Simplicity and informativeness in conceptual structure

Jon W. Carr

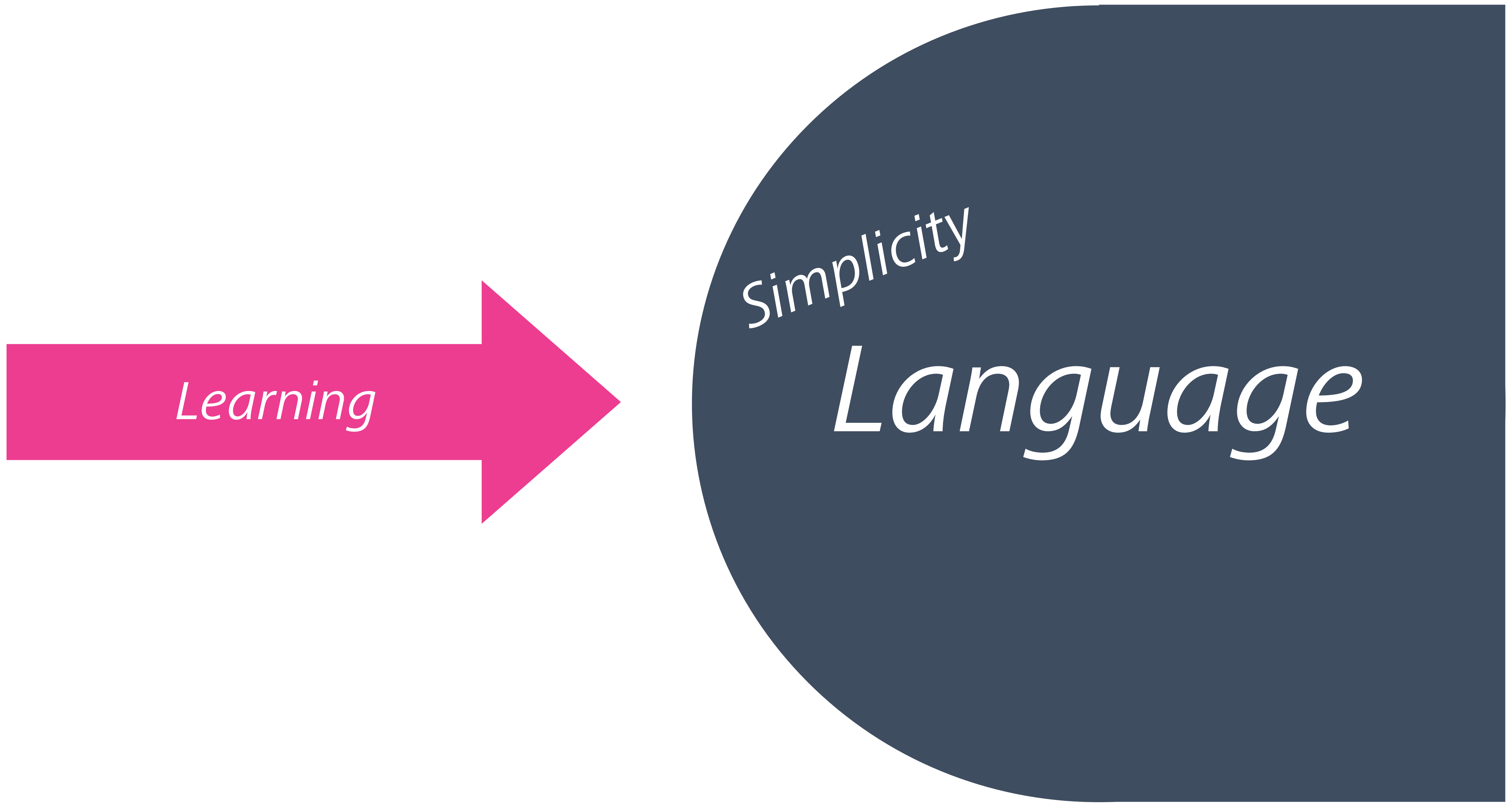
Centre for Language Evolution
School of Philosophy, Psychology and Language Sciences
University of Edinburgh



