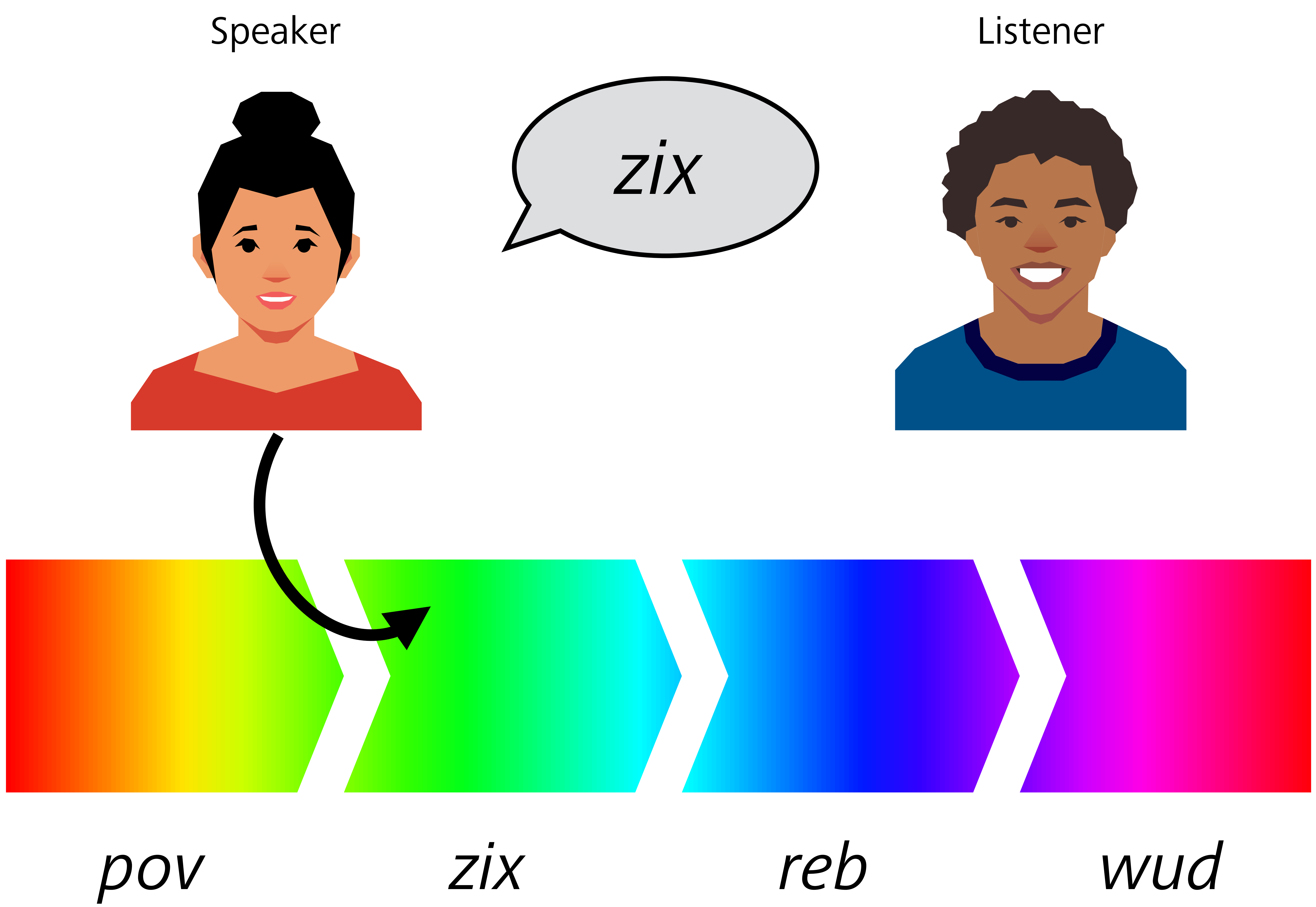
Regier et al.'s informativeness model



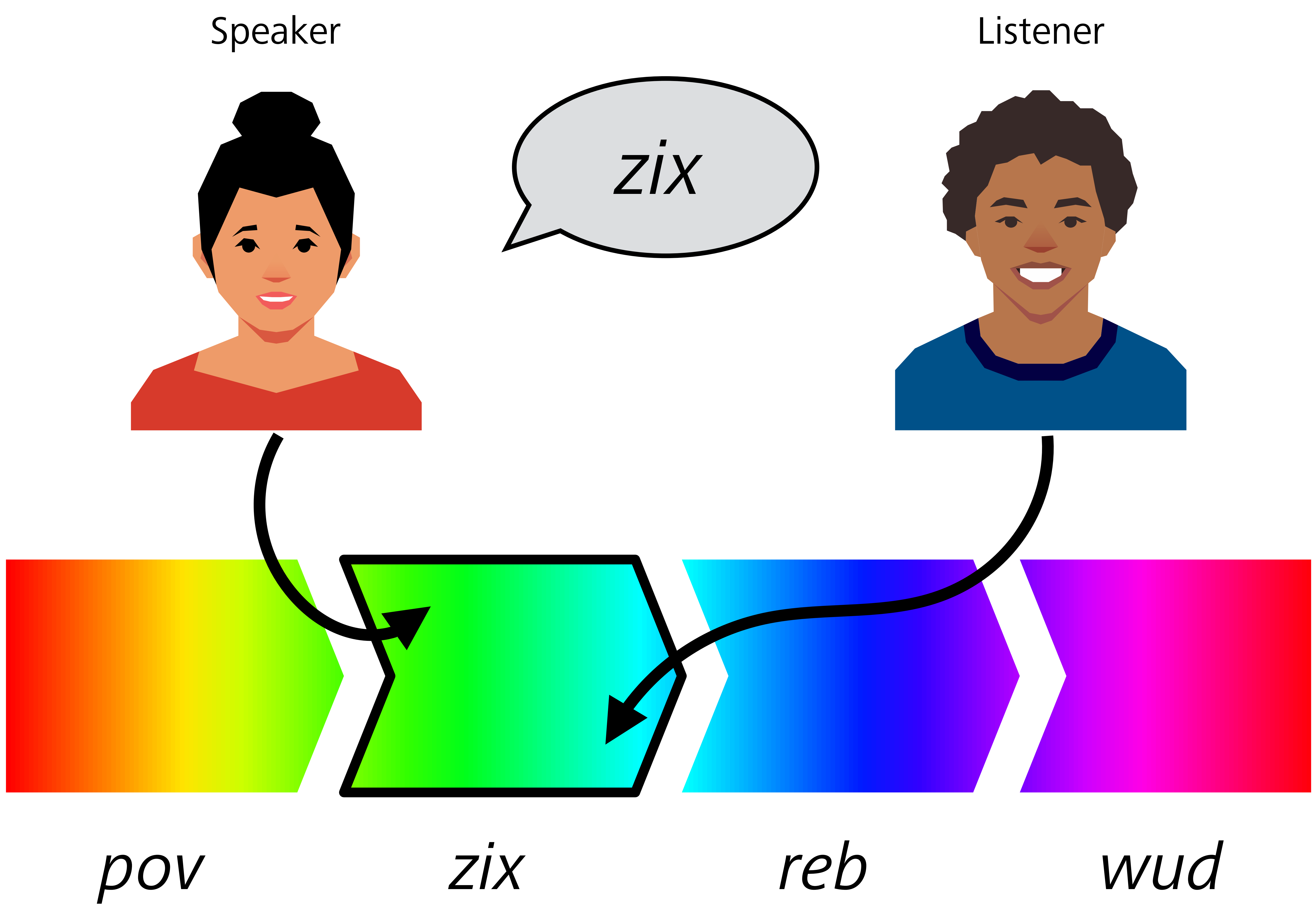
Regier et al.'s informativeness model



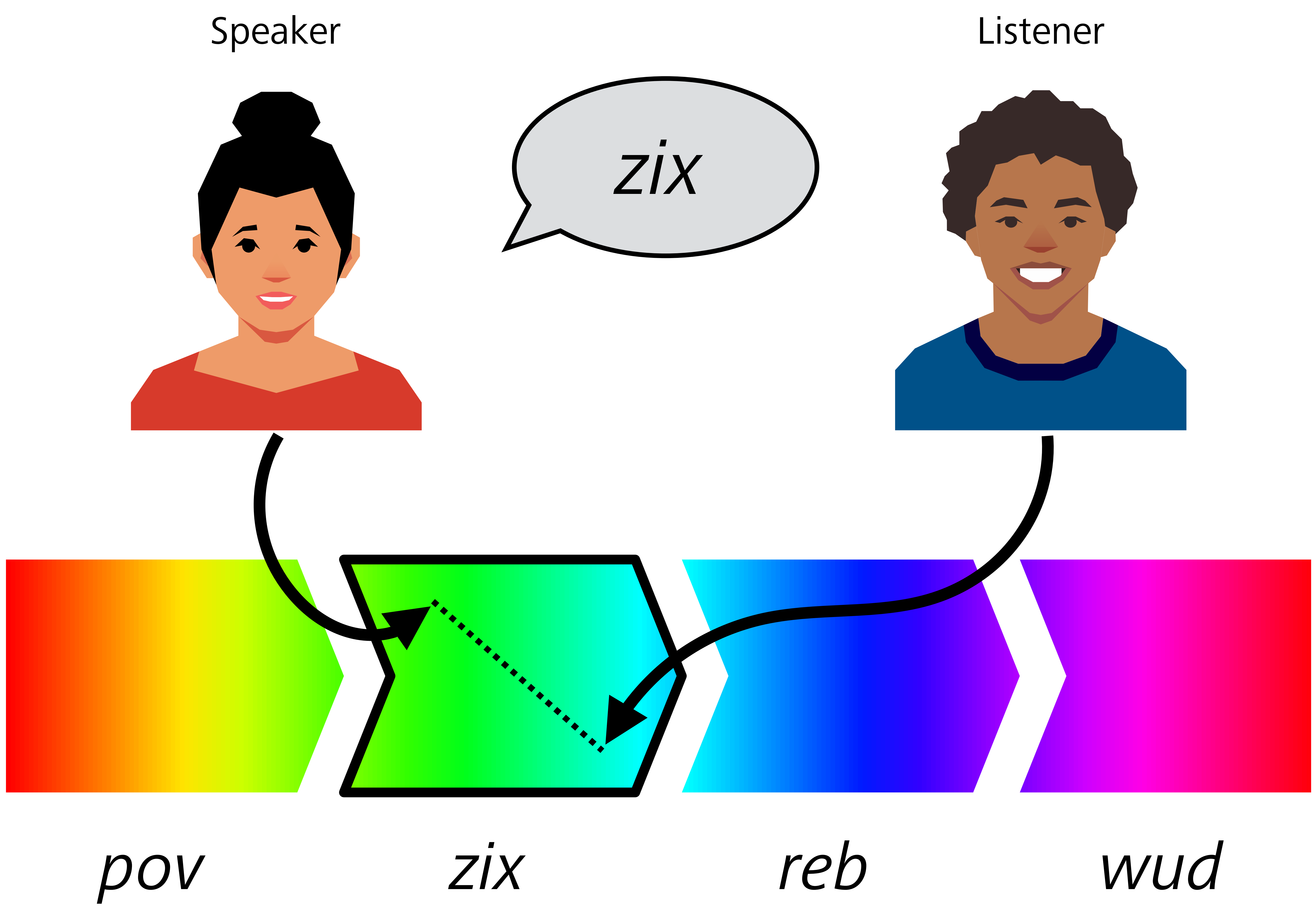
Regier et al.'s informativeness model



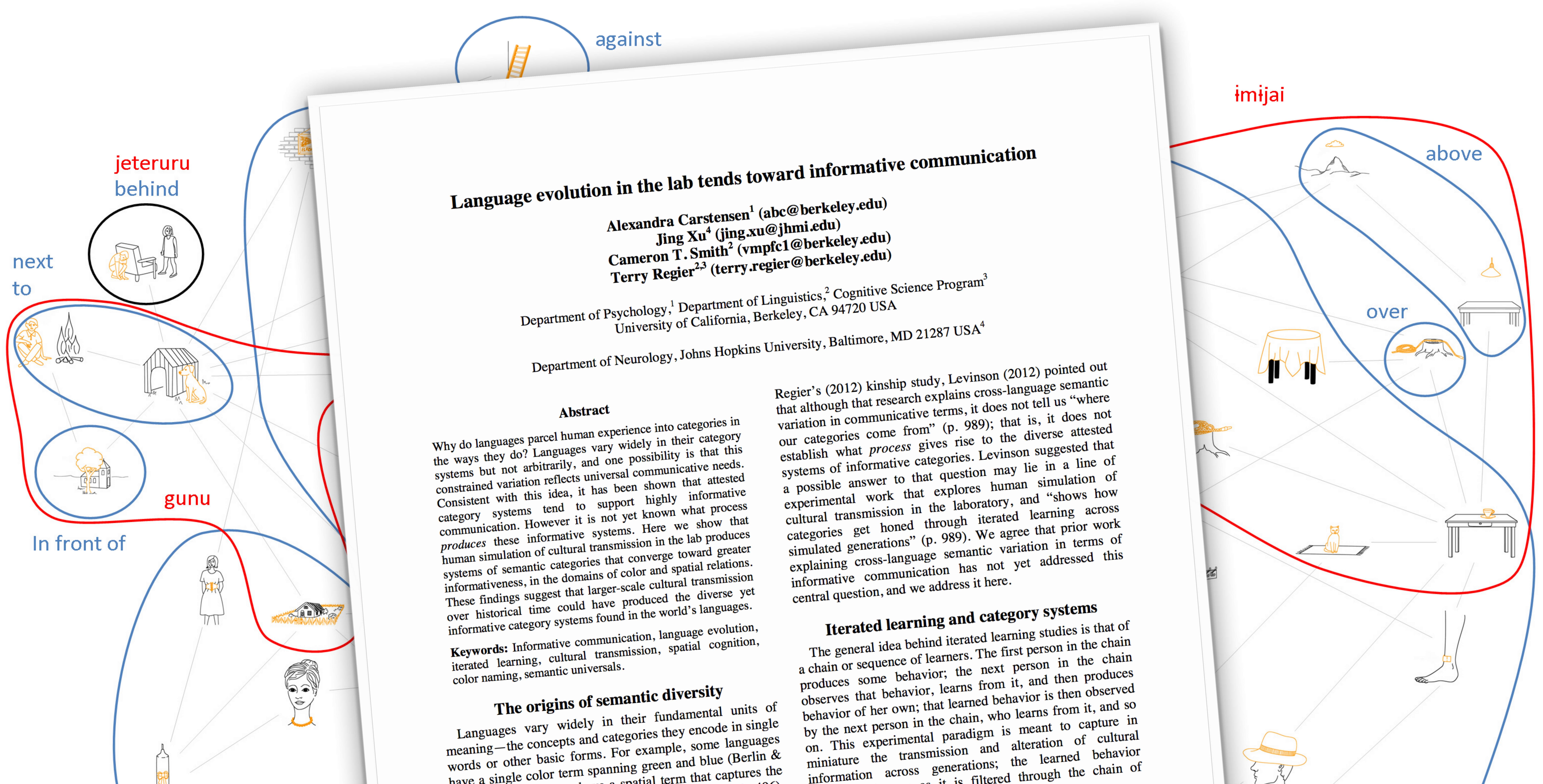
Regier et al.'s informativeness model



Regier et al.'s informativeness model



Can iterated learning give rise to informative categories?



Language evolution in the lab tends toward informative communication

Alexandra Carstensen[1] (abc@berkeley.edu)
Jing Xu[4] (jing.xu@jhmi.edu)
Cameron T. Smith[2] (vmpfc1@berkeley.edu)
Terry Regier[2,3] (terry.regier@berkeley.edu)

Department of Psychology,[1] Department of Linguistics,[2] Cognitive Science Program[3]
University of California, Berkeley, CA 94720 USA

Department of Neurology, Johns Hopkins University, Baltimore, MD 21287 USA[4]

Abstract

Why do languages parcel human experience into categories in the ways they do? Languages vary widely in their category systems but not arbitrarily, and one possibility is that this constrained variation reflects universal communicative needs. Consistent with this idea, it has been shown that attested category systems tend to support highly informative communication. However it is not yet known what process *produces* these informative systems. Here we show that human simulation of cultural transmission in the lab produces systems of semantic categories that converge toward greater informativeness, in the domains of color and spatial relations. These findings suggest that larger-scale cultural transmission over historical time could have produced the diverse yet informative category systems found in the world's languages.

Keywords: Informative communication, language evolution, iterated learning, cultural transmission, spatial cognition, color naming, semantic universals.

The origins of semantic diversity

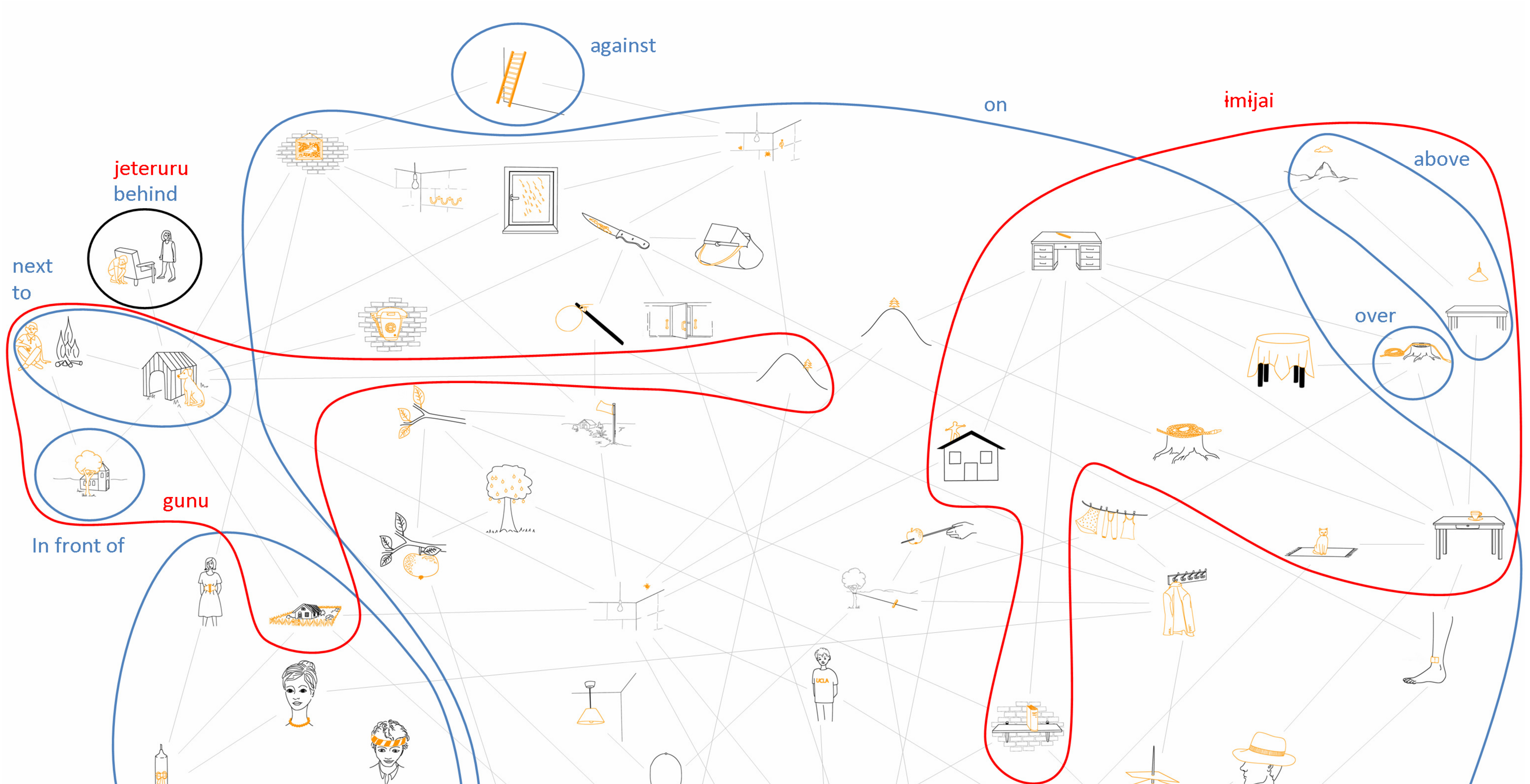
Languages vary widely in their fundamental units of meaning—the concepts and categories they encode in single words or other basic forms. For example, some languages have a single color term spanning green and blue (Berlin &

Regier's (2012) kinship study, Levinson (2012) pointed out that although that research explains cross-language semantic variation in communicative terms, it does not tell us "where our categories come from" (p. 989); that is, it does not establish what *process* gives rise to the diverse attested systems of informative categories. Levinson suggested that a possible answer to that question may lie in a line of experimental work that explores human simulation of cultural transmission in the laboratory, and "shows how categories get honed through iterated learning across simulated generations" (p. 989). We agree that prior work explaining cross-language semantic variation in terms of informative communication has not yet addressed this central question, and we address it here.