Pressures shaping language

Language

Pressures shaping language

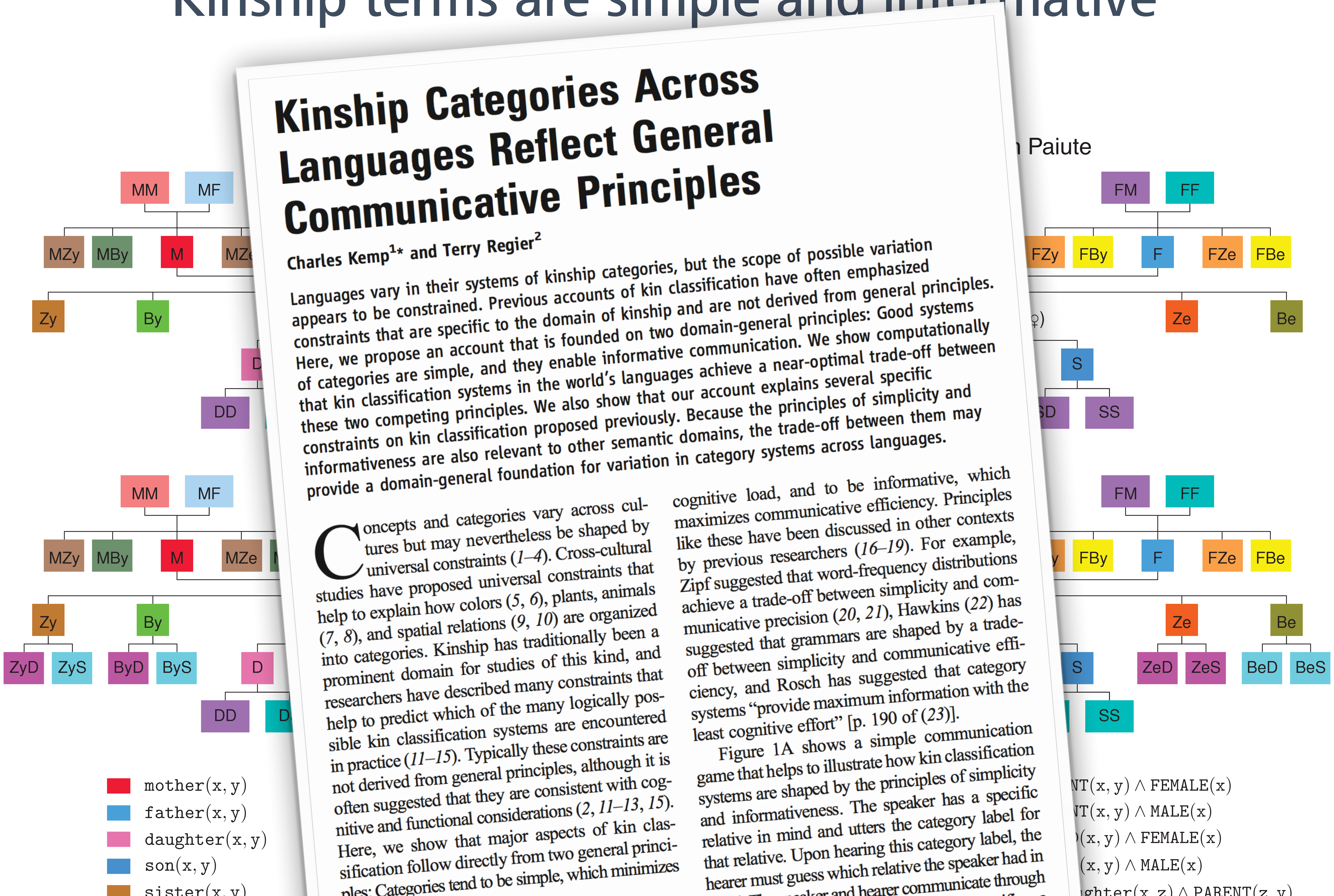


Pressures shaping language

Pressures shaping language

The simplicity–informativeness tradeoff

Kinship terms are simple and informative



Kinship Categories Across Languages Reflect General Communicative Principles

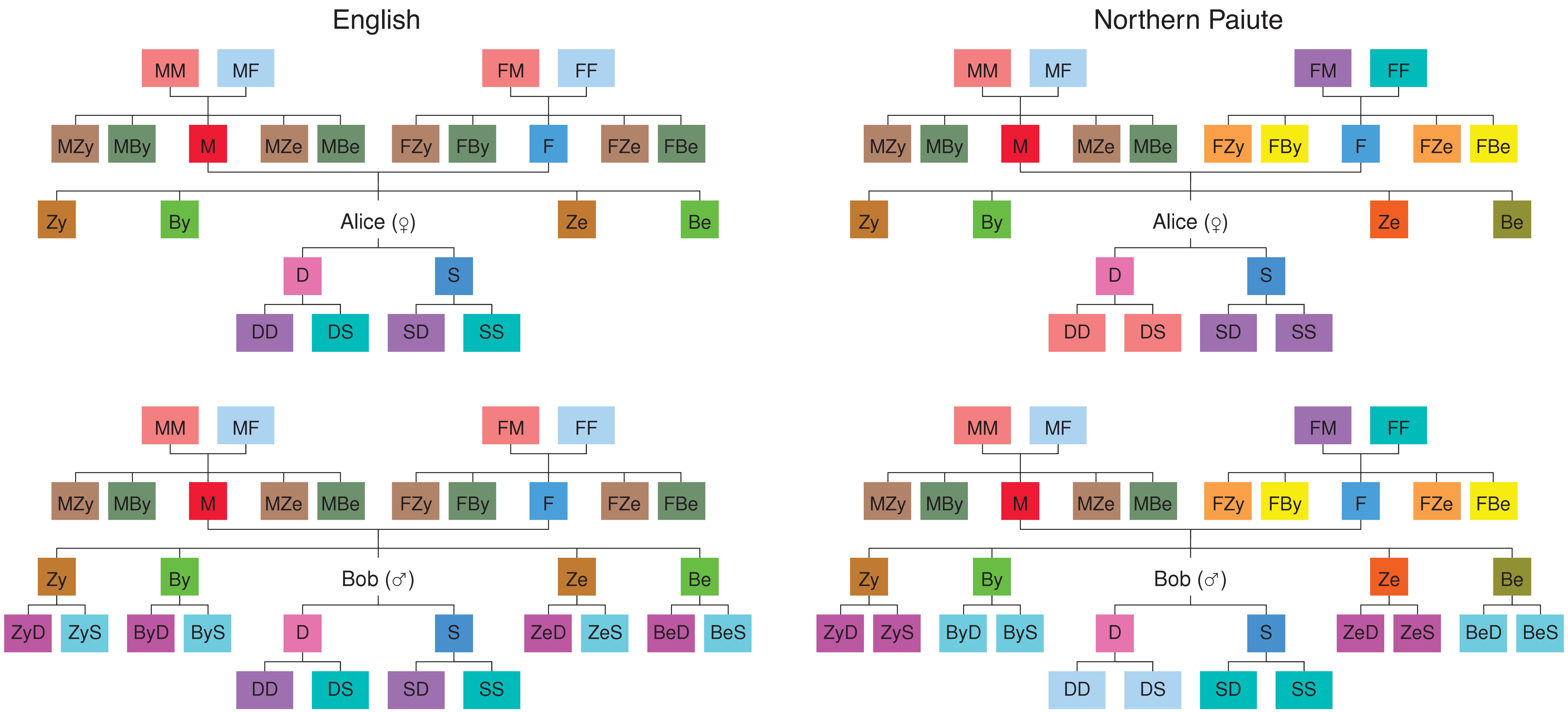
Charles Kemp[1*] and Terry Regier[2]