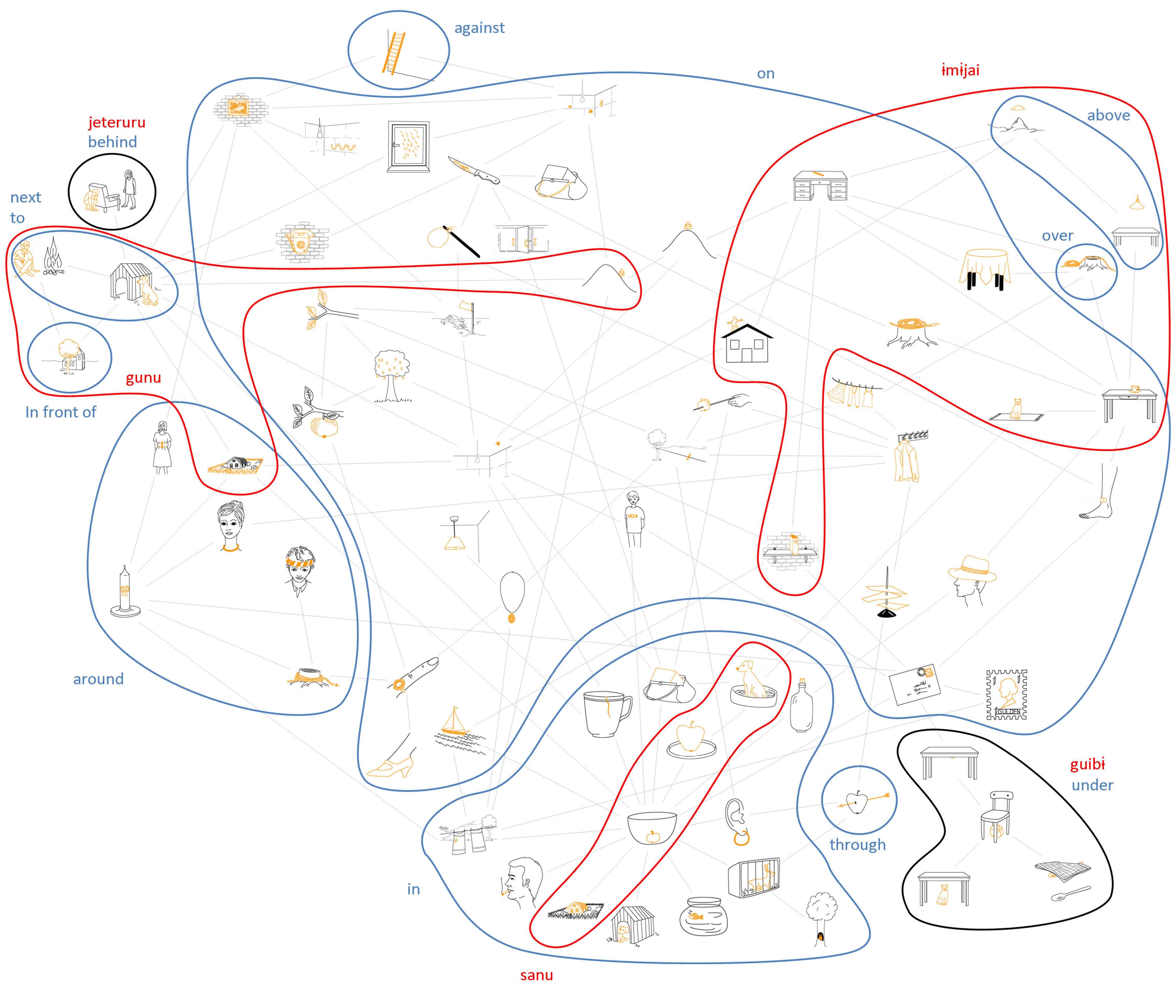
Iterated learning and category systems

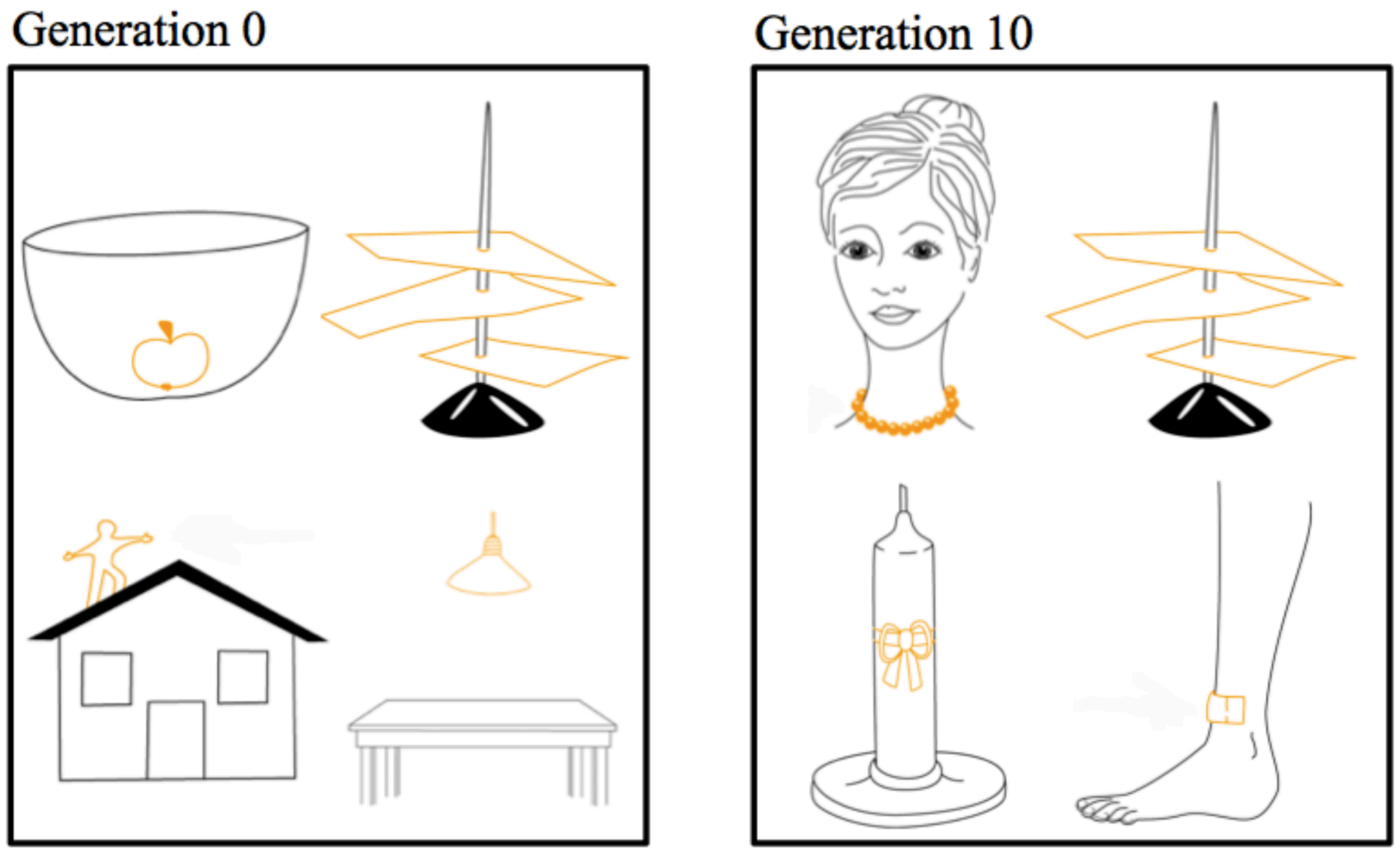
The general idea behind iterated learning studies is that of a chain or sequence of learners. The first person in the chain produces some behavior; the next person in the chain observes that behavior, learns from it, and then produces behavior of her own; that learned behavior is then observed by the next person in the chain, who learns from it, and so on. This experimental paradigm is meant to capture in miniature the transmission and alteration of cultural information across generations; the learned behavior

Can iterated learning give rise to informative categories?



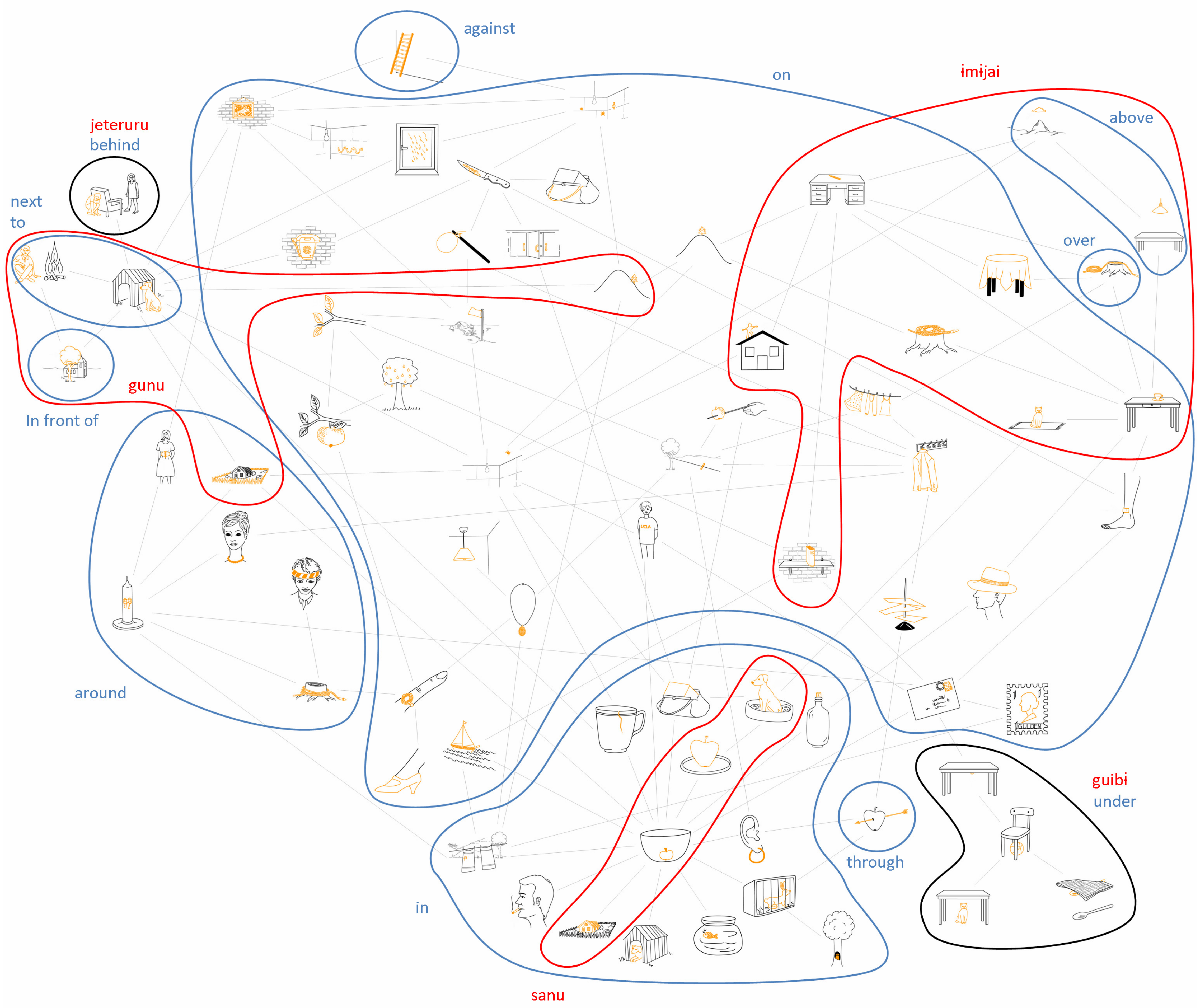
Can iterated learning give rise to informative categories?

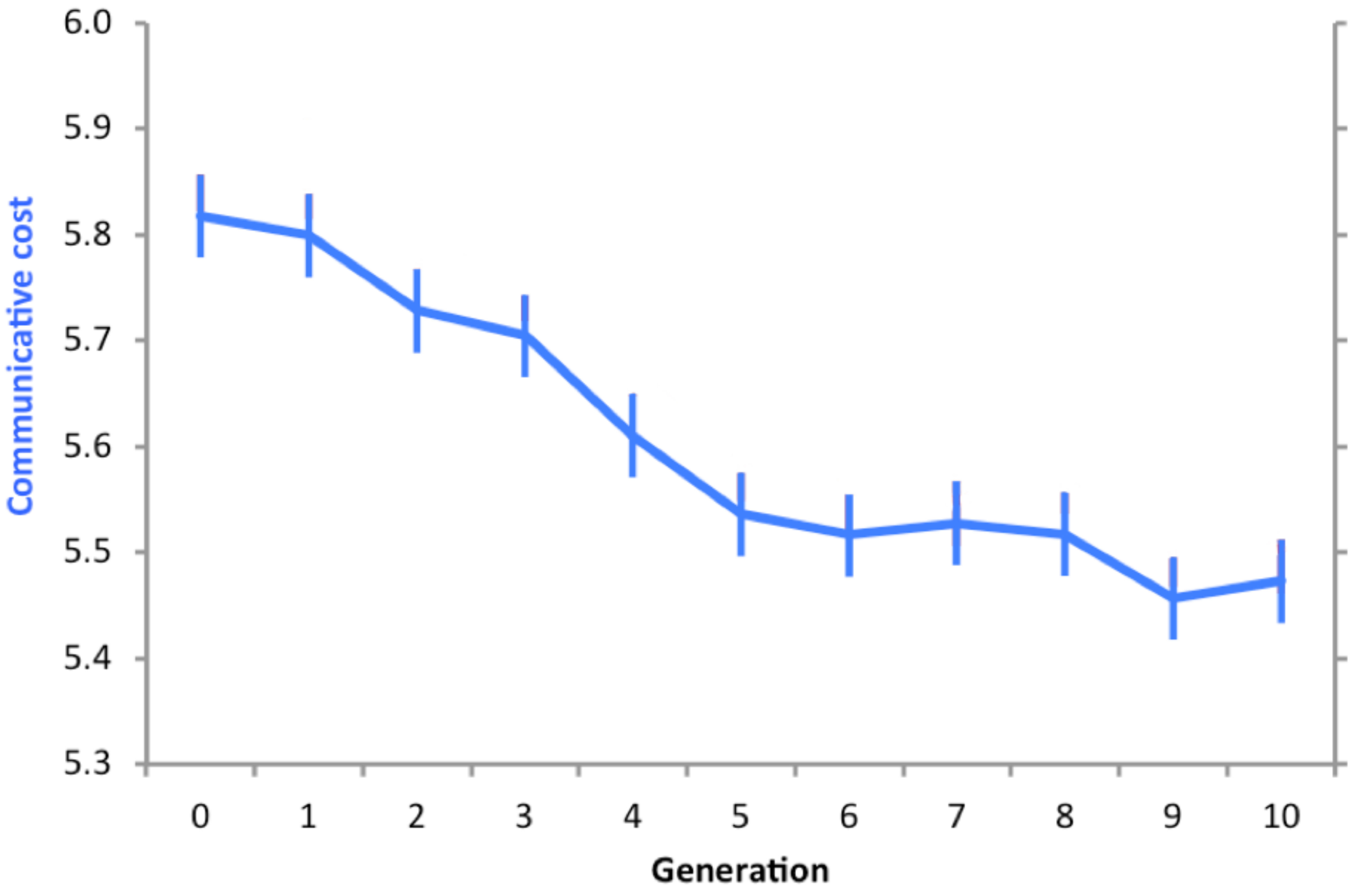




Carstensen, Xu, Smith, Regier (2015)

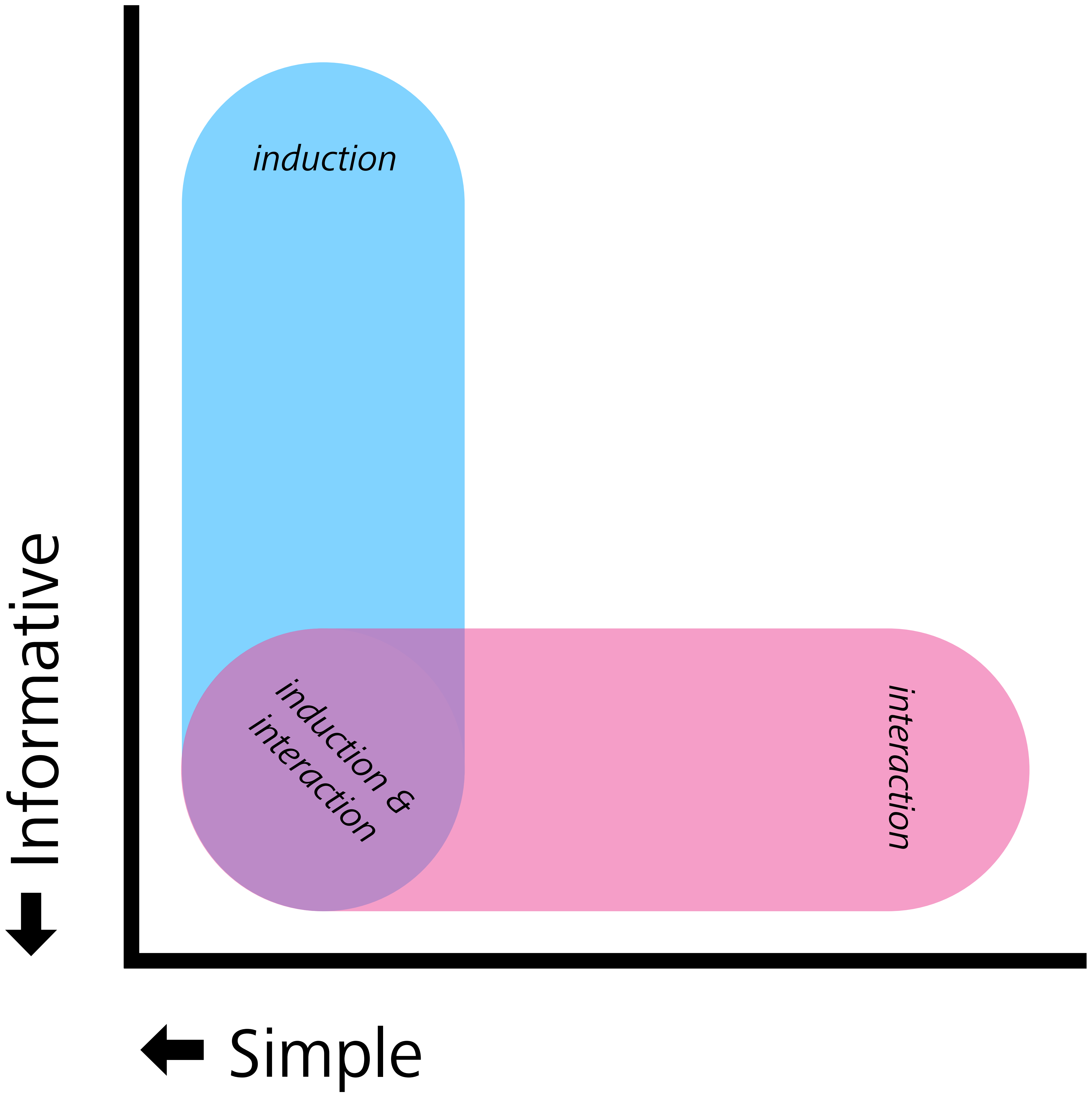
Can iterated learning give rise to informative categories?



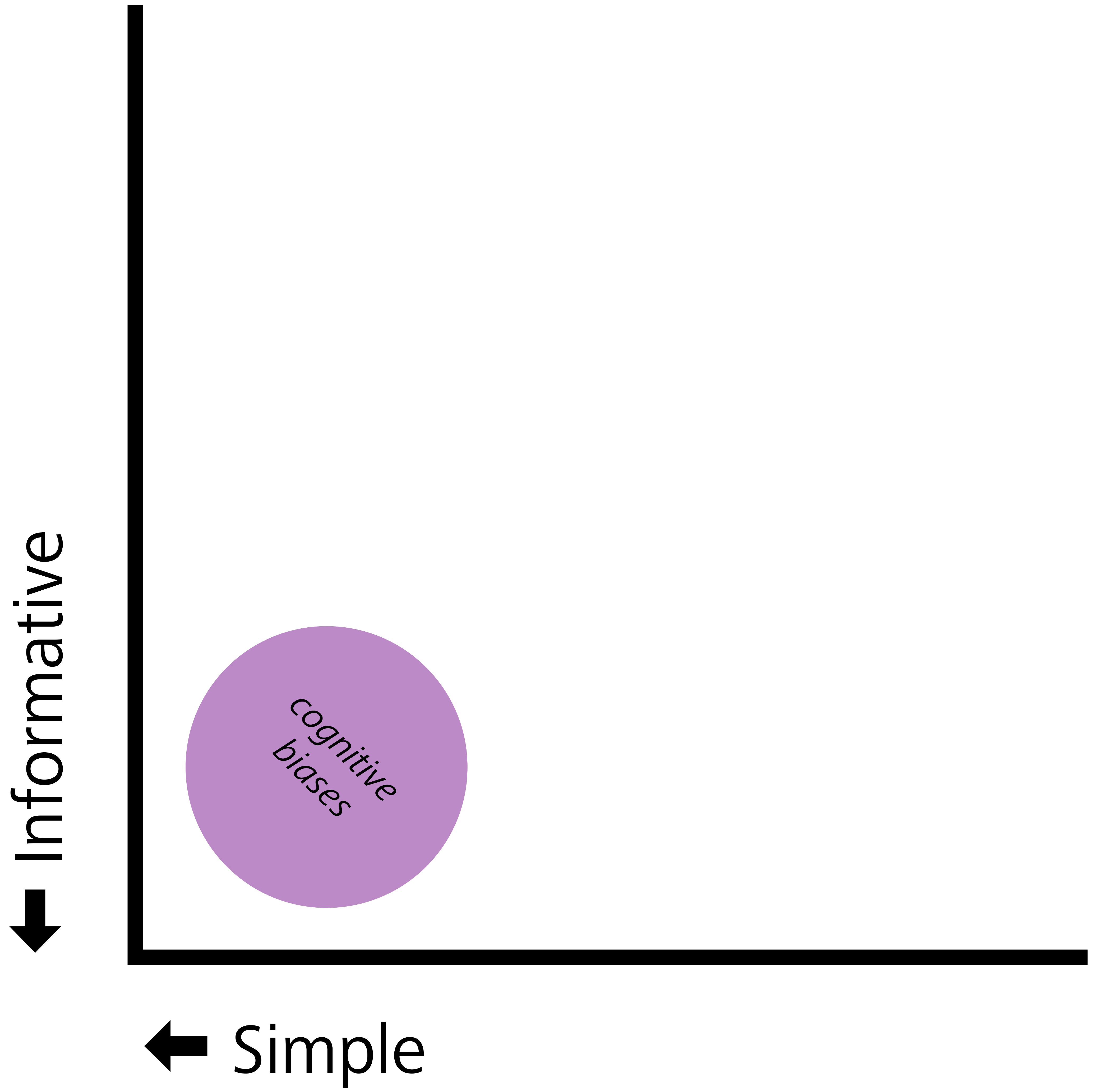


Carstensen, Xu, Smith, Regier (2015)

Informativeness from learning biases

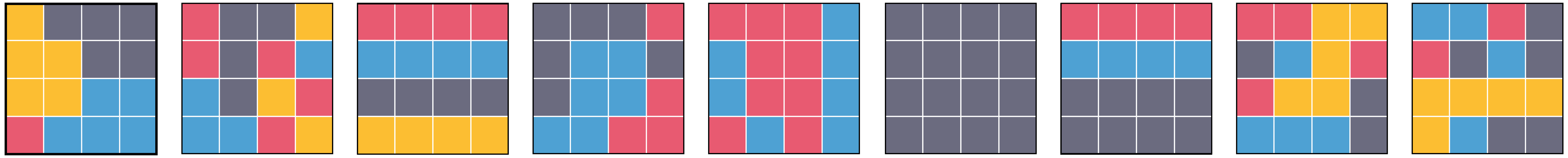


Informativeness from learning biases

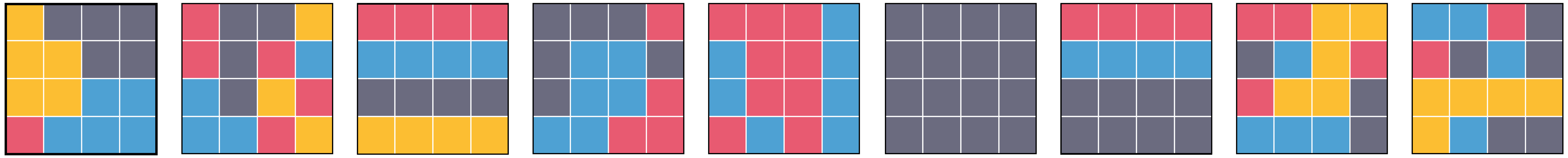


Bayesian model

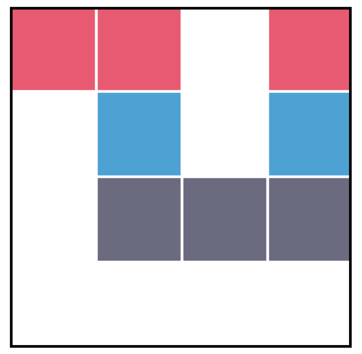
Bayesian inference

$\mathcal{L} = \{$  $\cdots \}$

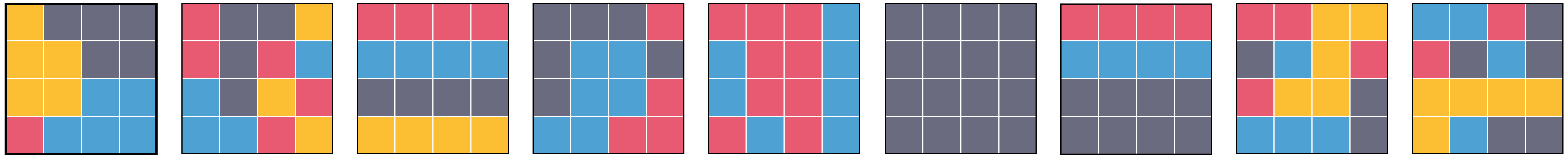
Bayesian inference

$\mathcal{L} = \{$  $\cdots\}$

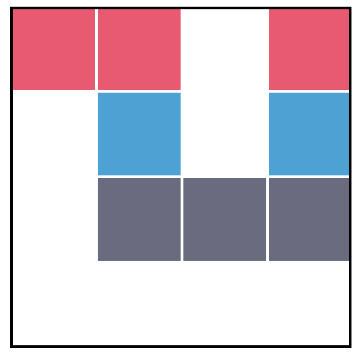
$$D = [\langle m_1, s_1\rangle, \langle m_2, s_2\rangle, \langle m_3, s_3\rangle, \ldots, \langle m_n, s_n\rangle]$$



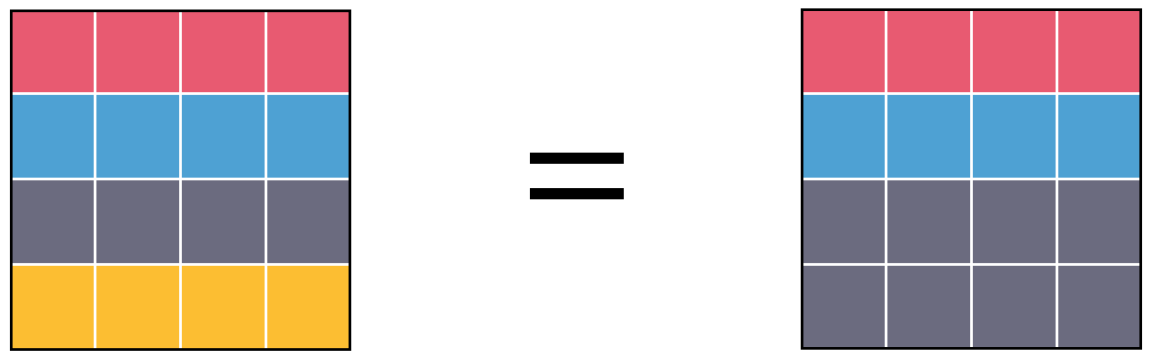
Bayesian inference

$\mathcal{L} = \{$  $\cdots\}$

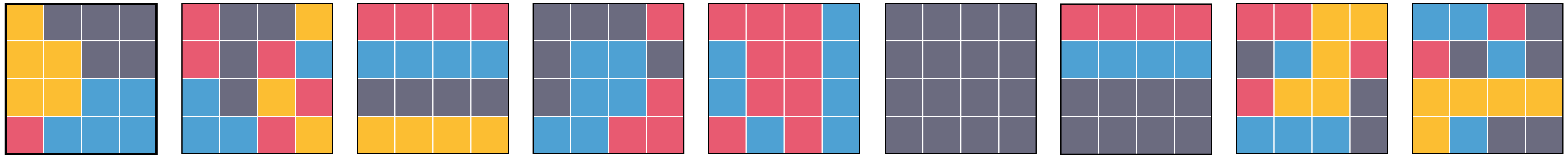
$$D = [\langle m_1, s_1\rangle, \langle m_2, s_2\rangle, \langle m_3, s_3\rangle, \ldots, \langle m_n, s_n\rangle]$$



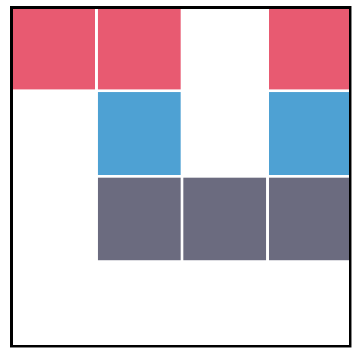
$$\text{likelihood}(D|L) = \prod_{\langle m,s\rangle} P(s|L, m)$$



Bayesian inference

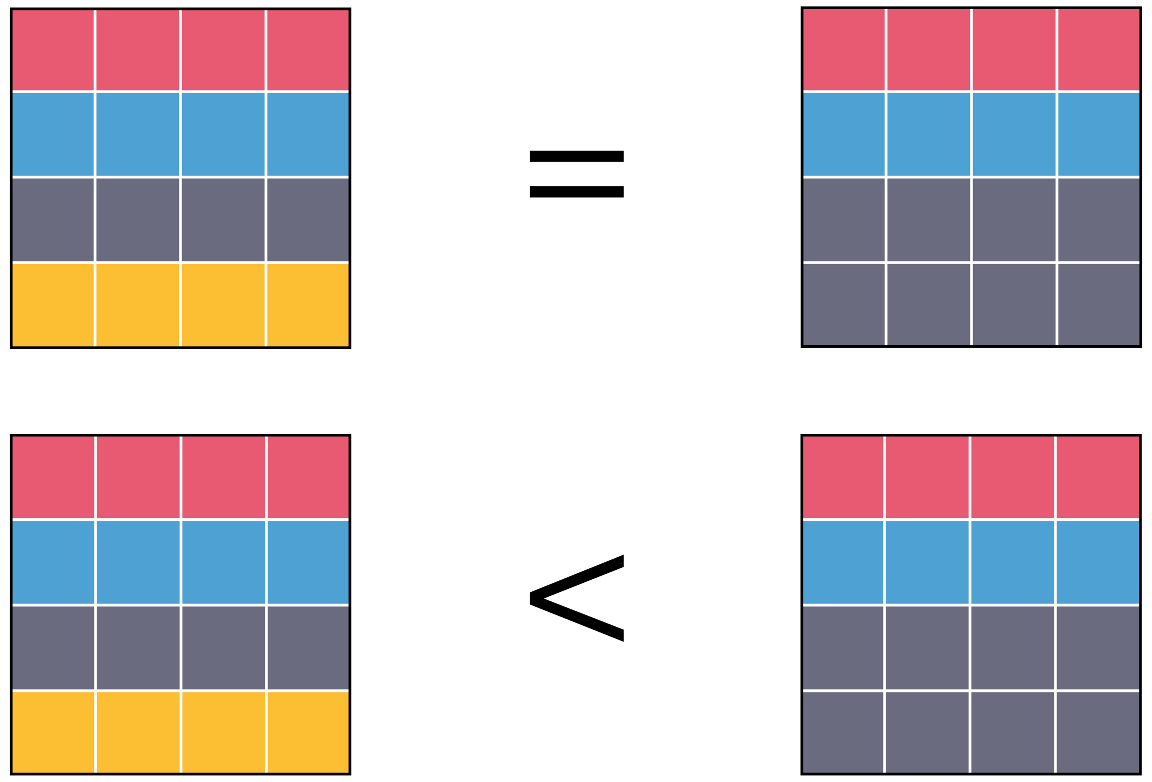
$\mathcal{L} = \{$  $\cdots\}$

$D = [\langle m_1, s_1\rangle, \langle m_2, s_2\rangle, \langle m_3, s_3\rangle, \ldots, \langle m_n, s_n\rangle]$

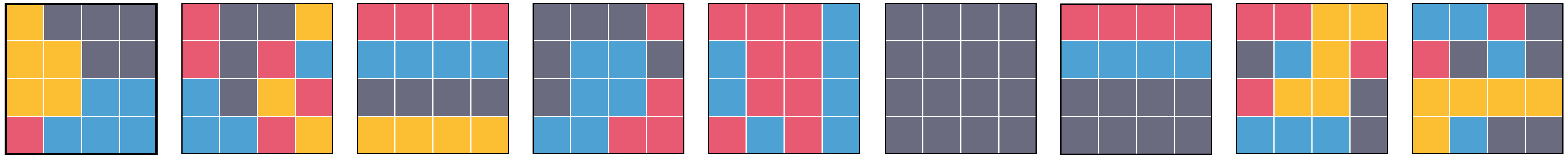


$$\text{likelihood}(D|L) = \prod_{\langle m,s\rangle} P(s|L,m)$$

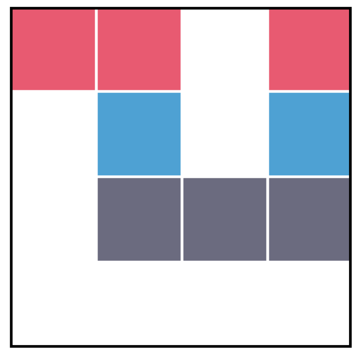
$$\text{prior}(L) \propto 2^{-\text{complexity}(L)}$$



Bayesian inference

$\mathcal{L} = \{$  $\cdots\}$

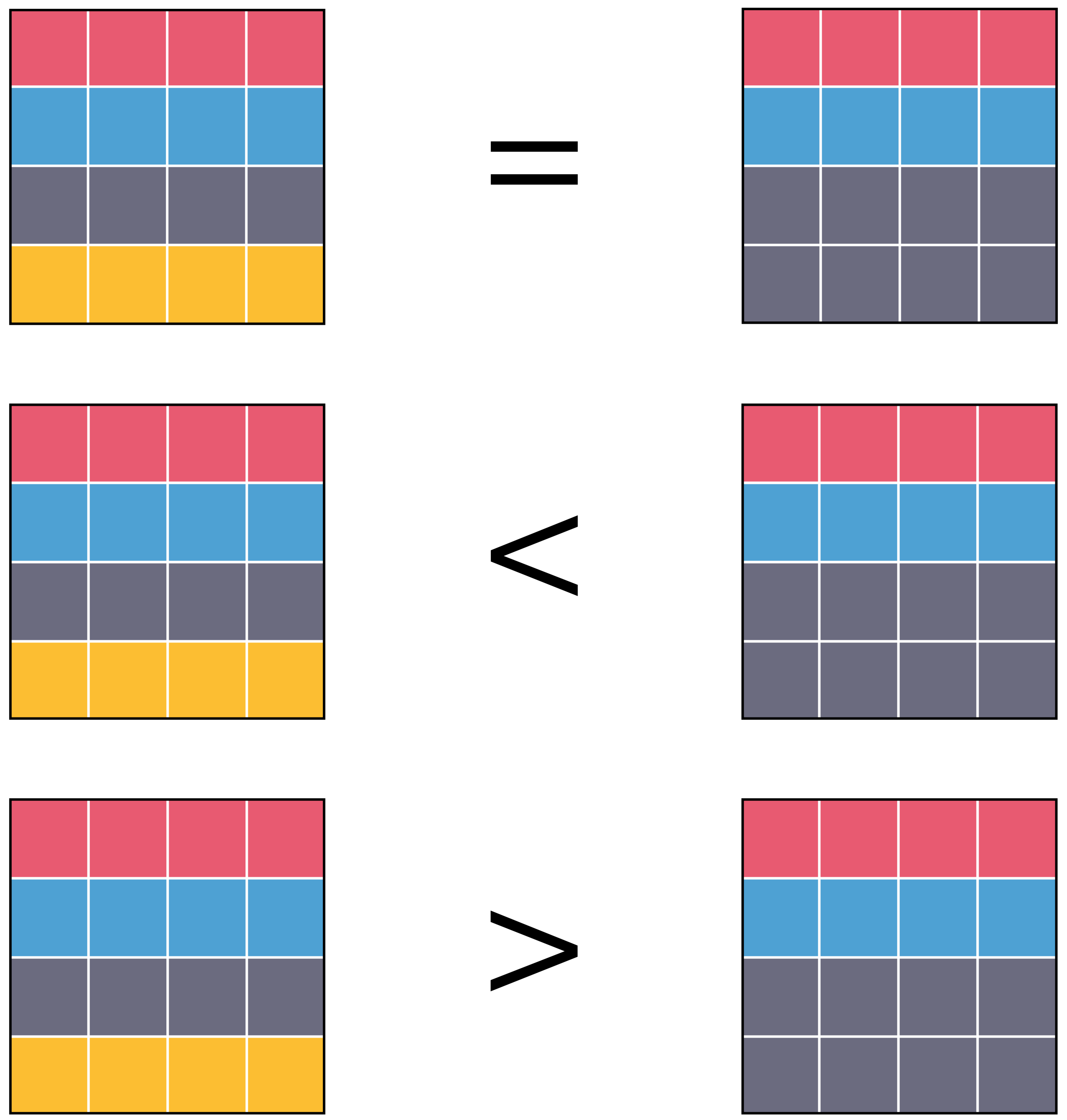
$D = [\langle m_1, s_1 \rangle, \langle m_2, s_2 \rangle, \langle m_3, s_3 \rangle, ..., \langle m_n, s_n \rangle]$



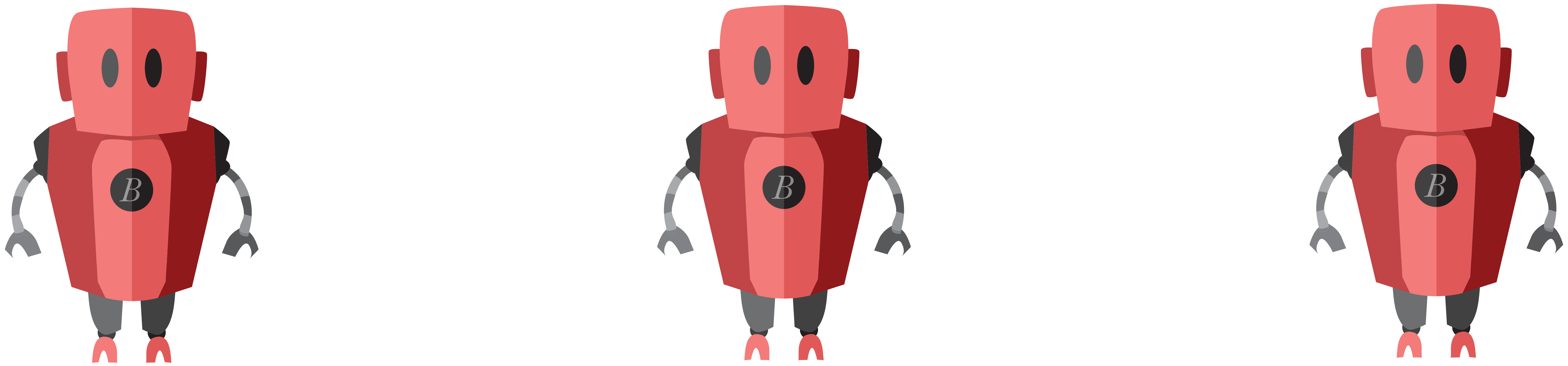
$$\text{likelihood}(D|L) = \prod_{\langle m,s \rangle} P(s|L, m)$$

$$\text{prior}(L) \propto 2^{-\text{complexity}(L)}$$

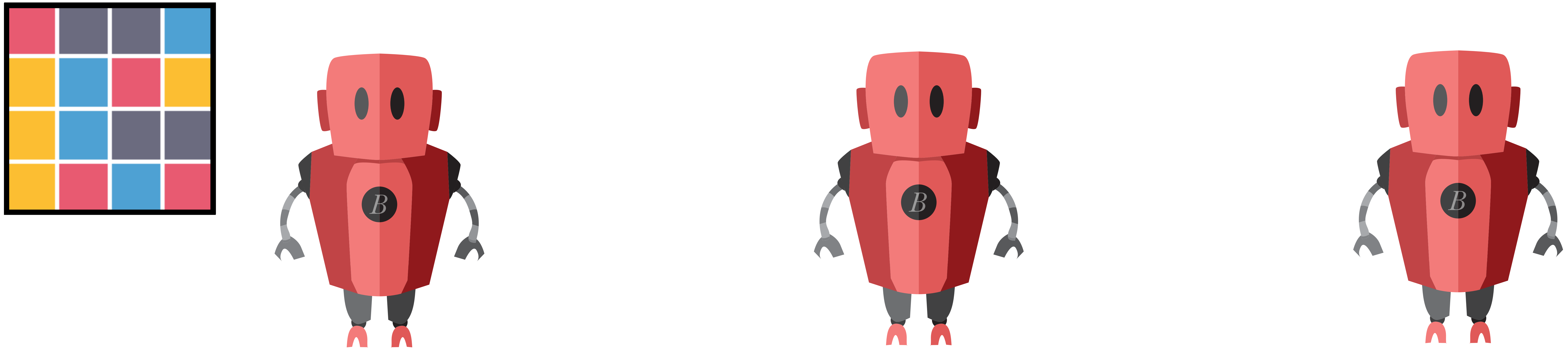
$$\text{prior}(L) \propto 2^{-\text{cost}(L)}$$



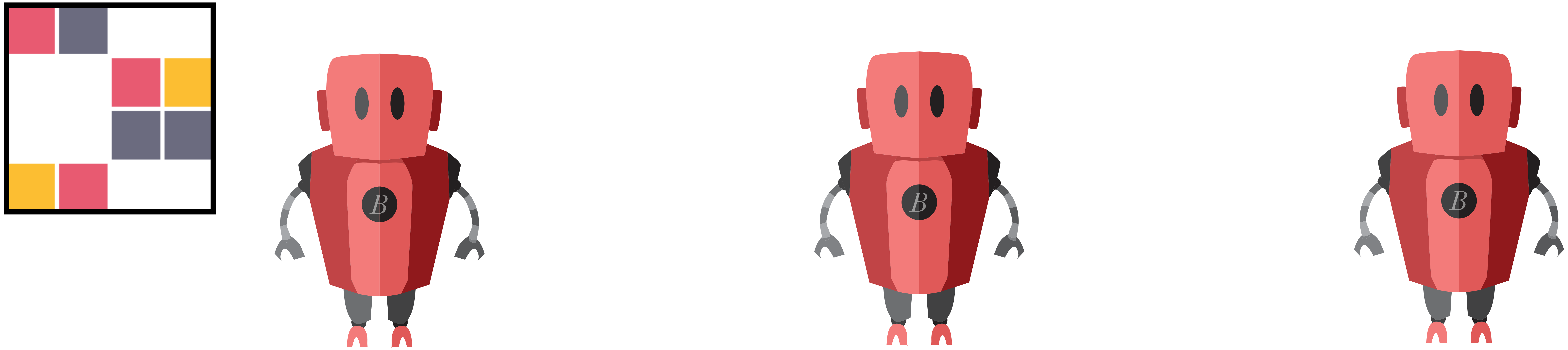
Bayesian iterated learning under a **simplicity prior**



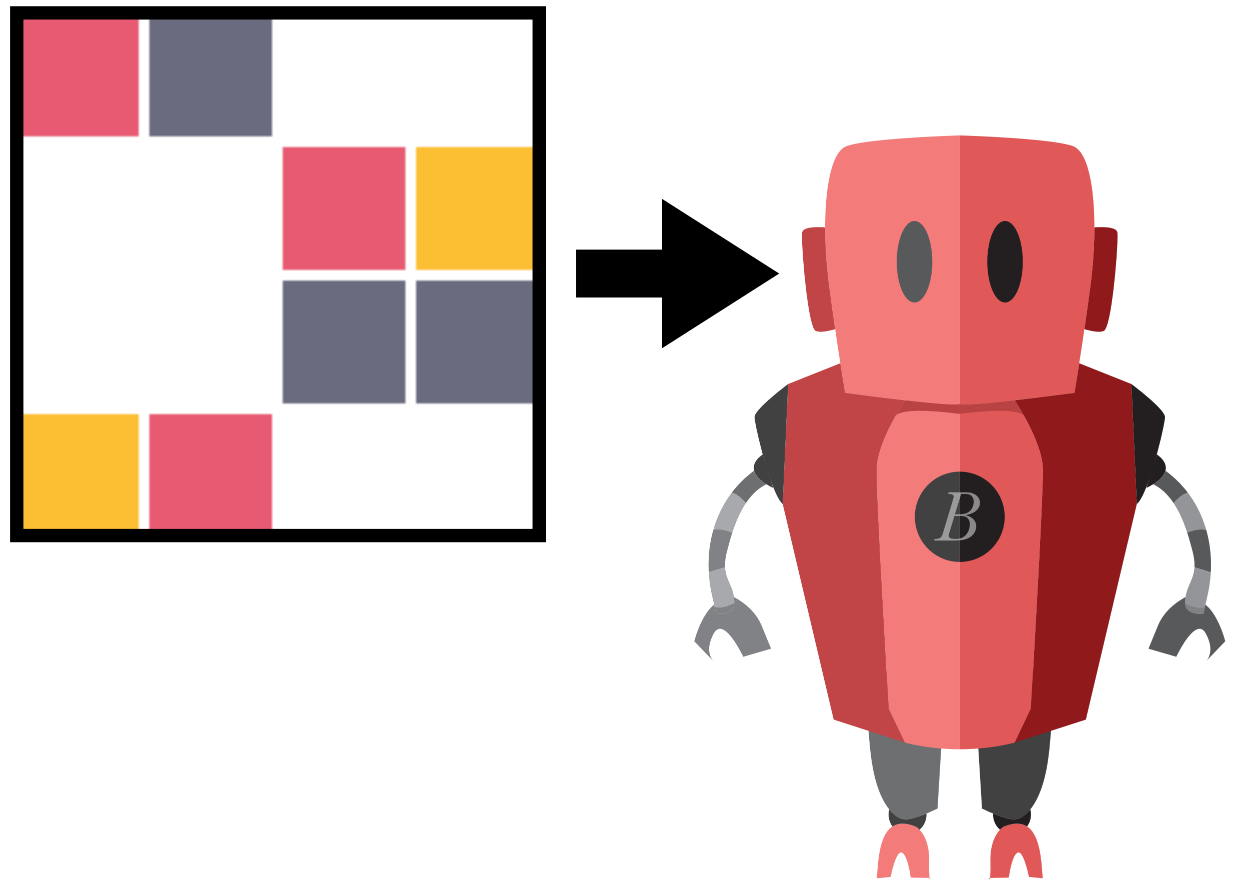
Bayesian iterated learning under a **simplicity prior**



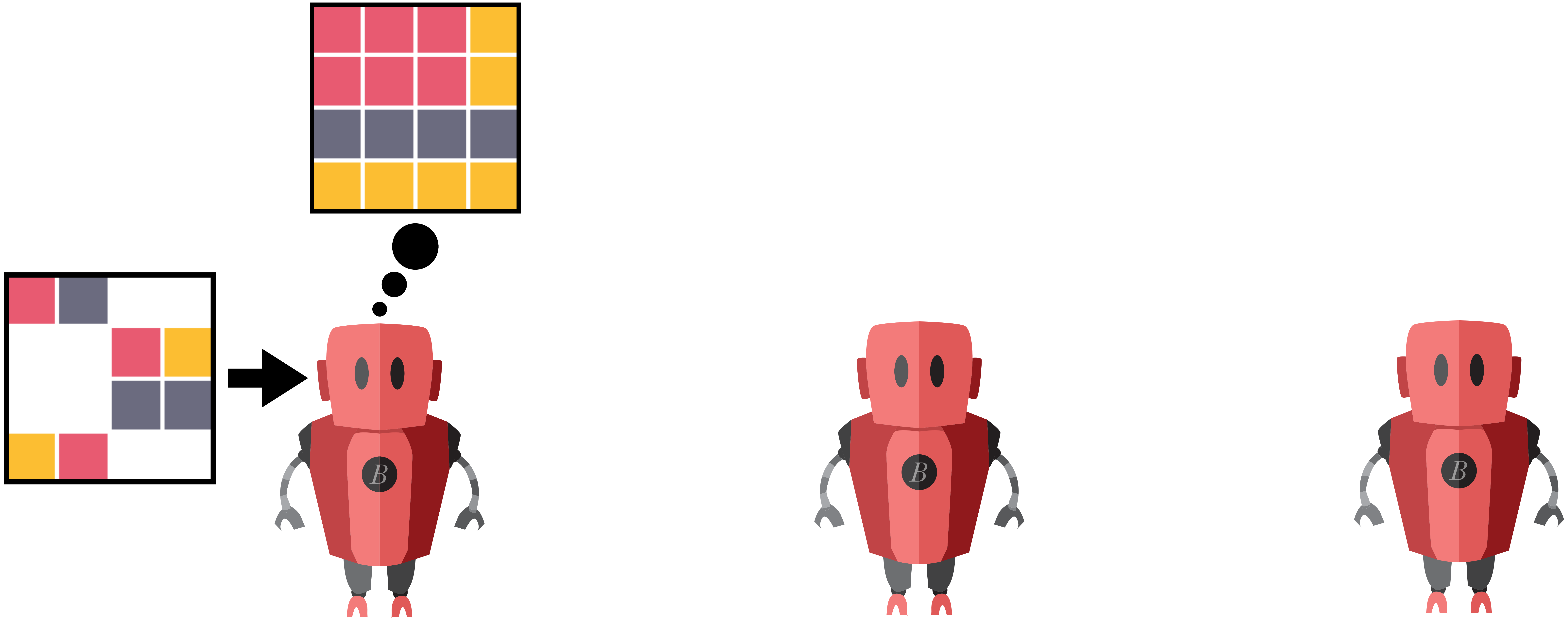
Bayesian iterated learning under a **simplicity prior**