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

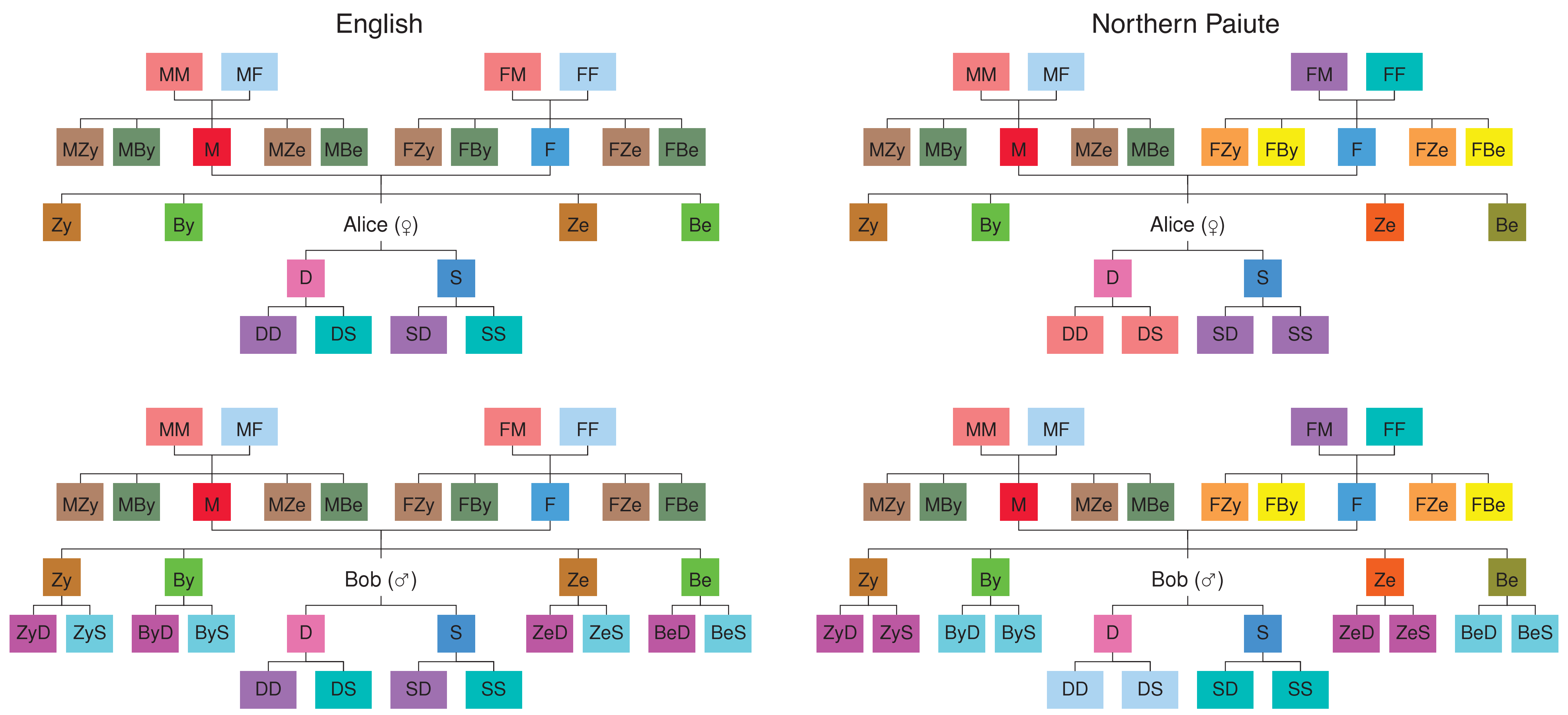
Kinship terms are simple and informative