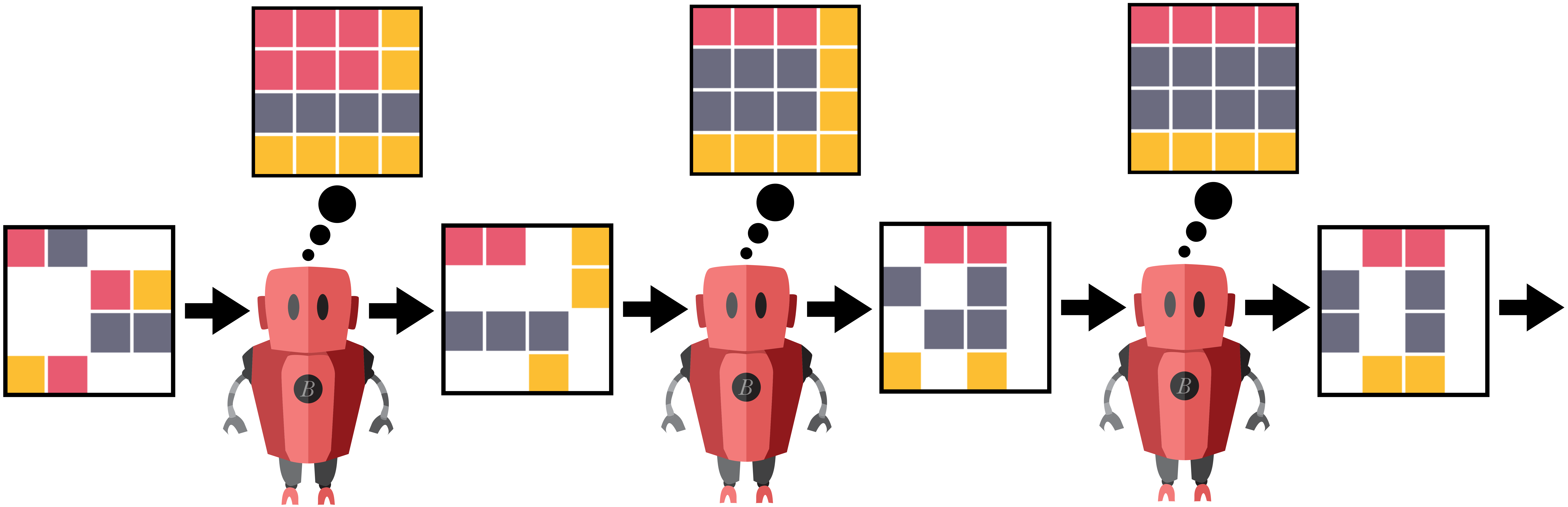
Bayesian iterated learning under a **simplicity prior**



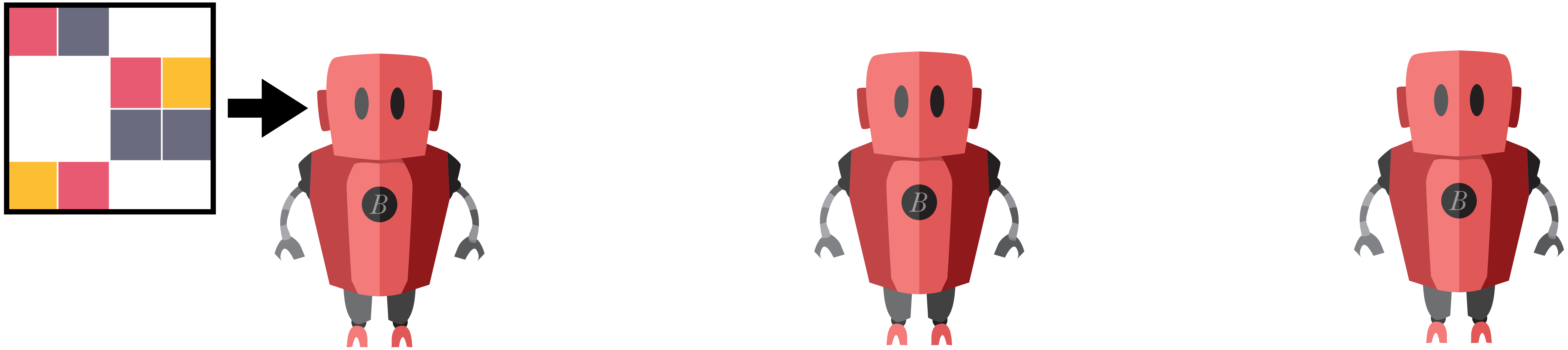
Bayesian iterated learning under a **simplicity prior**



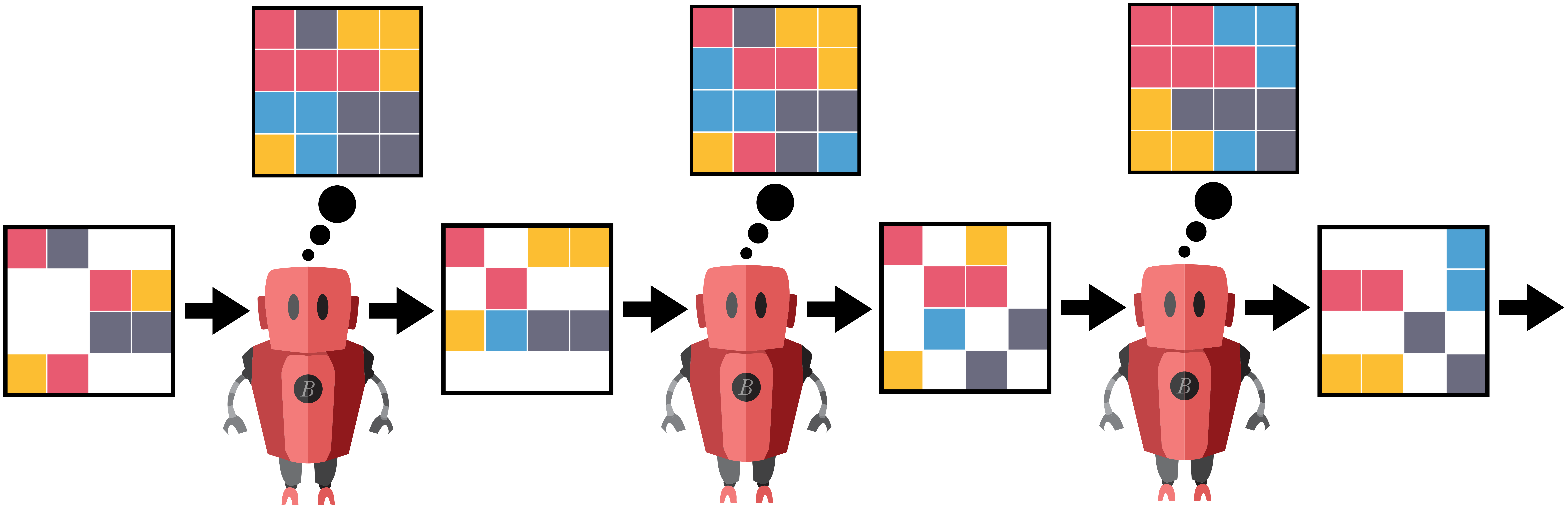
Bayesian iterated learning under a **simplicity prior**



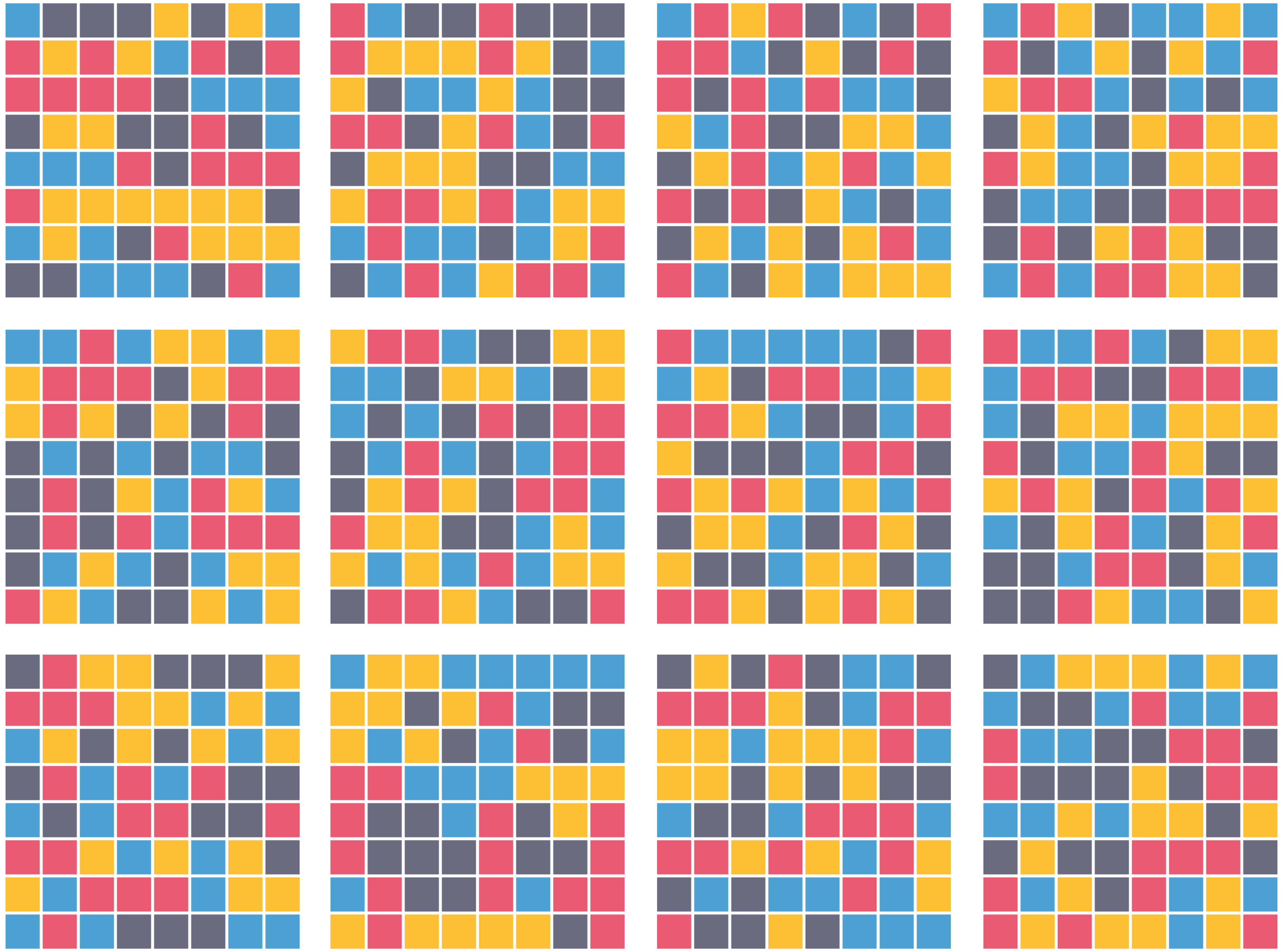
Bayesian iterated learning under an **informativeness prior**



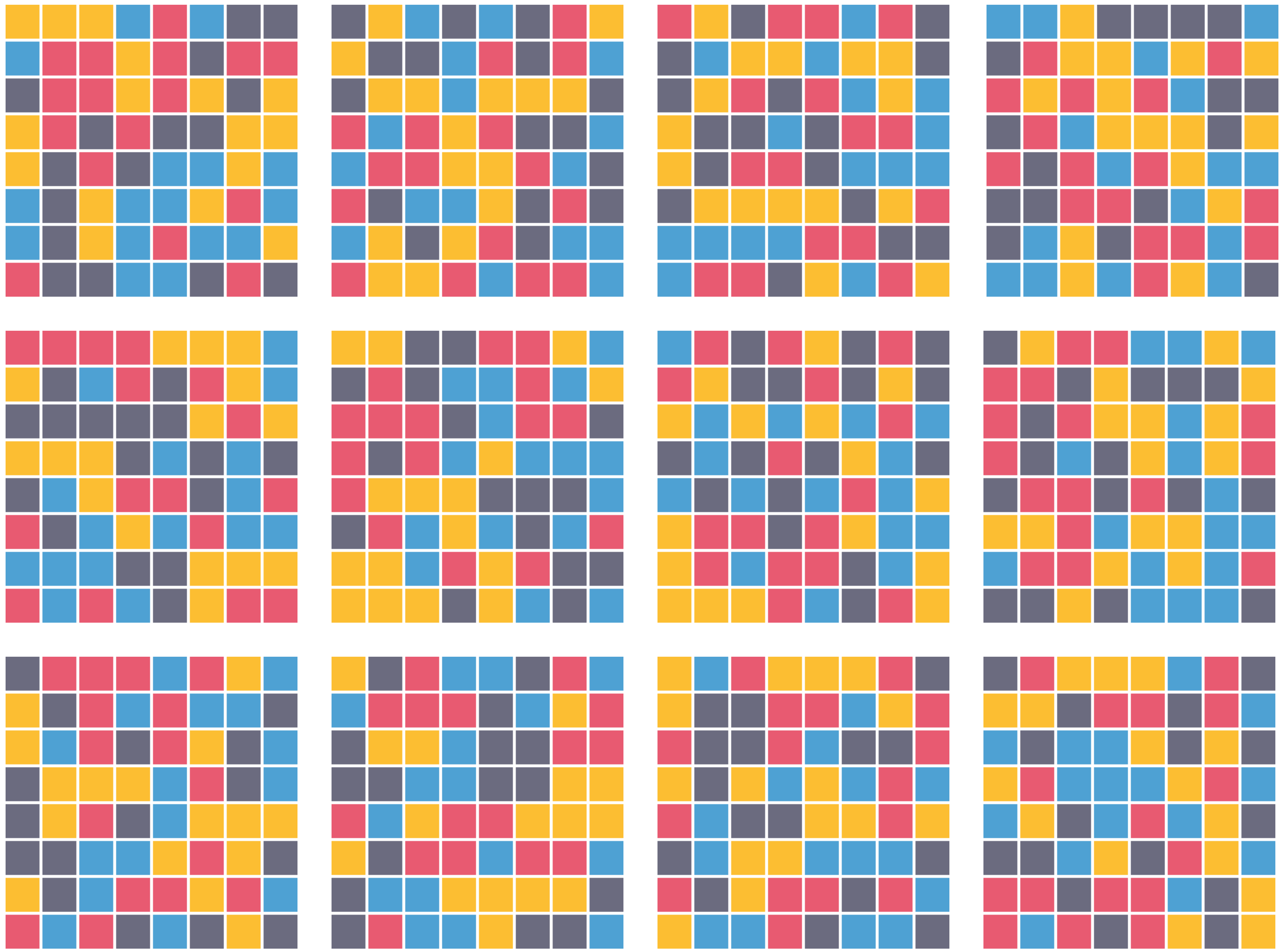
Bayesian iterated learning under an **informativeness prior**



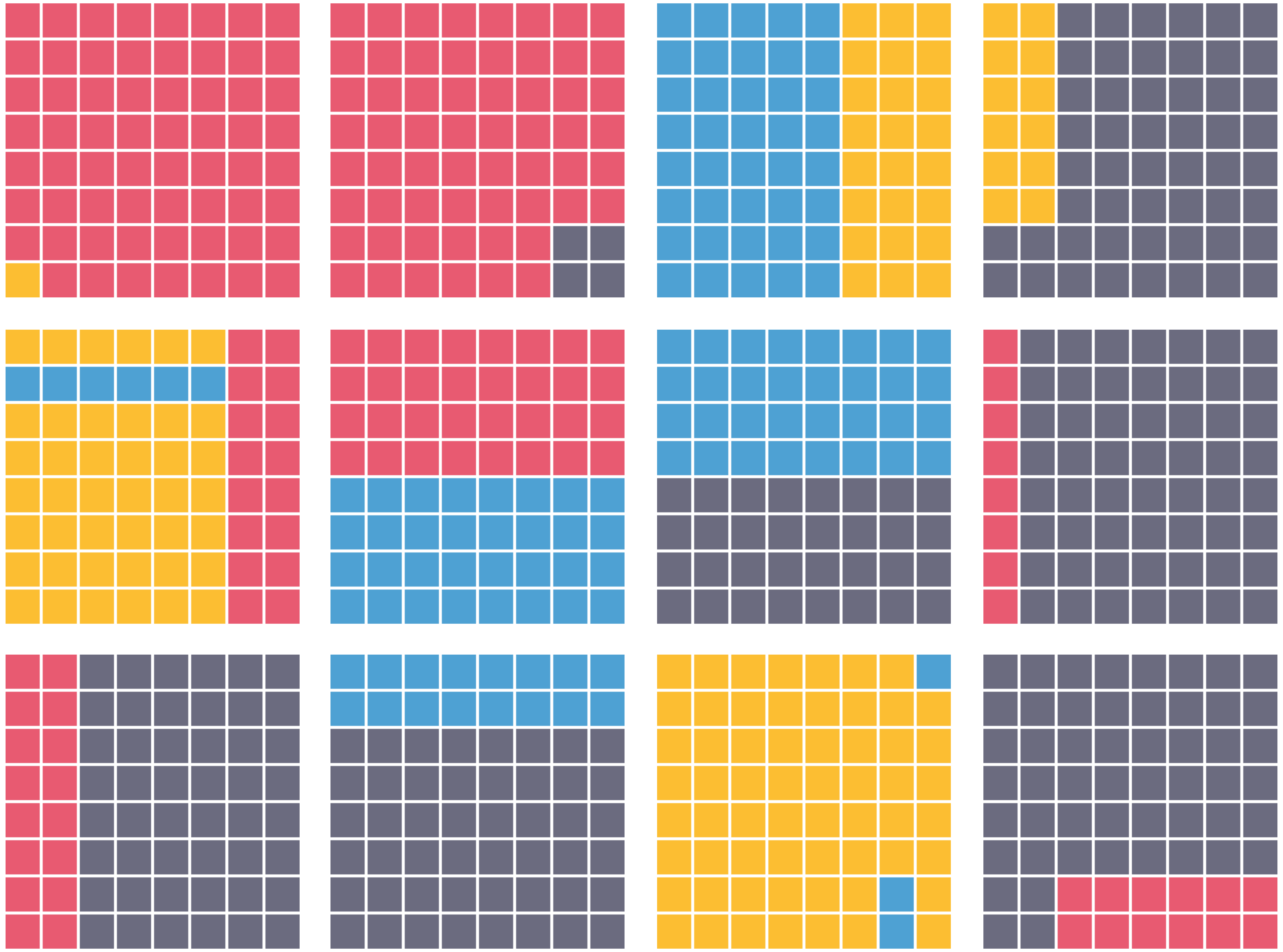
Simplicity prior



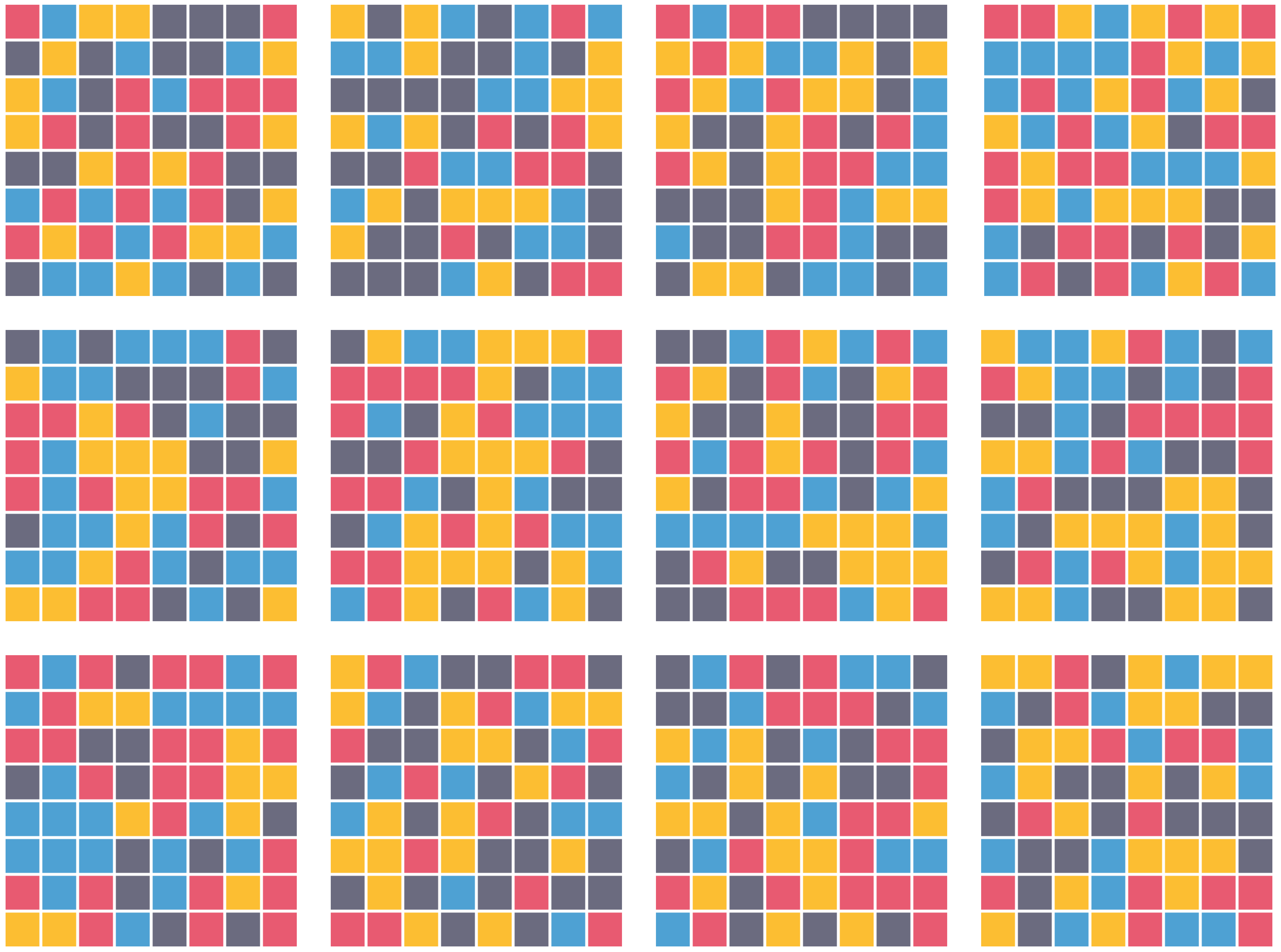
Informativeness prior



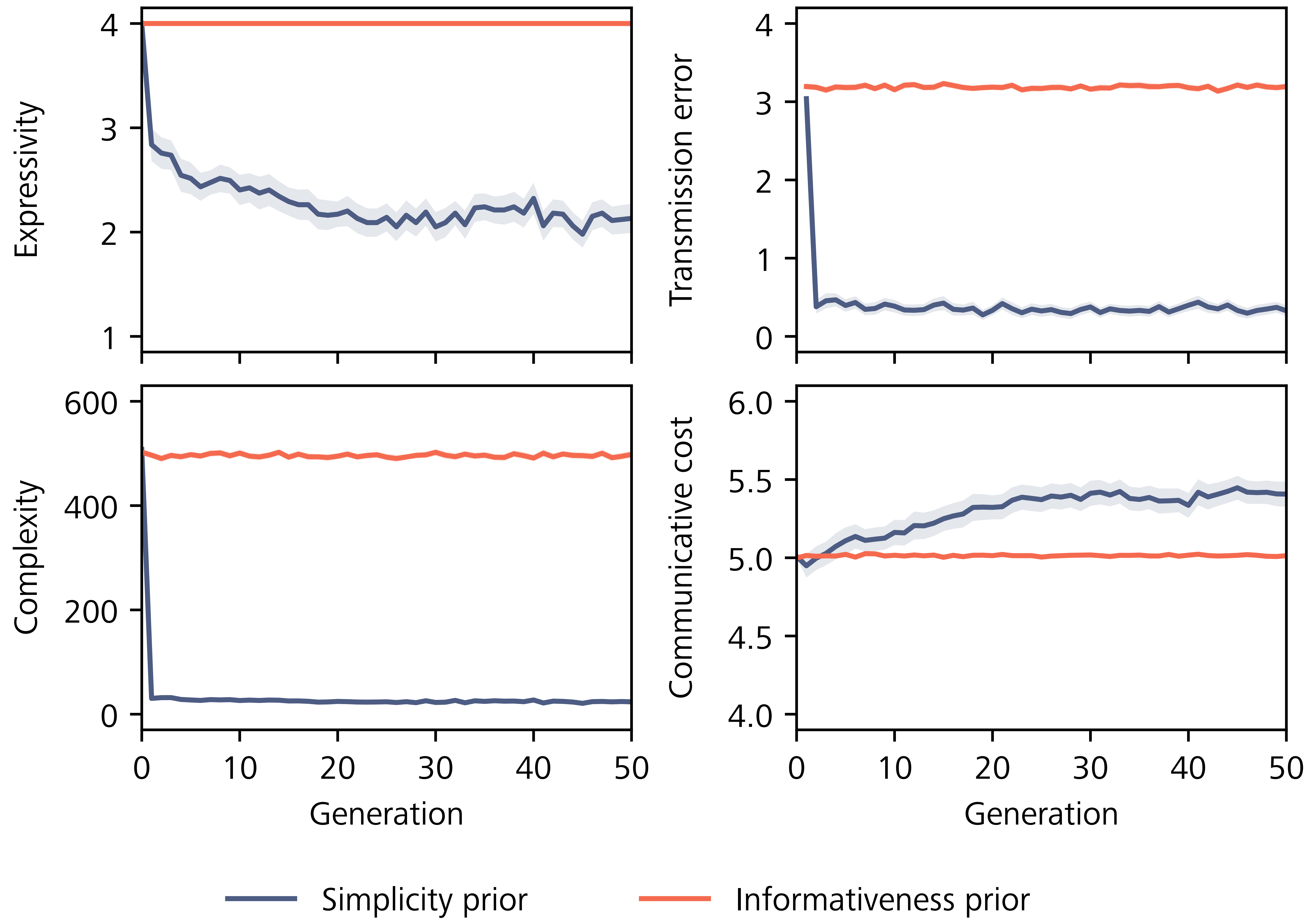
Simplicity prior



Informativeness prior



Model results



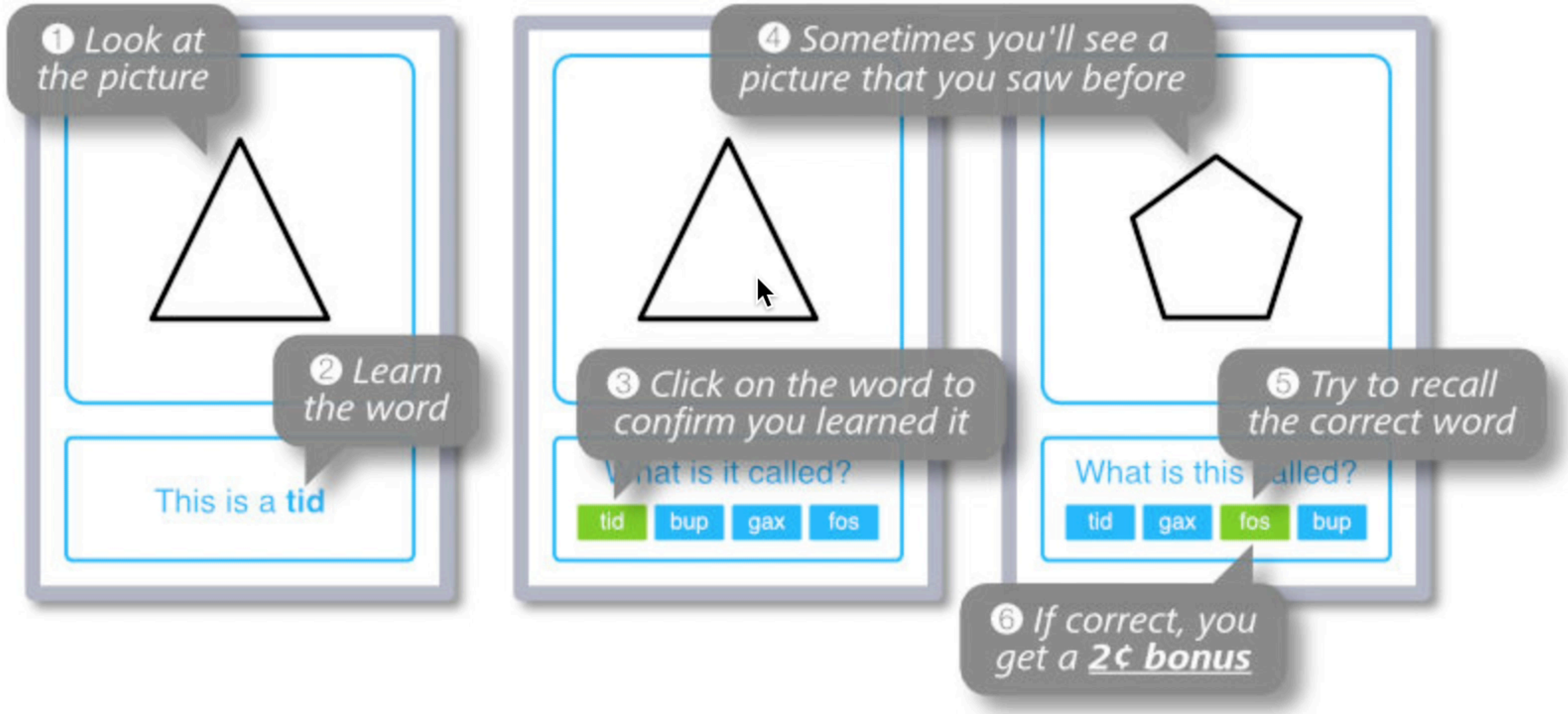
Experiment

localhost/~jon/shepard/

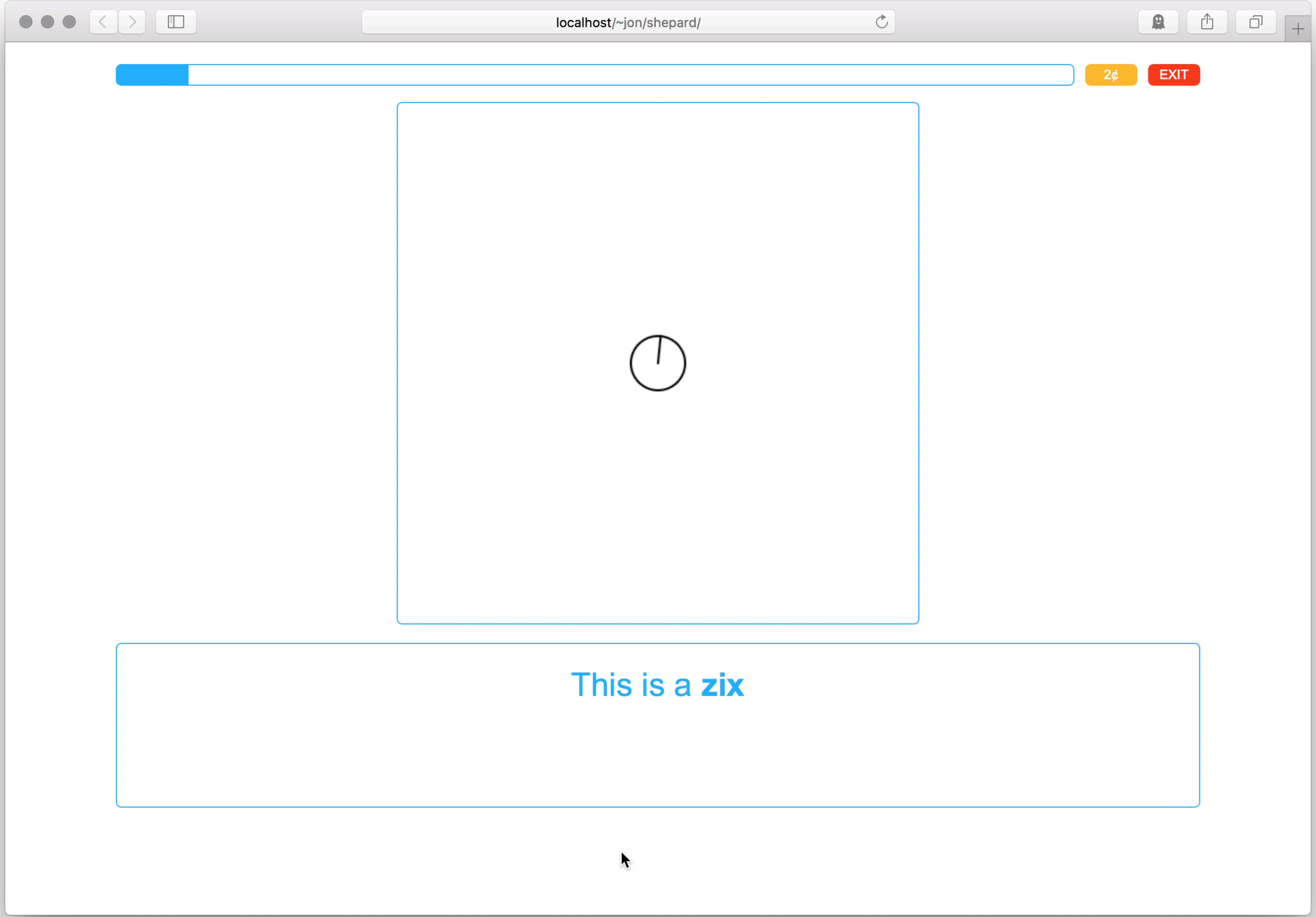
Stage 1: Training

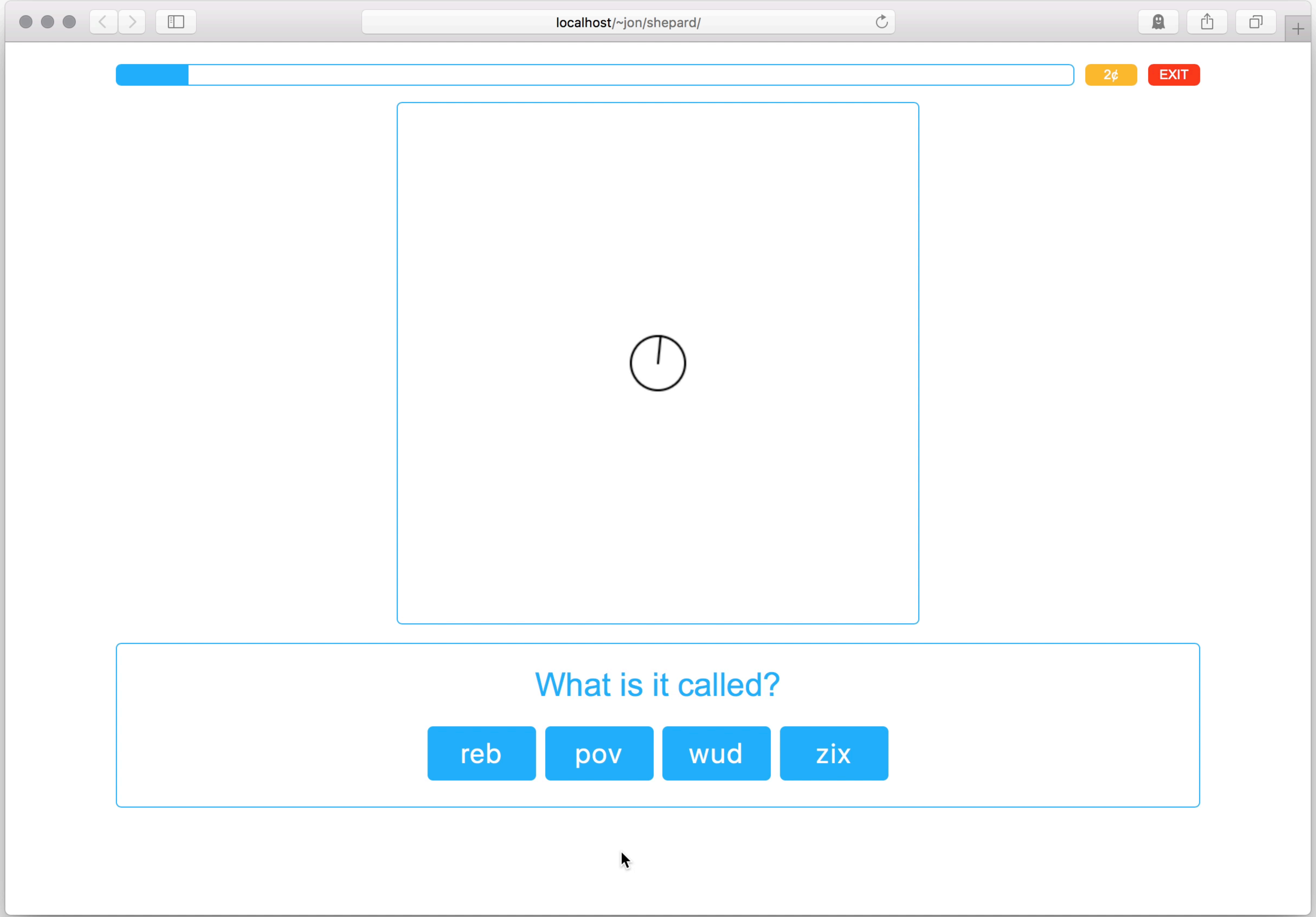
15 minutes

You are going to learn a simple language. **We will train you on 4 words** in the language and **we will test how well you are learning the words**. Try to learn the language as well as you can and **aim to be accurate in your answers**. You will receive a **2¢ bonus payment** for every correct test answer. If you decide to stop the task, please click the EXIT button so that someone else can take part.



START


localhost/~jon/shepard/
2¢
EXIT
This is a **zix**


localhost/~jon/shepard/
2¢
EXIT
What is it called?
reb
pov
wud
zix

localhost/~jon/shepard/

Stage 2: Test

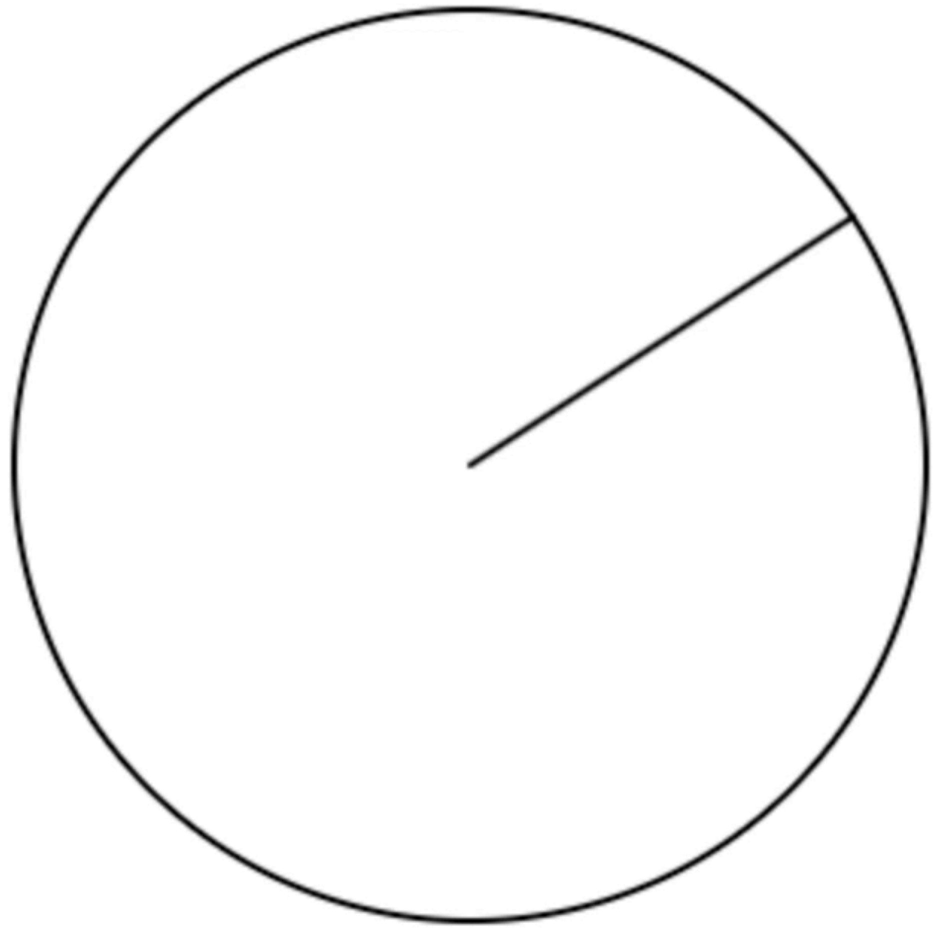
5 minutes

You have now completed the training stage! Next **we will test you on the language** that you just learned. For each picture, try to click on the correct word. You will get a **2¢ bonus payment** for every correct answer. It is therefore possible to earn **up to $1.28** in this stage of the task. However, this time **we will <u>not</u> tell you if you are correct or incorrect**. You will find out at the end how many you got correct.