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)

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



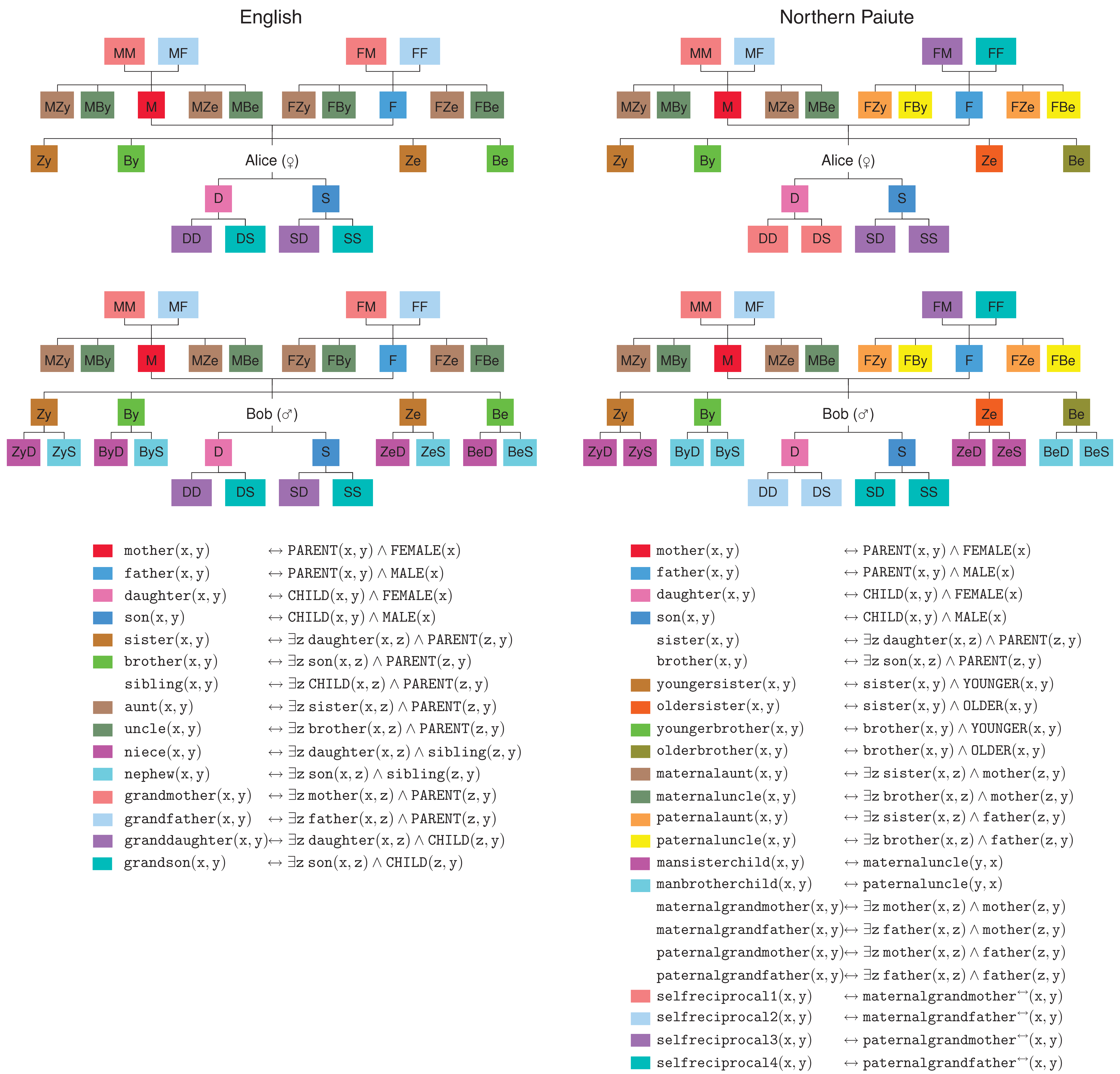
English

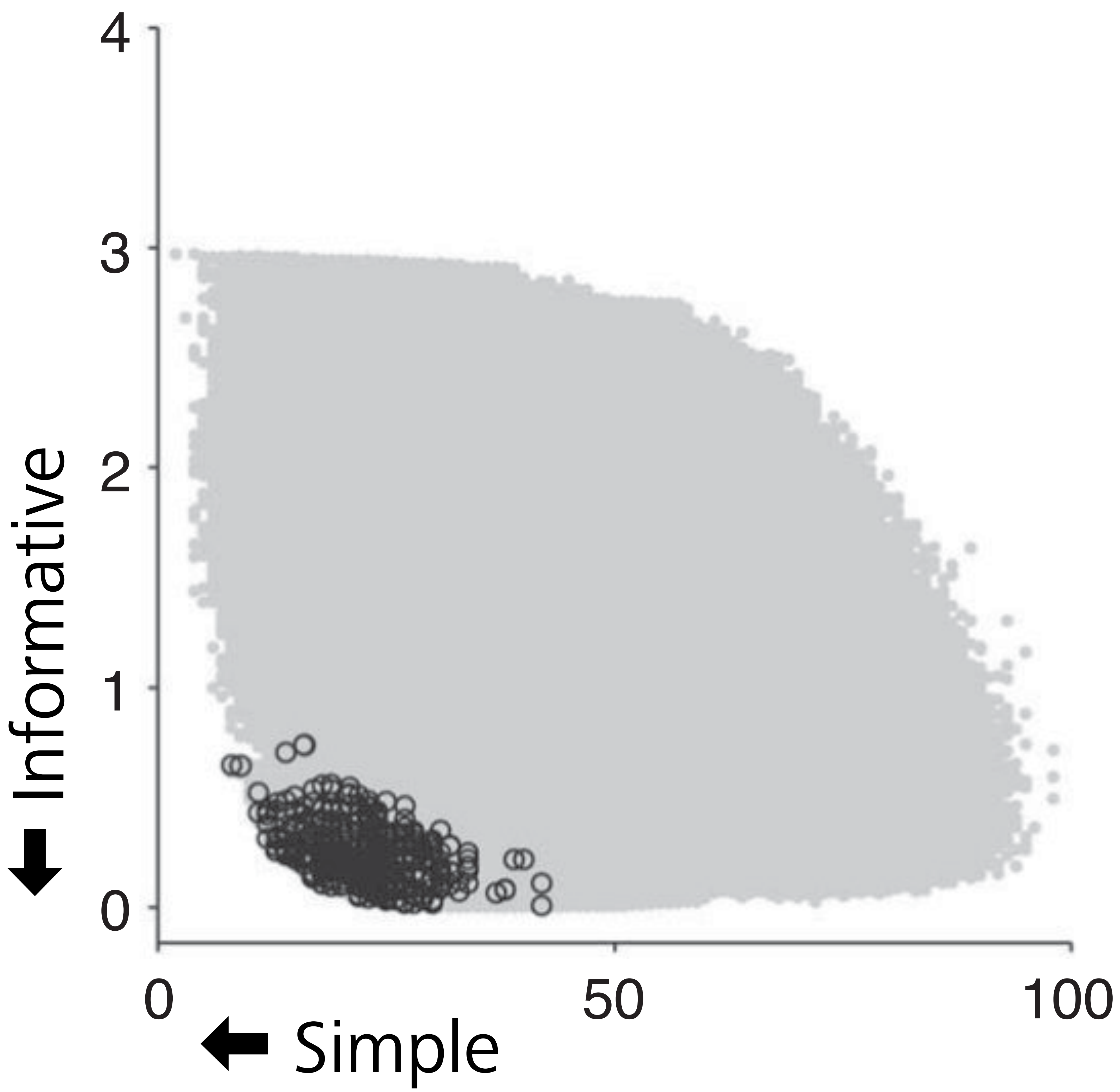
$\text{mother}(x,y)$	$\leftrightarrow \text{PARENT}(x,y) \wedge \text{FEMALE}(x)$
$\text{father}(x,y)$	$\leftrightarrow \text{PARENT}(x,y) \wedge \text{MALE}(x)$
$\text{daughter}(x,y)$	$\leftrightarrow \text{CHILD}(x,y) \wedge \text{FEMALE}(x)$
$\text{son}(x,y)$	$\leftrightarrow \text{CHILD}(x,y) \wedge \text{MALE}(x)$
$\text{sister}(x,y)$	$\leftrightarrow \exists z\, \text{daughter}(x,z) \wedge \text{PARENT}(z,y)$
$\text{brother}(x,y)$	$\leftrightarrow \exists z\, \text{son}(x,z) \wedge \text{PARENT}(z,y)$
$\text{sibling}(x,y)$	$\leftrightarrow \exists z\, \text{CHILD}(x,z) \wedge \text{PARENT}(z,y)$
$\text{aunt}(x,y)$	$\leftrightarrow \exists z\, \text{sister}(x,z) \wedge \text{PARENT}(z,y)$
$\text{uncle}(x,y)$	$\leftrightarrow \exists z\, \text{brother}(x,z) \wedge \text{PARENT}(z,y)$
$\text{niece}(x,y)$	$\leftrightarrow \exists z\, \text{daughter}(x,z) \wedge \text{sibling}(z,y)$
$\text{nephew}(x,y)$	$\leftrightarrow \exists z\, \text{son}(x,z) \wedge \text{sibling}(z,y)$
$\text{grandmother}(x,y)$	$\leftrightarrow \exists z\, \text{mother}(x,z) \wedge \text{PARENT}(z,y)$
$\text{grandfather}(x,y)$	$\leftrightarrow \exists z\, \text{father}(x,z) \wedge \text{PARENT}(z,y)$
$\text{granddaughter}(x,y)$	$\leftrightarrow \exists z\, \text{daughter}(x,z) \wedge \text{CHILD}(z,y)$
$\text{grandson}(x,y)$	$\leftrightarrow \exists z\, \text{son}(x,z) \wedge \text{CHILD}(z,y)$