START

localhost/~jon/shepard/

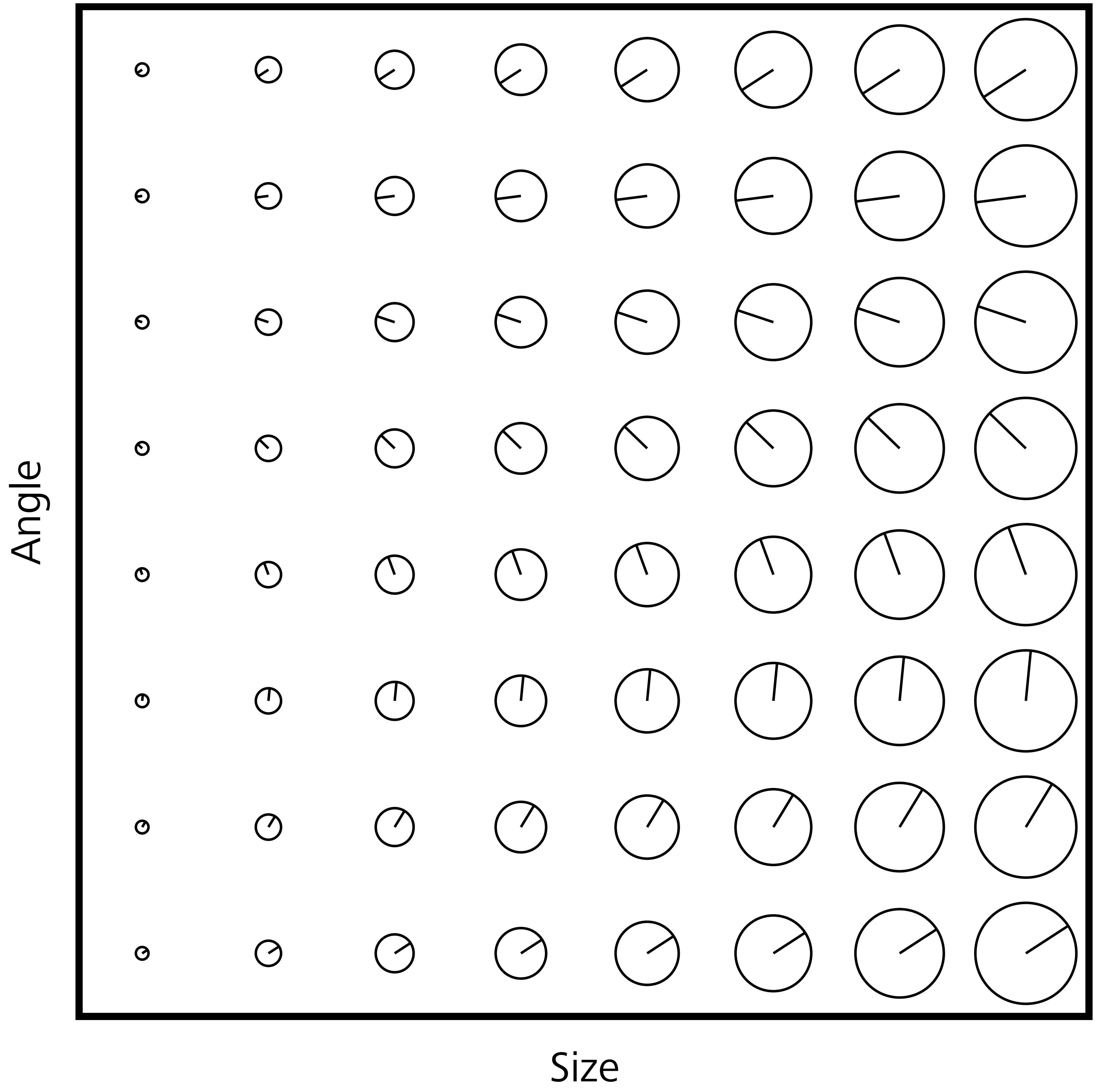
EXIT



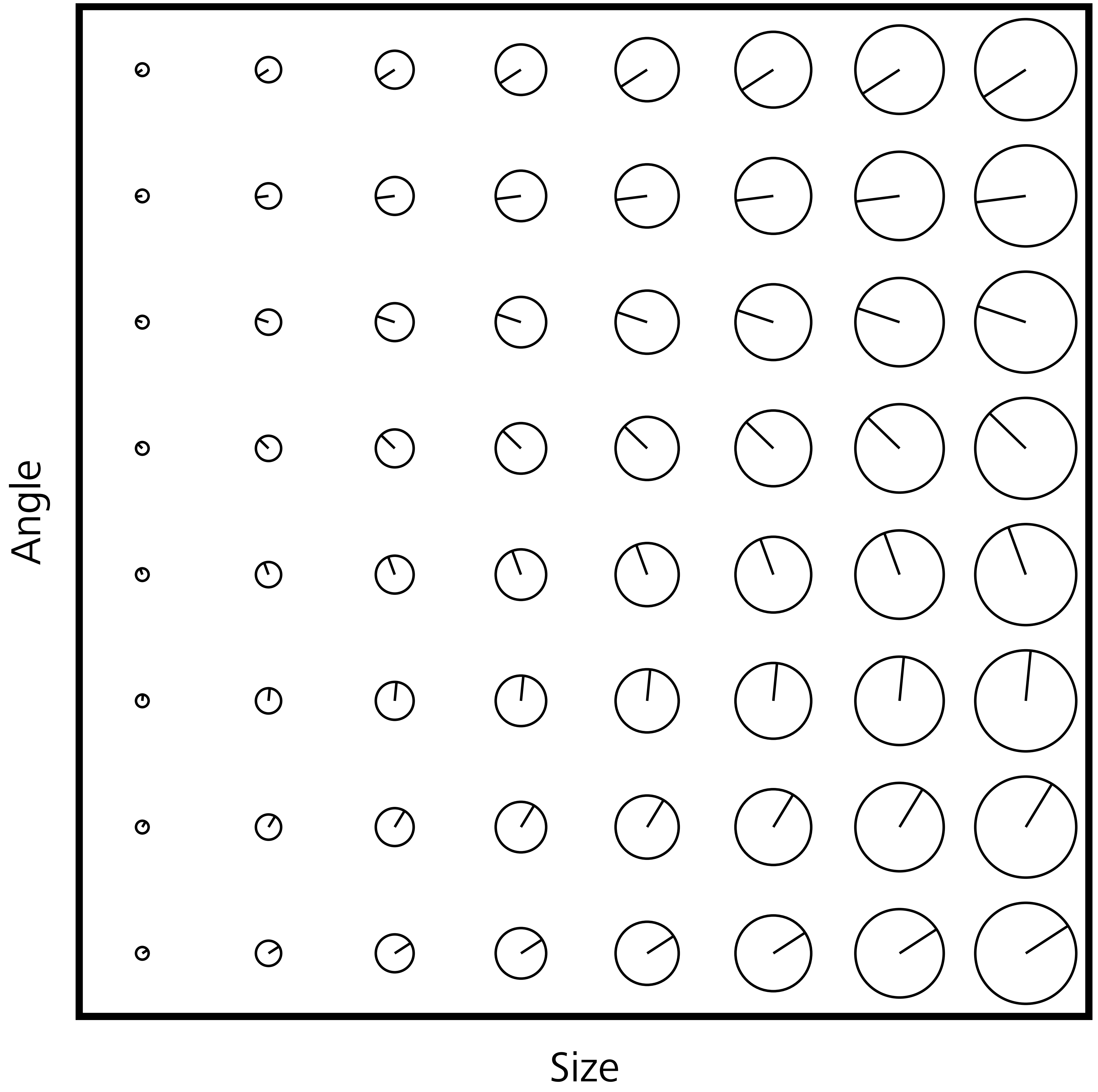
What is this called? +2¢ if correct

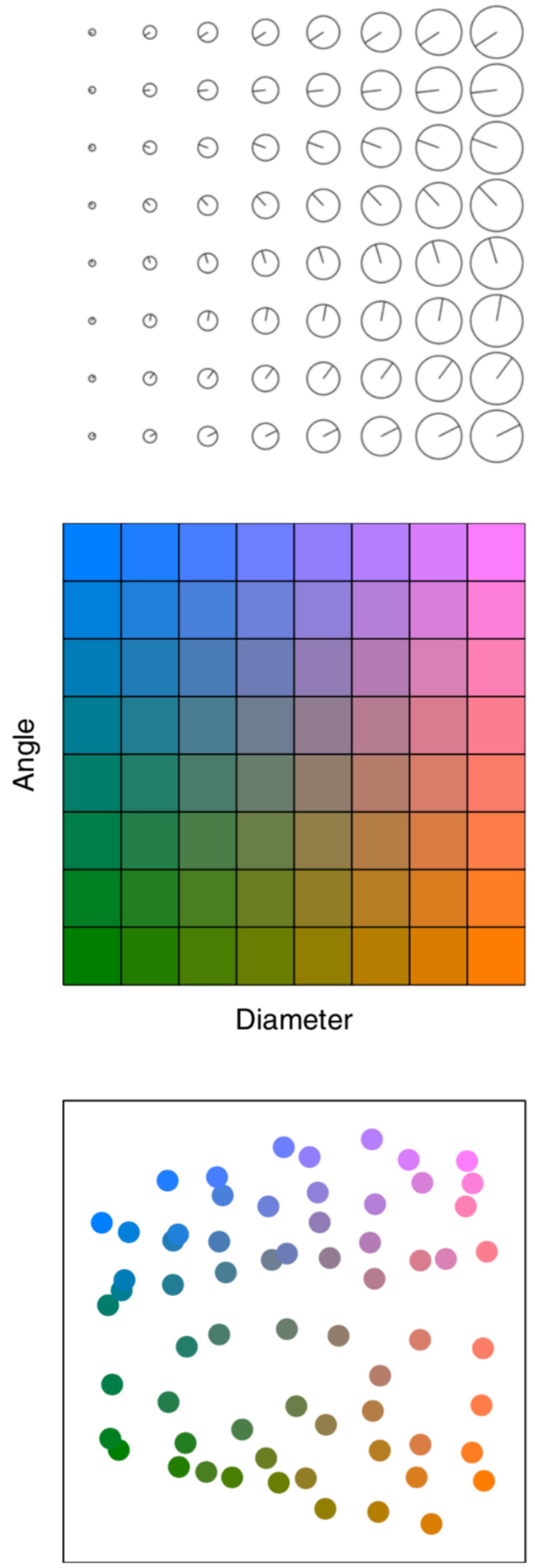
pov wud reb zix

Experimental stimuli



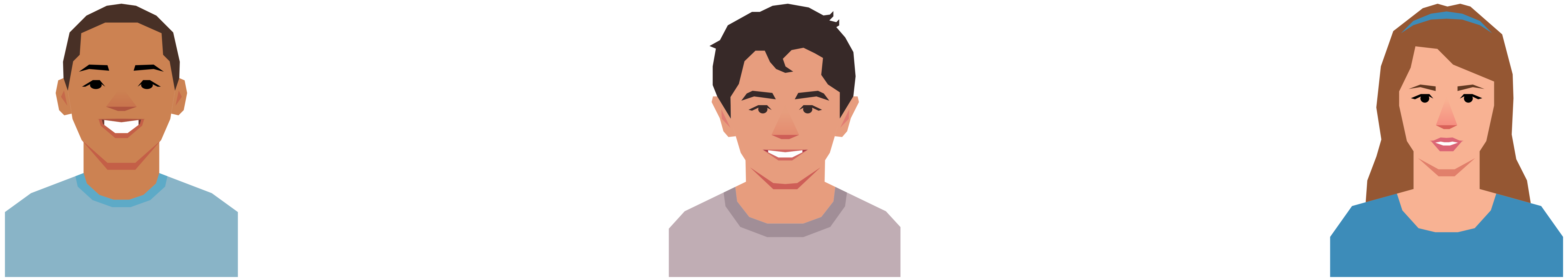
Experimental stimuli



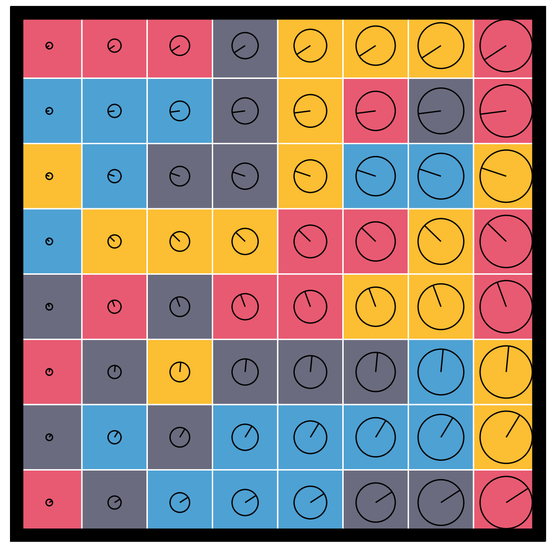
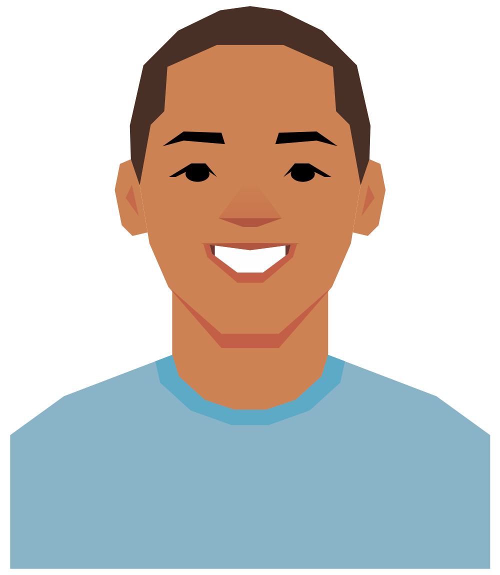
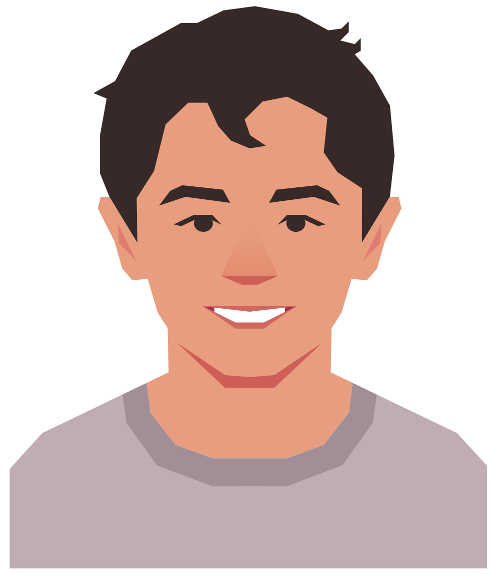


Canini et al. (2014)

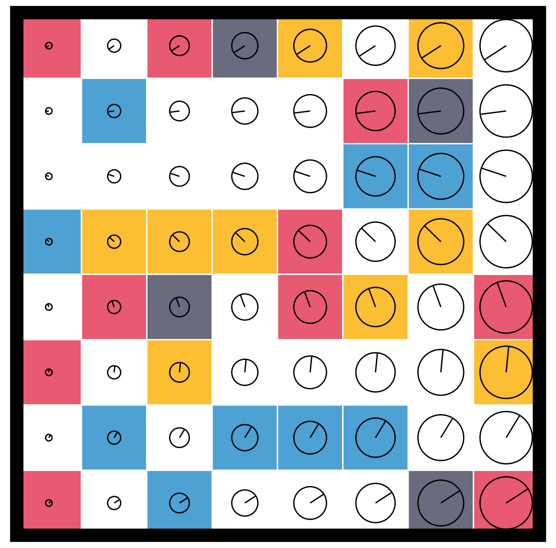
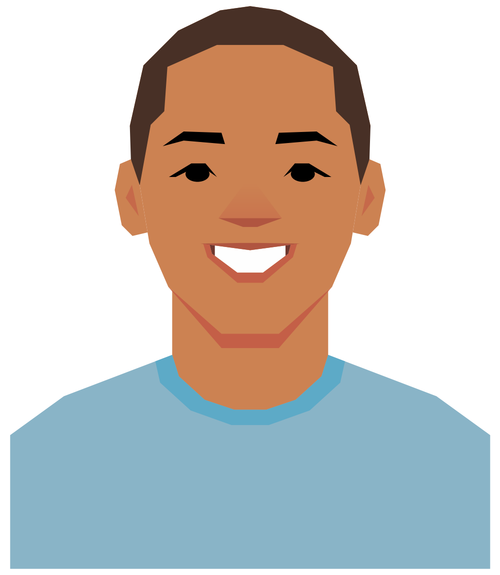
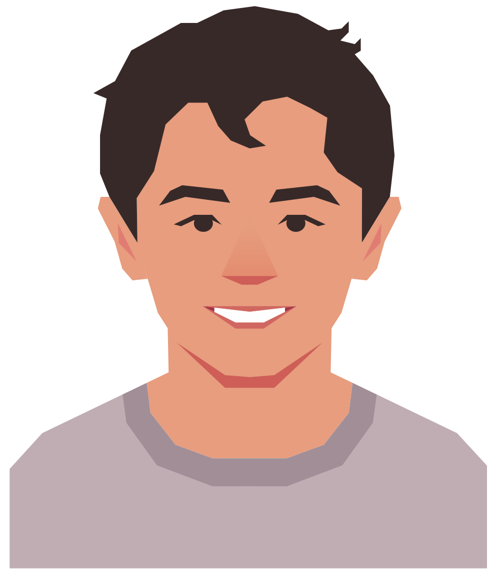
Iterated learning with humans