Northern Paiute

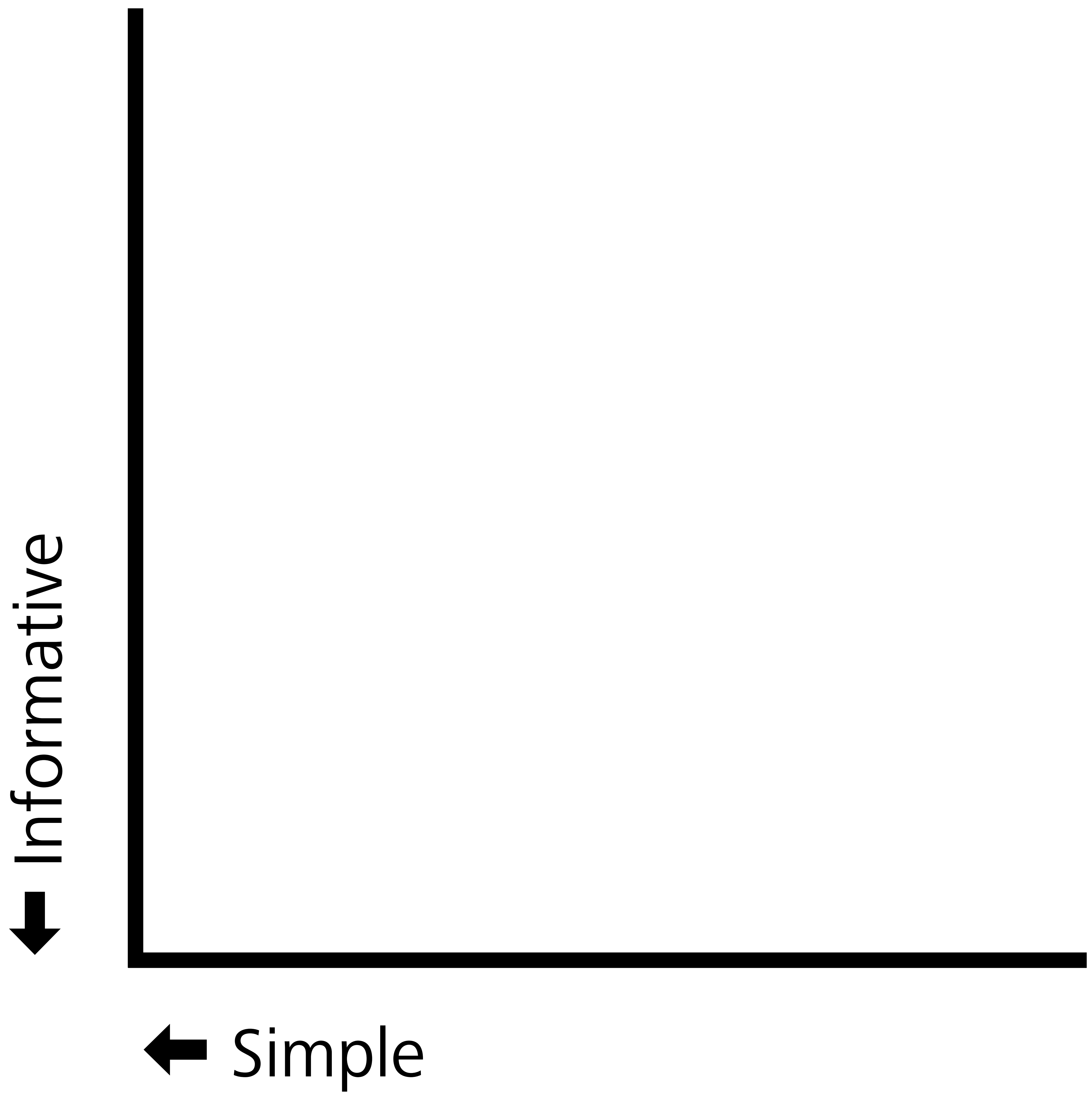
$\text{mother}(x,y)$	$\leftrightarrow \text{PARENT}(x,y) \wedge \text{FEMALE}(x)$
$\text{father}(x,y)$	$\leftrightarrow \text{PARENT}(x,y) \wedge \text{MALE}(x)$
$\text{daughter}(x,y)$	$\leftrightarrow \text{CHILD}(x,y) \wedge \text{FEMALE}(x)$
$\text{son}(x,y)$	$\leftrightarrow \text{CHILD}(x,y) \wedge \text{MALE}(x)$
$\text{sister}(x,y)$	$\leftrightarrow \exists z\, \text{daughter}(x,z) \wedge \text{PARENT}(z,y)$
$\text{brother}(x,y)$	$\leftrightarrow \exists z\, \text{son}(x,z) \wedge \text{PARENT}(z,y)$
$\text{youngersister}(x,y)$	$\leftrightarrow \text{sister}(x,y) \wedge \text{YOUNGER}(x,y)$
$\text{oldersister}(x,y)$	$\leftrightarrow \text{sister}(x,y) \wedge \text{OLDER}(x,y)$
$\text{youngerbrother}(x,y)$	$\leftrightarrow \text{brother}(x,y) \wedge \text{YOUNGER}(x,y)$
$\text{olderbrother}(x,y)$	$\leftrightarrow \text{brother}(x,y) \wedge \text{OLDER}(x,y)$
$\text{maternalaunt}(x,y)$	$\leftrightarrow \exists z\, \text{sister}(x,z) \wedge \text{mother}(z,y)$
$\text{maternaluncle}(x,y)$	$\leftrightarrow \exists z\, \text{brother}(x,z) \wedge \text{mother}(z,y)$
$\text{paternalaunt}(x,y)$	$\leftrightarrow \exists z\, \text{sister}(x,z) \wedge \text{father}(z,y)$
$\text{paternaluncle}(x,y)$	$\leftrightarrow \exists z\, \text{brother}(x,z) \wedge \text{father}(z,y)$
$\text{mansisterchild}(x,y)$	$\leftrightarrow \text{maternaluncle}(y,x)$
$\text{manbrotherchild}(x,y)$	$\leftrightarrow \text{paternaluncle}(y,x)$
$\text{maternalgrandmother}(x,y)$	$\leftrightarrow \exists z\, \text{mother}(x,z) \wedge \text{mother}(z,y)$
$\text{maternalgrandfather}(x,y)$	$\leftrightarrow \exists z\, \text{father}(x,z) \wedge \text{mother}(z,y)$
$\text{paternalgrandmother}(x,y)$	$\leftrightarrow \exists z\, \text{mother}(x,z) \wedge \text{father}(z,y)$
$\text{paternalgrandfather}(x,y)$	$\leftrightarrow \exists z\, \text{father}(x,z) \wedge \text{father}(z,y)$
$\text{selfreciprocal1}(x,y)$	$\leftrightarrow \text{maternalgrandmother}^{\leftrightarrow}(x,y)$
$\text{selfreciprocal2}(x,y)$	$\leftrightarrow \text{maternalgrandfather}^{\leftrightarrow}(x,y)$
$\text{selfreciprocal3}(x,y)$	$\leftrightarrow \text{paternalgrandmother}^{\leftrightarrow}(x,y)$
$\text{selfreciprocal4}(x,y)$	$\leftrightarrow \text{paternalgrandfather}^{\leftrightarrow}(x,y)$