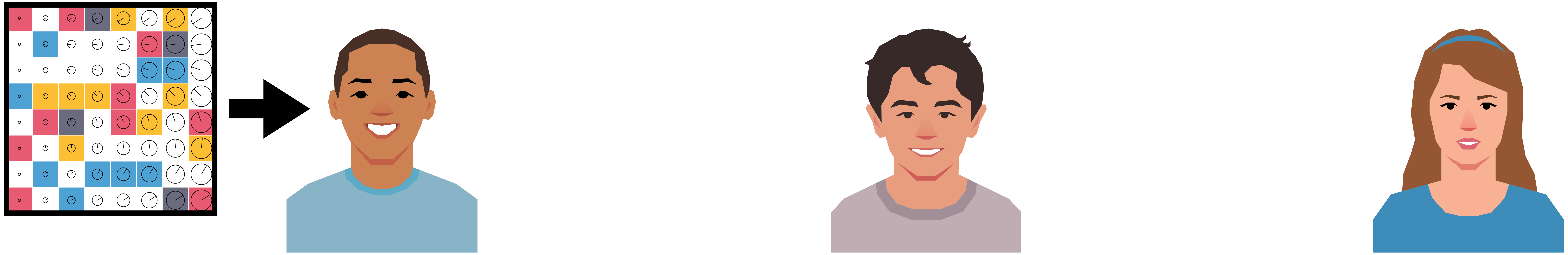
Iterated learning with humans

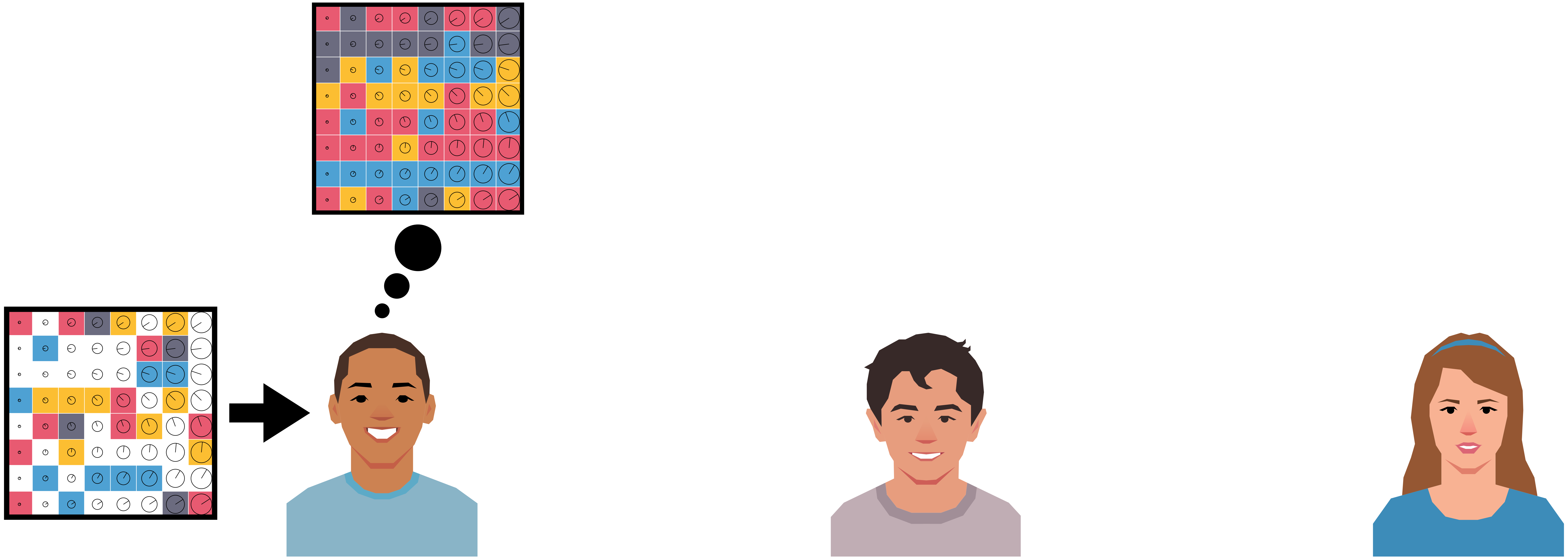

Iterated learning with humans

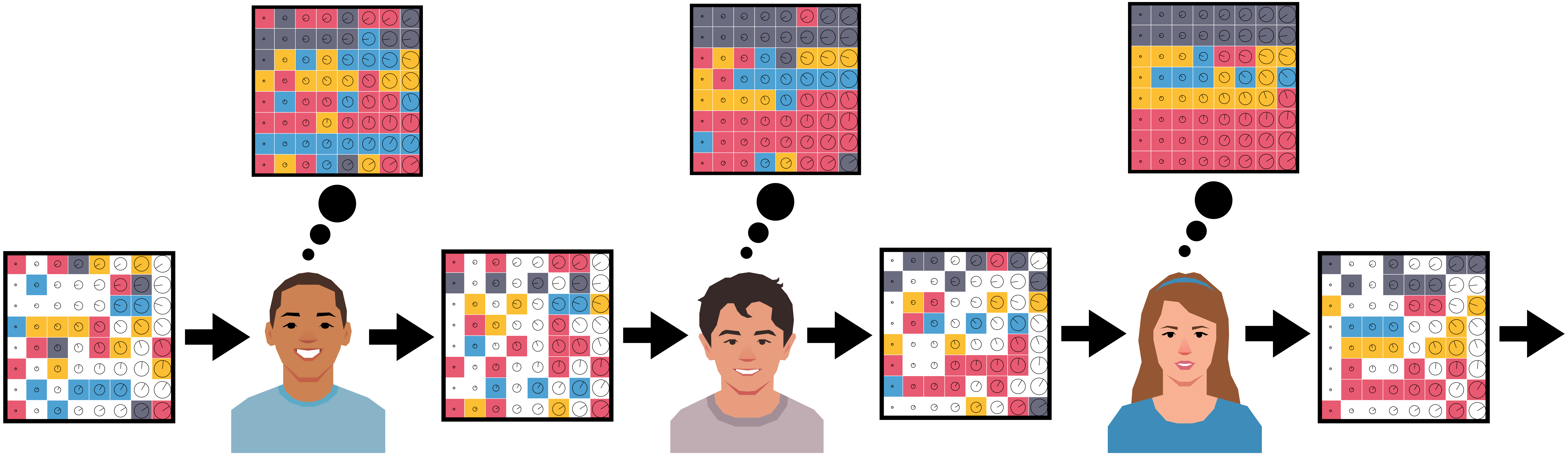

Iterated learning with humans

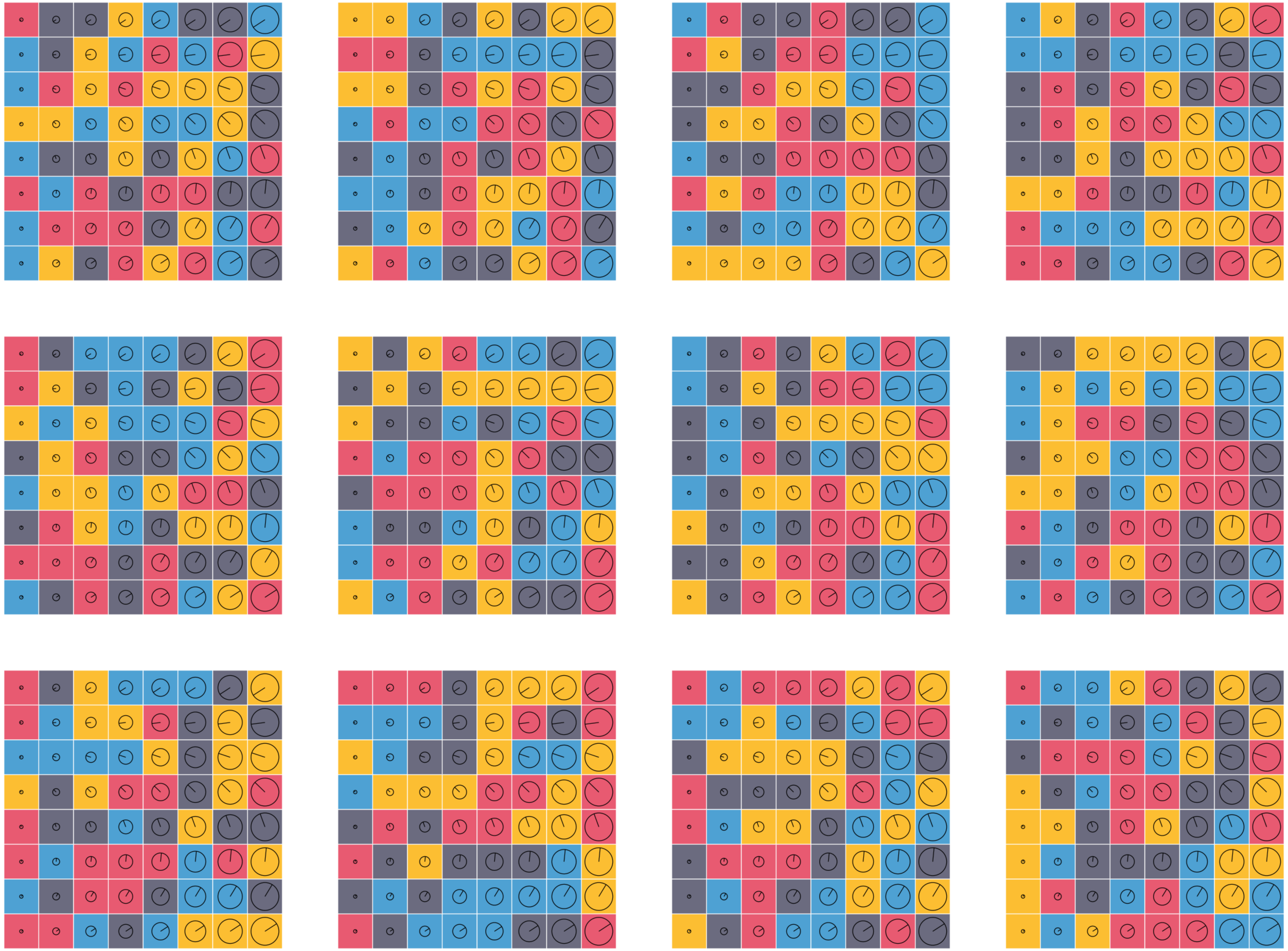


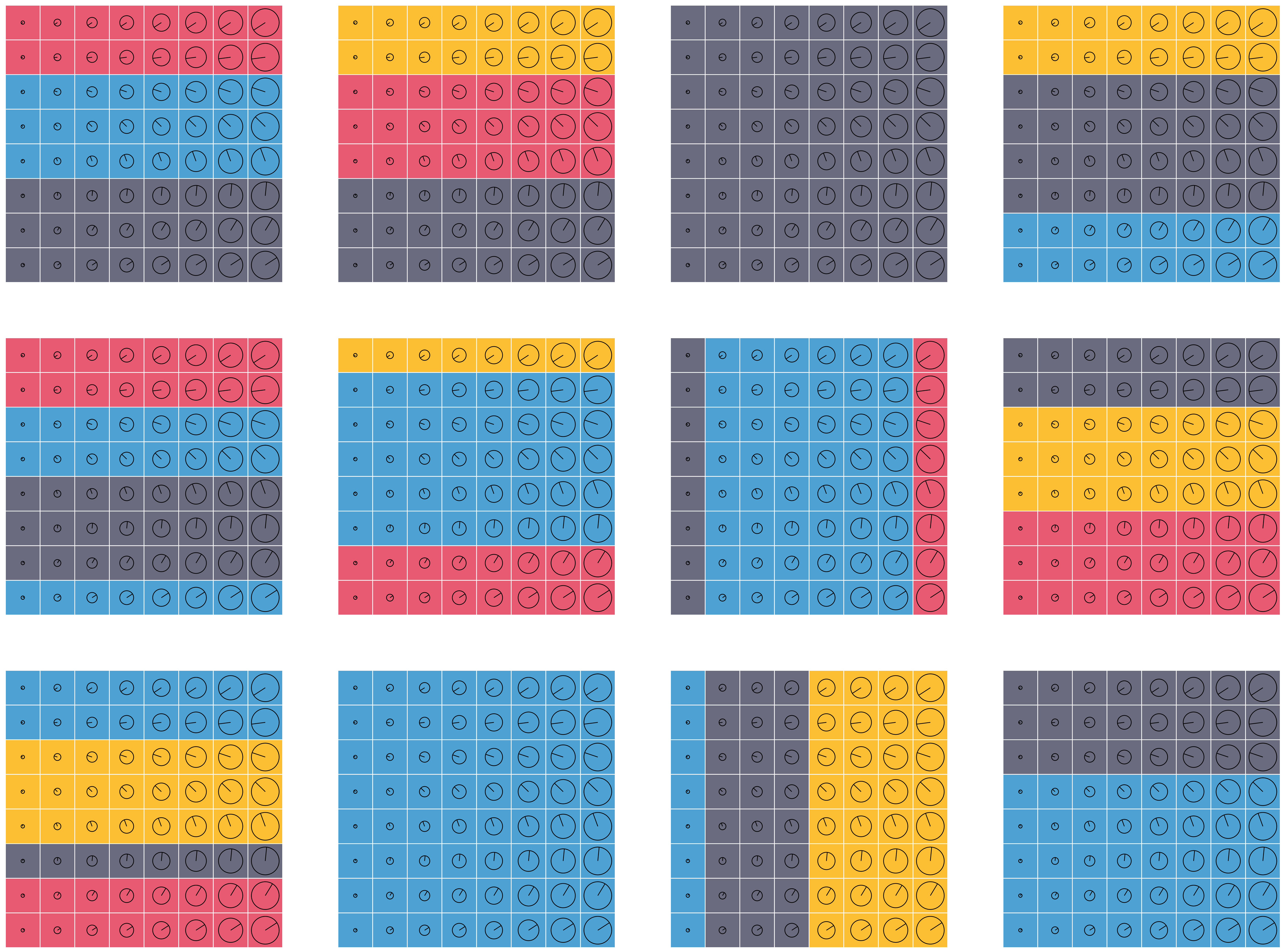
Iterated learning with humans



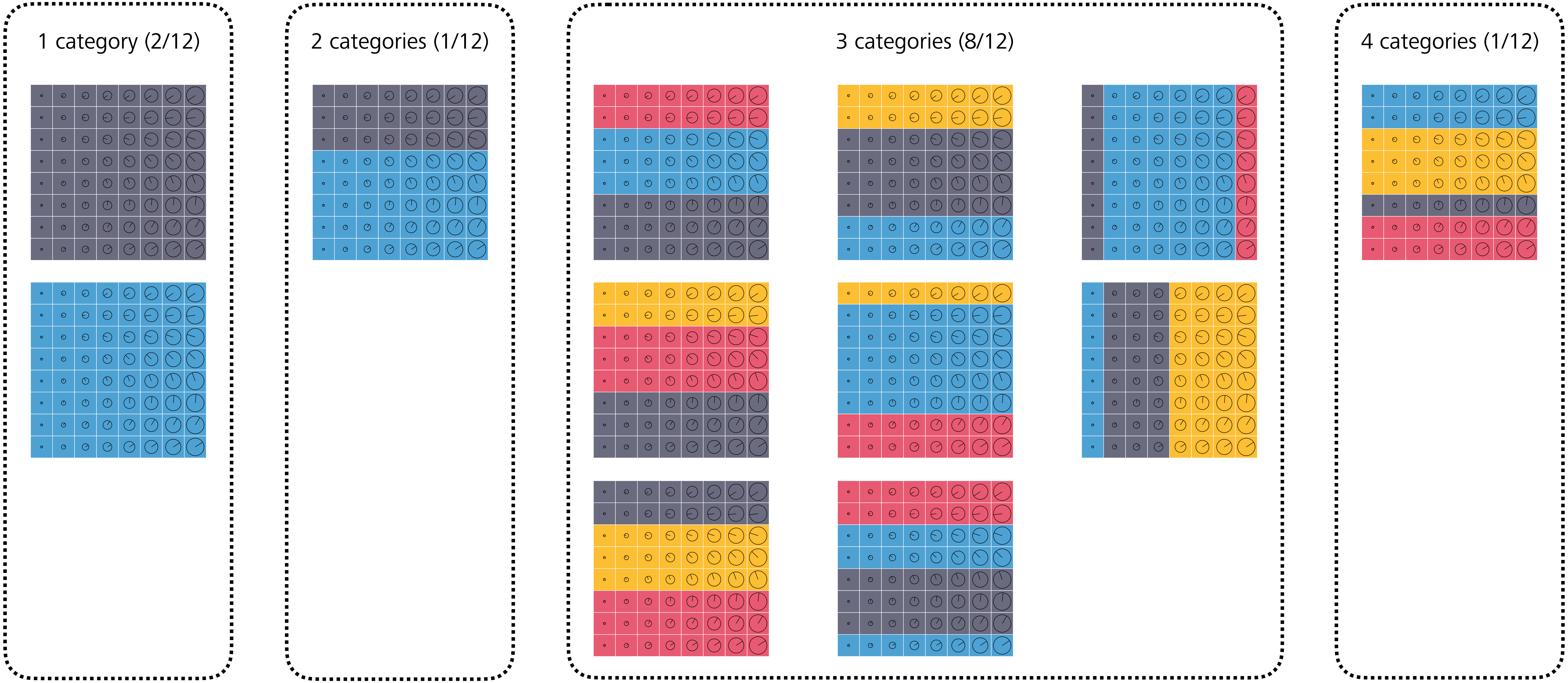
Iterated learning with humans



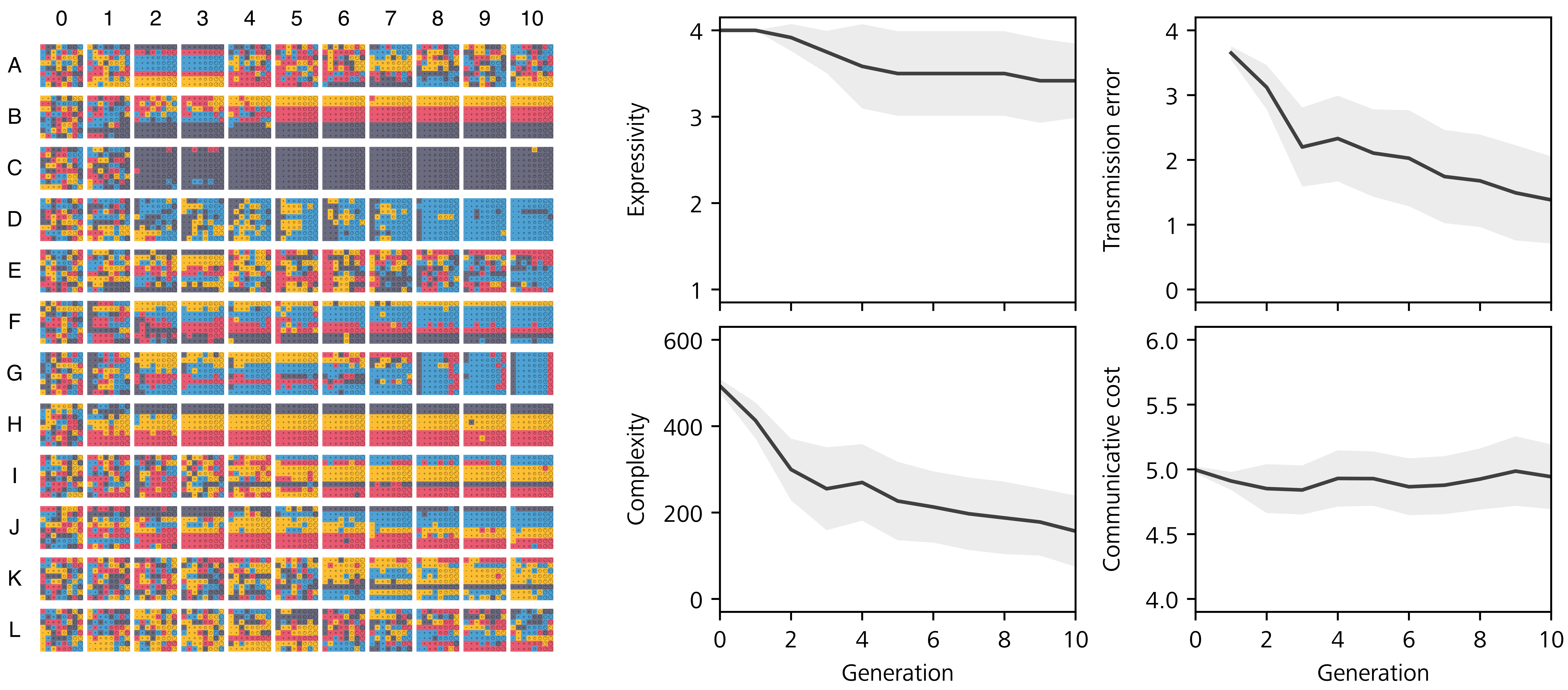




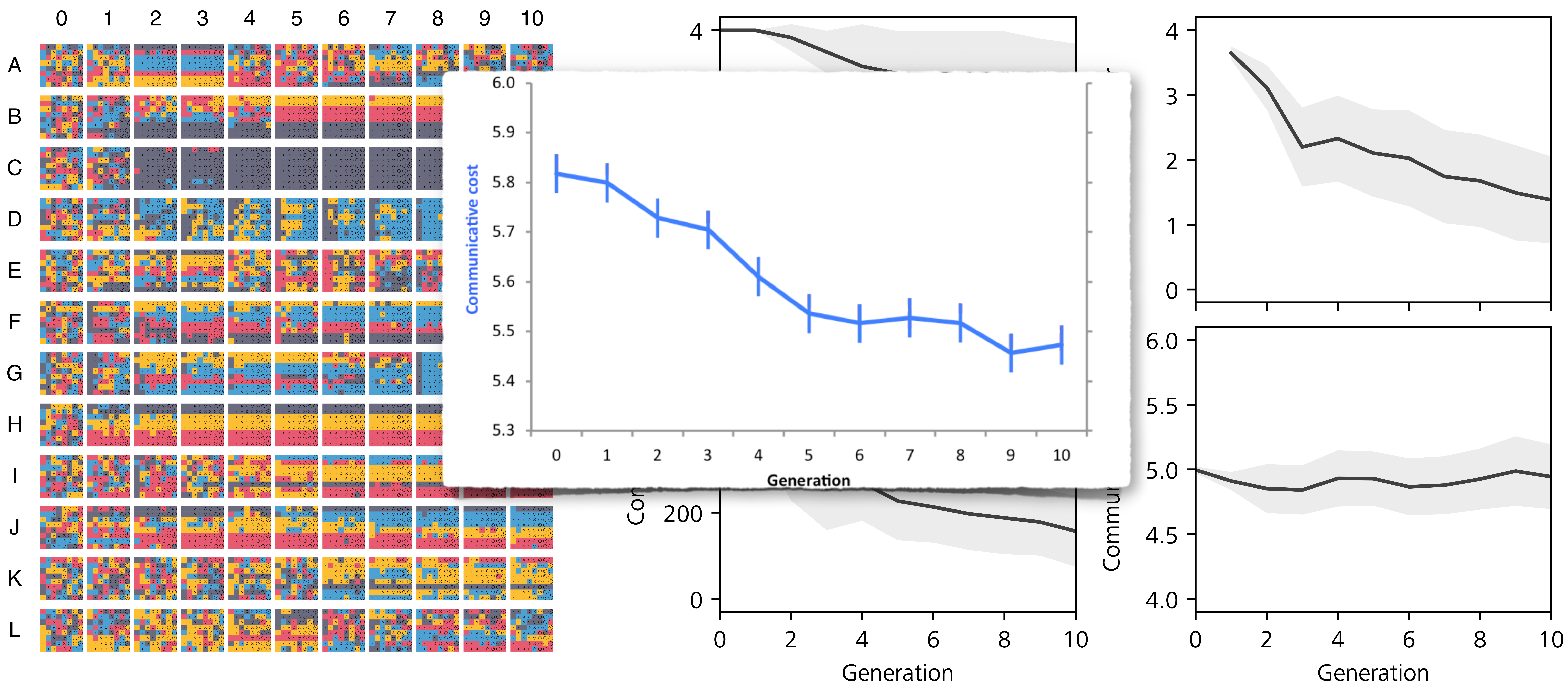
Category systems that were converged on



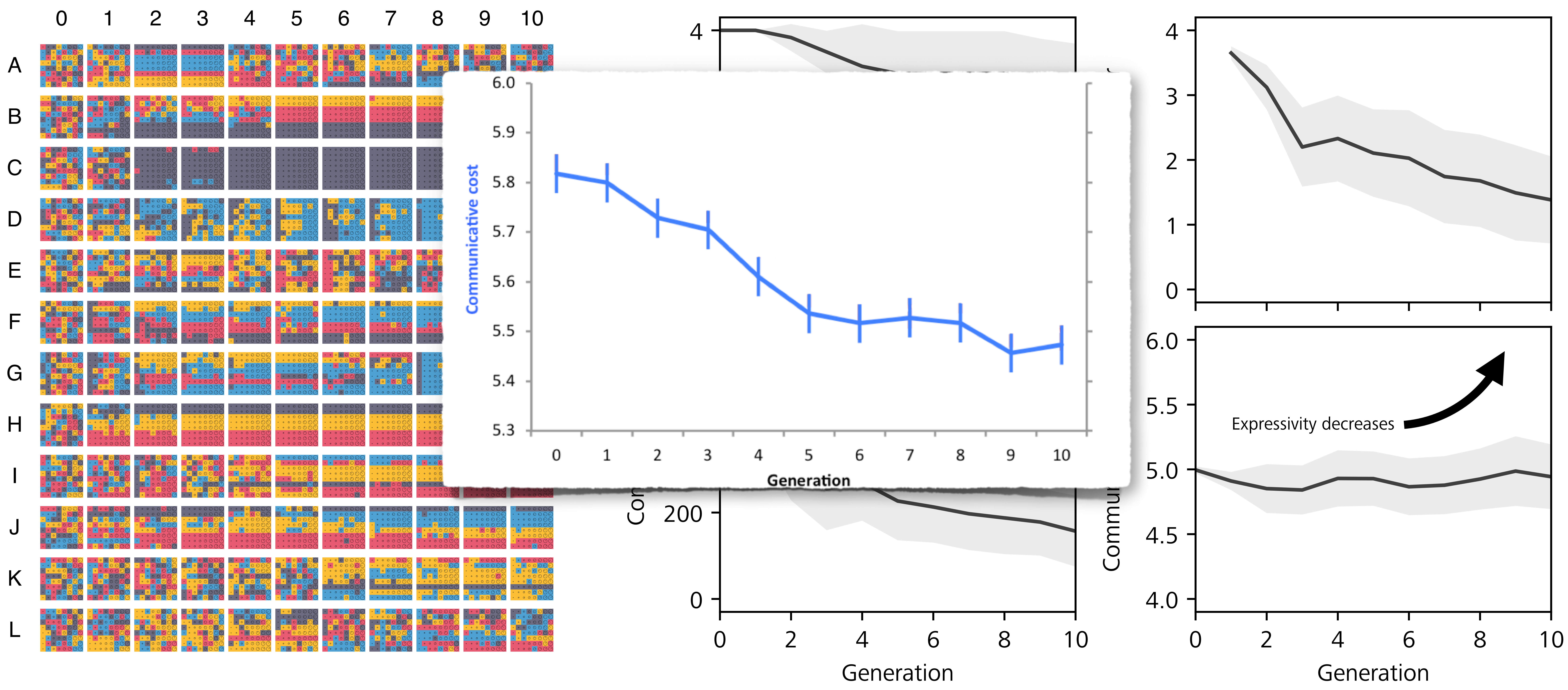
Experimental results



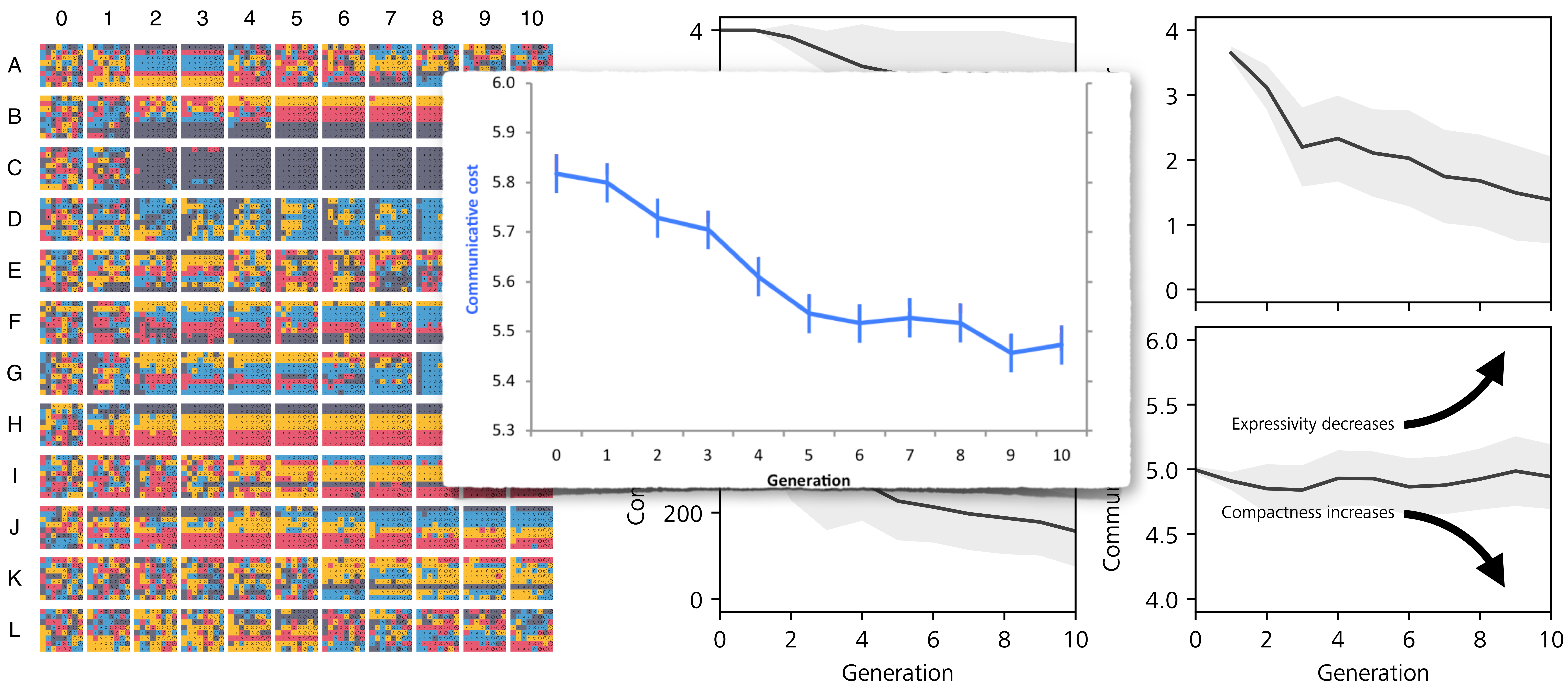
Experimental results



Experimental results



Experimental results

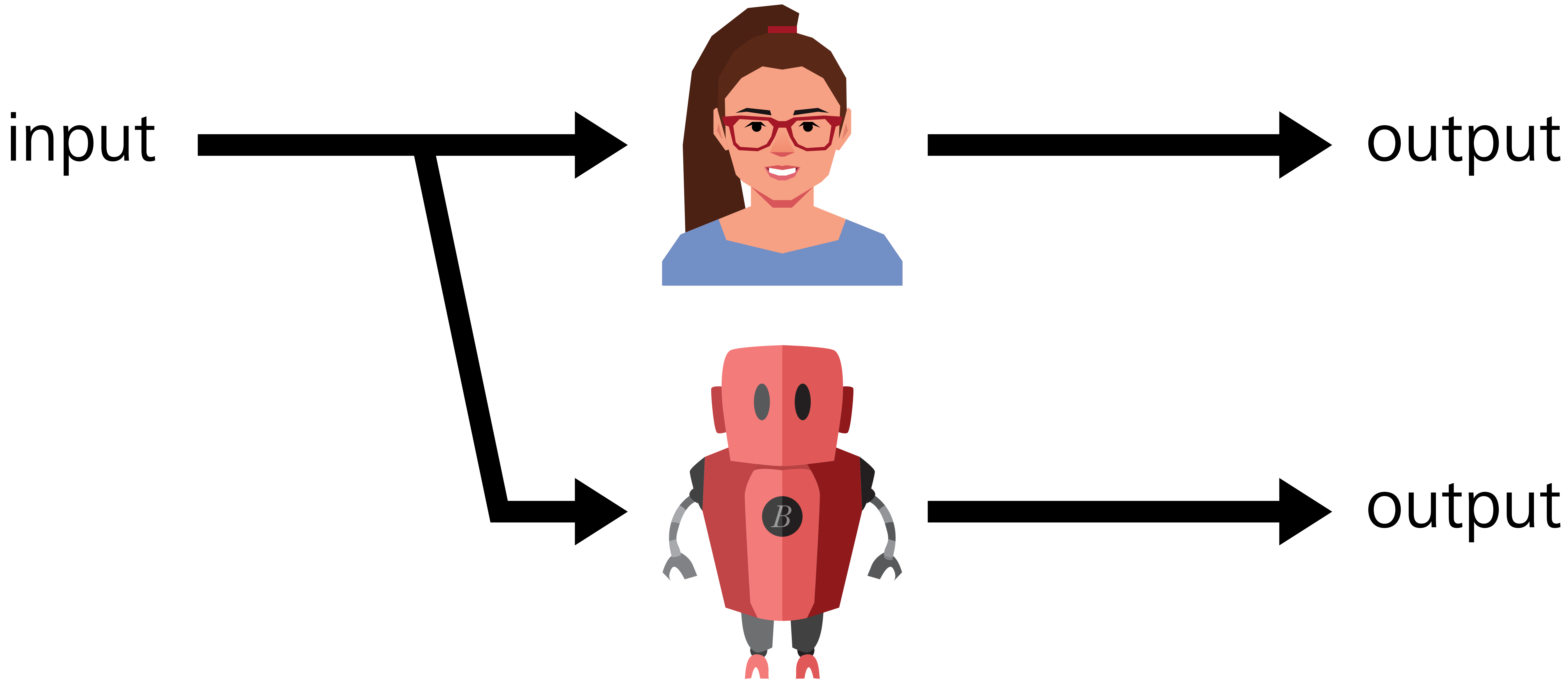


Model fit

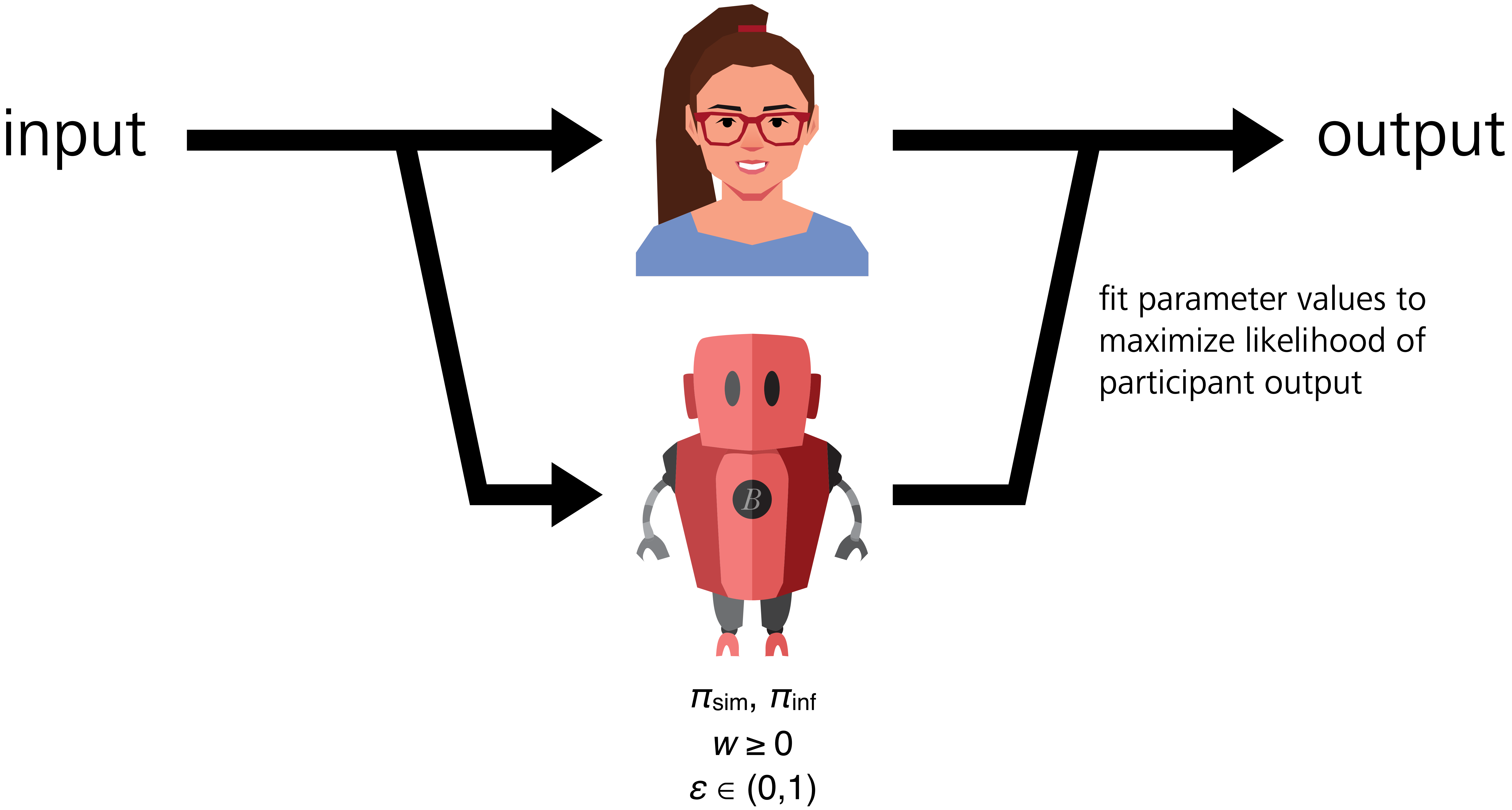
Estimating unknown parameters of the model

input →  → output

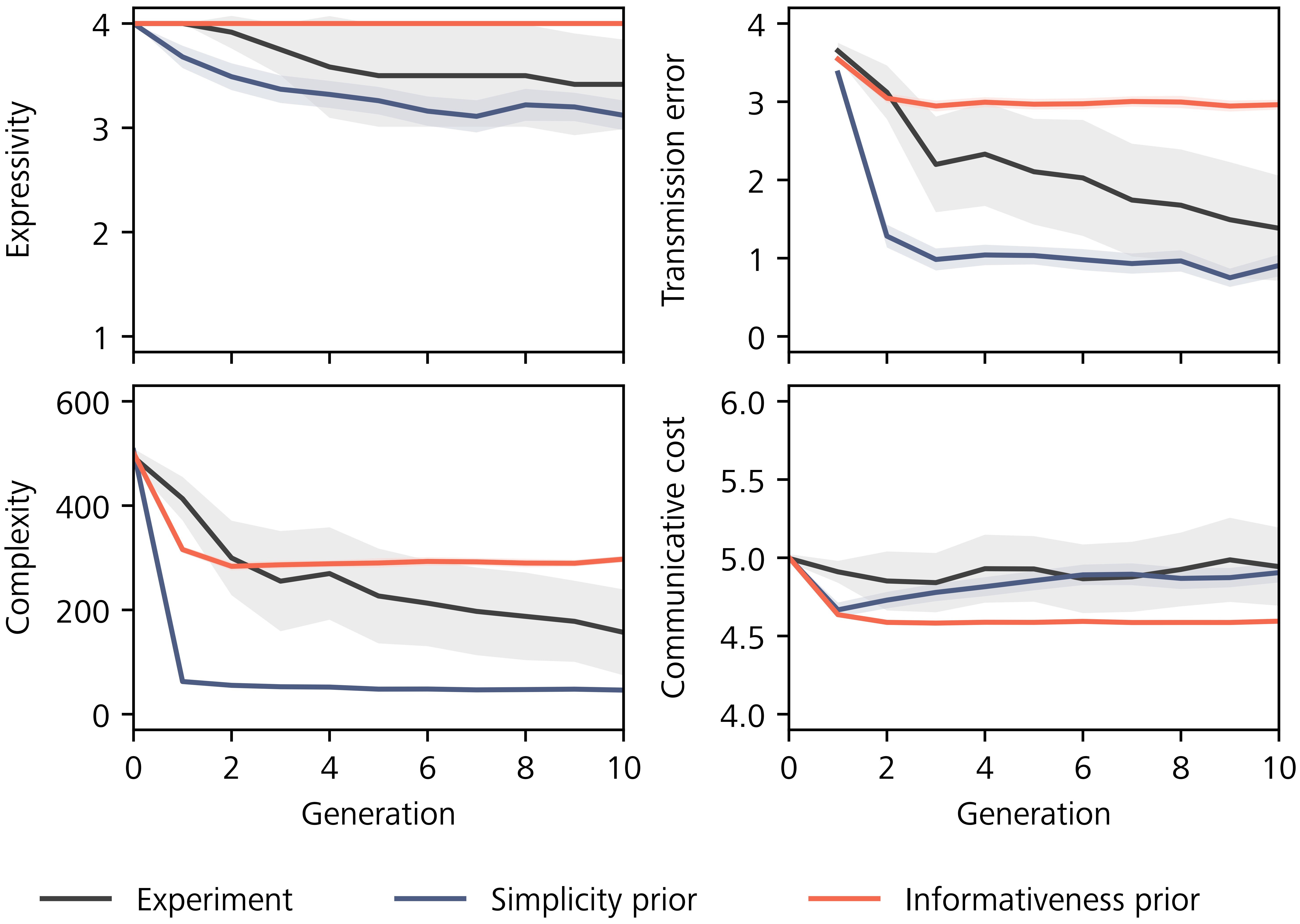
Estimating unknown parameters of the model



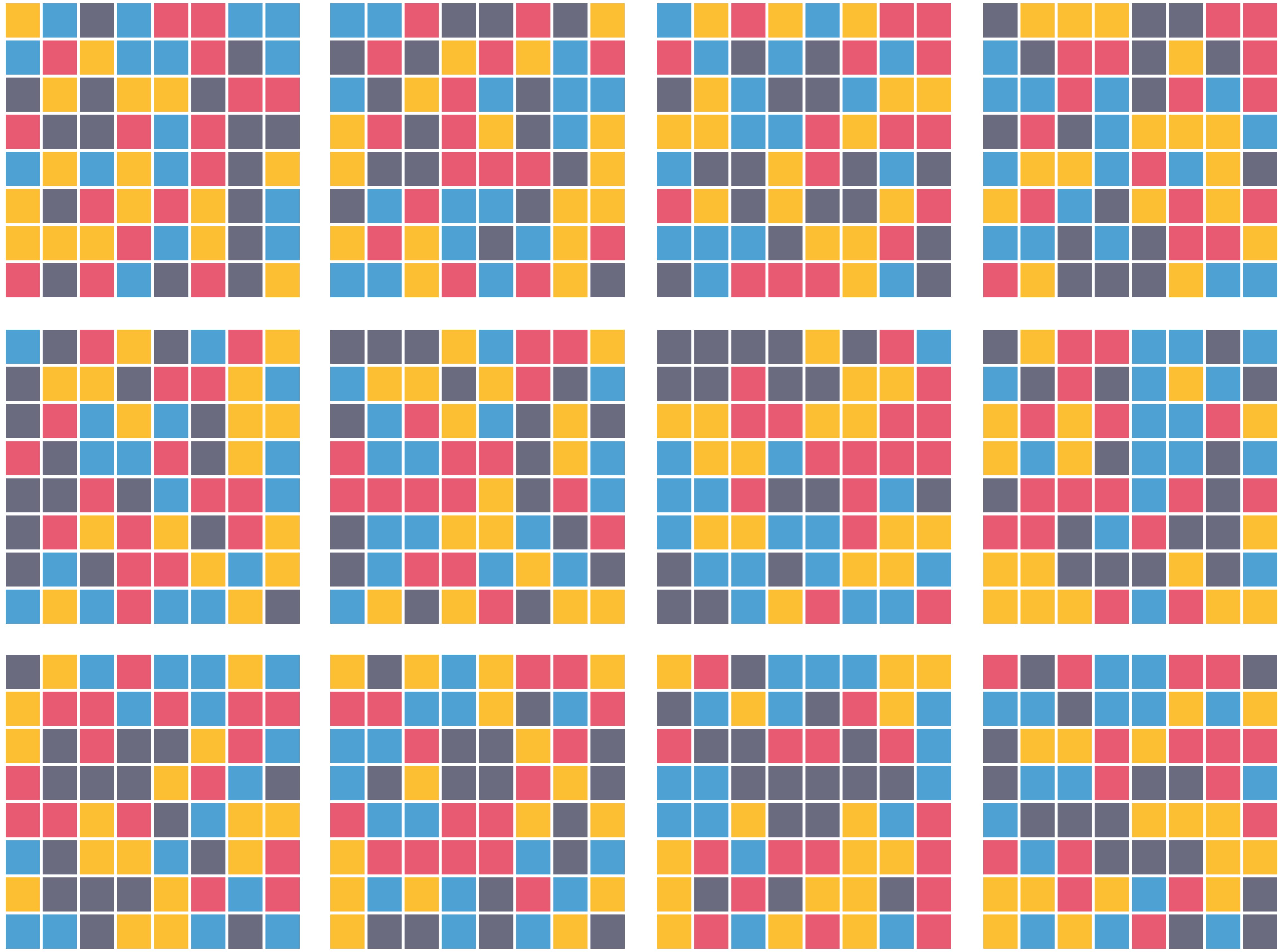
Estimating unknown parameters of the model



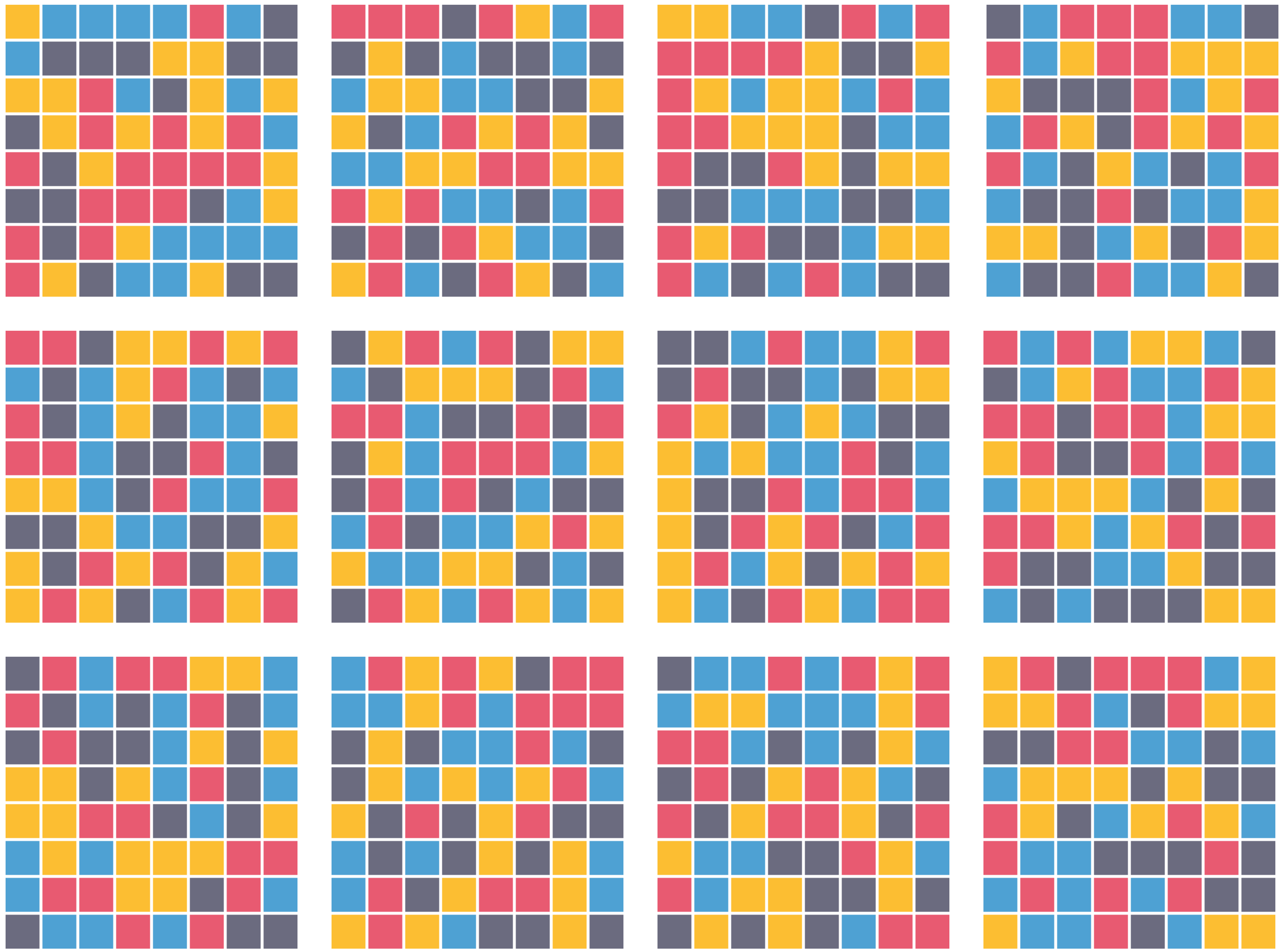
Model results with best-fit parameters



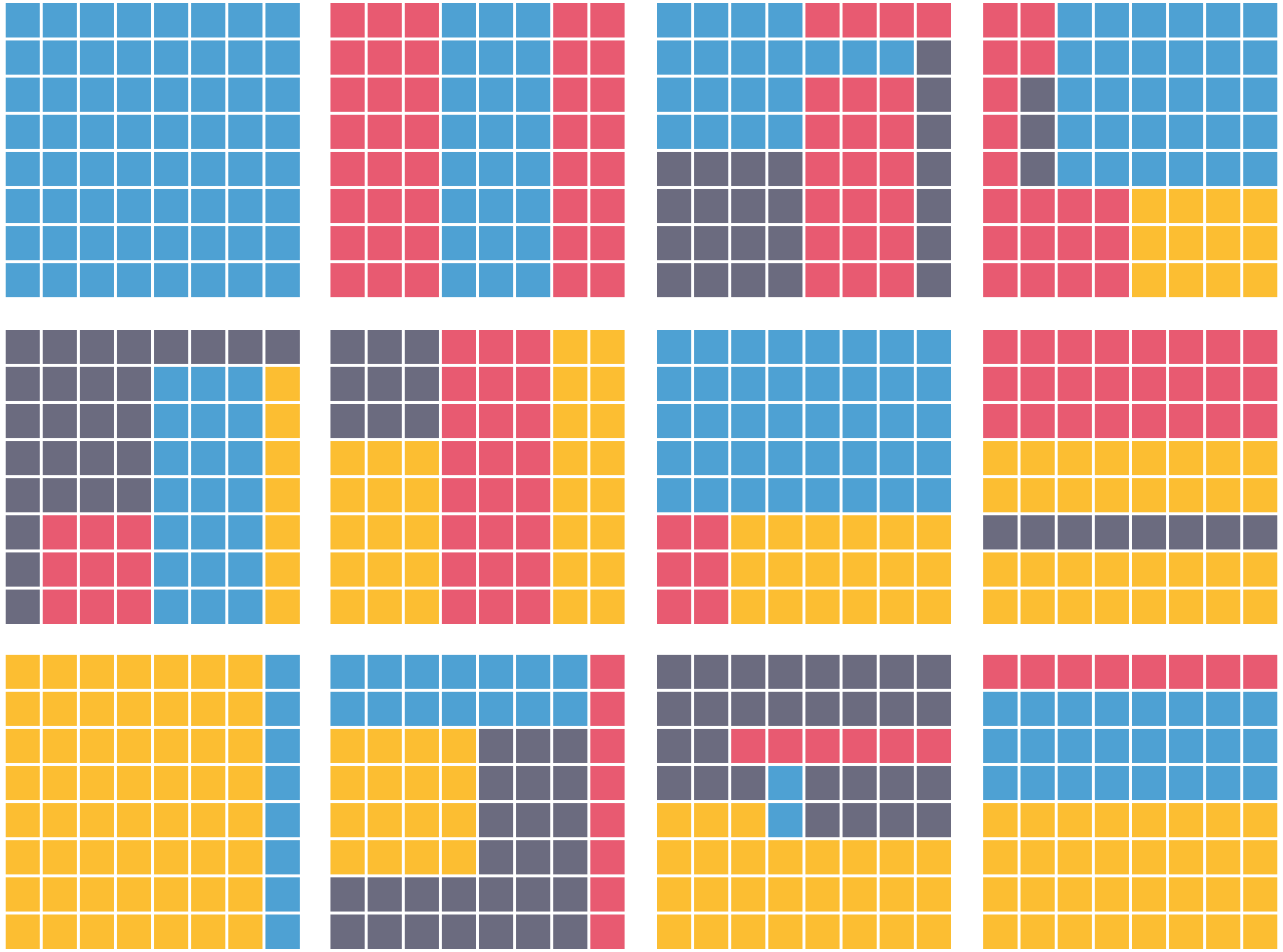
Simplicity prior



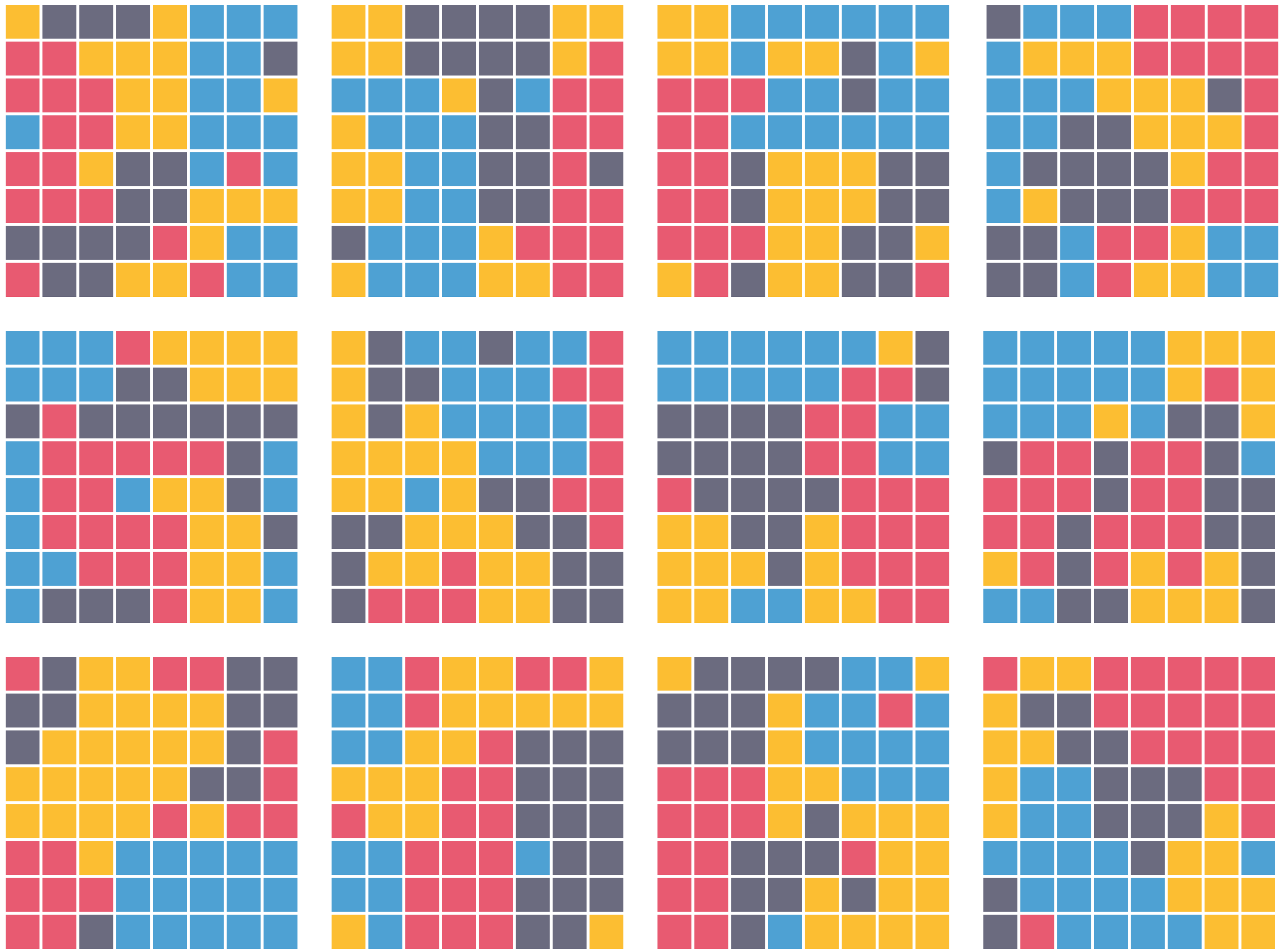
Informativeness prior



Simplicity prior



Informativeness prior

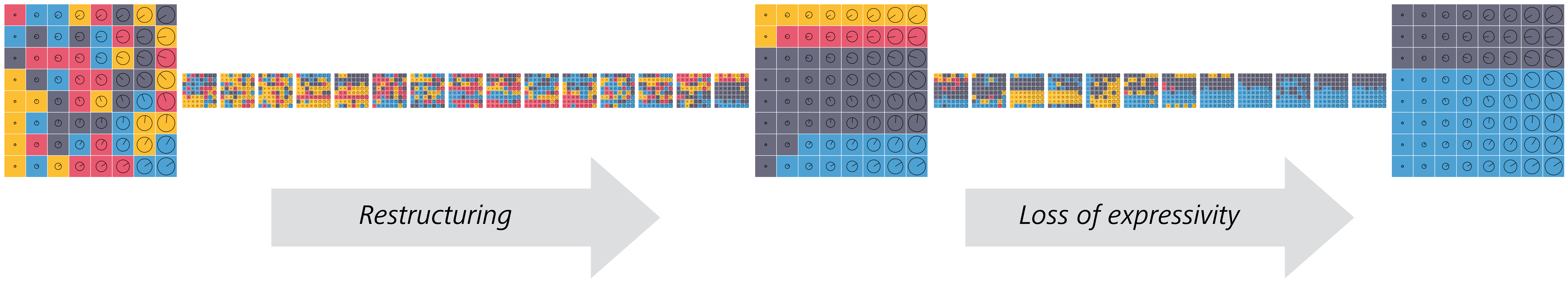


Conclusions

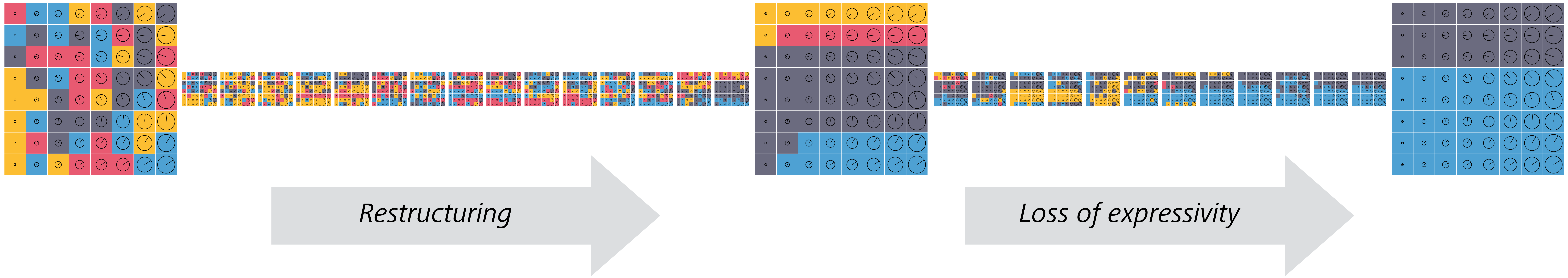
Languages are shaped under the simplicity–informativeness tradeoff by pressures from induction and interaction

A rational learner with no prior expectations ought to apply Occam's razor to domain-general problems of induction

Iterated learning results in simple categorization systems through two mechanisms

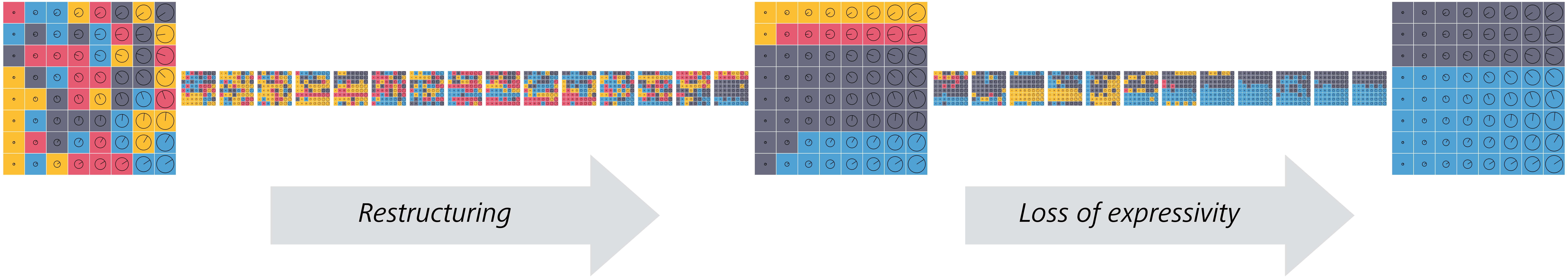


Iterated learning results in simple categorization systems through two mechanisms



Iterated learning can give rise to informative languages without positing a bias for informativeness

Iterated learning results in simple categorization systems through two mechanisms



Iterated learning can give rise to informative languages without positing a bias for informativeness

But! Unconstrained, iterated learning results in degenerate languages, so there's still a role for interaction

Thank you!

Take-home messages

Languages are shaped under the simplicity–informativeness tradeoff by pressures from induction and interaction

A rational learner with no prior expectations ought to apply a simplicity principle to domain-general problems of induction

Iterated learning (repeated induction) results in simple categorization systems through two mechanisms:

> **Compact categories:** Restructuring of the space ⇒ more informative
>
> **Loss of expressivity:** Loss of words/concepts ⇒ less informative

Iterated learning can give rise to informative(-ish) categories without actually positing a bias for informativeness; the languages are actually evolving to become simpler

Nevertheless, interactional dynamics restrain languages from total degeneration