Kinship terms are simple and informative

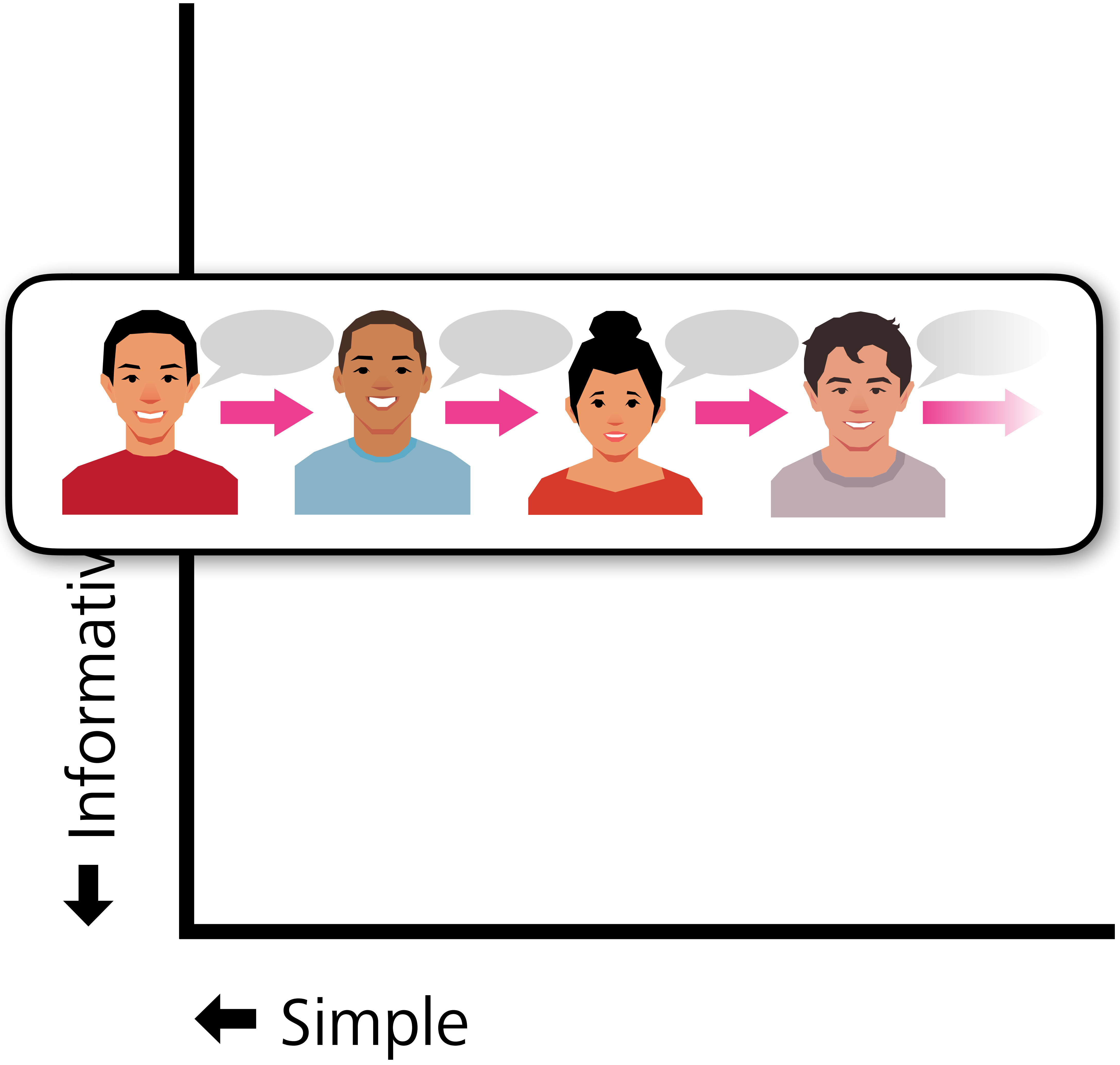




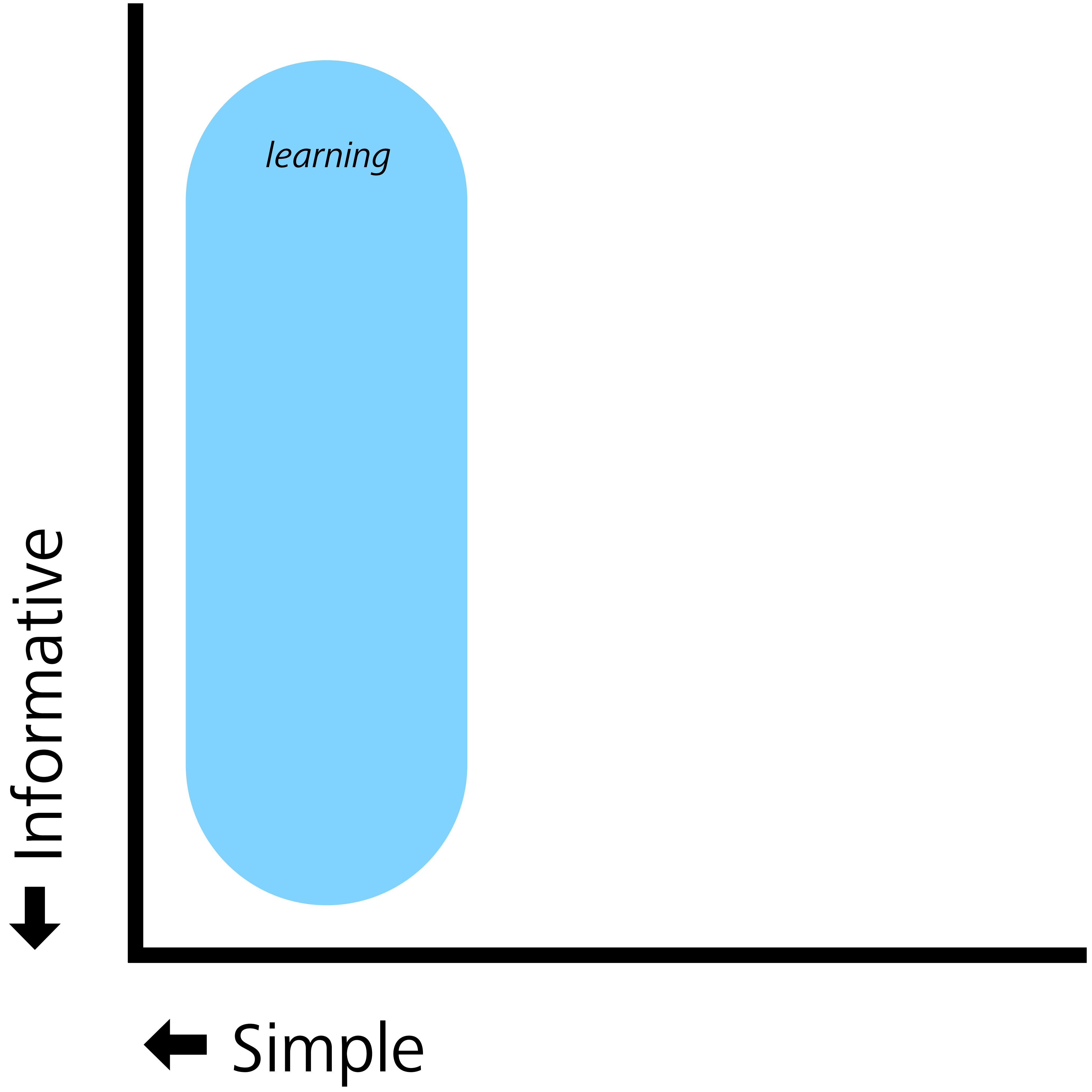
Learning and interaction pressures



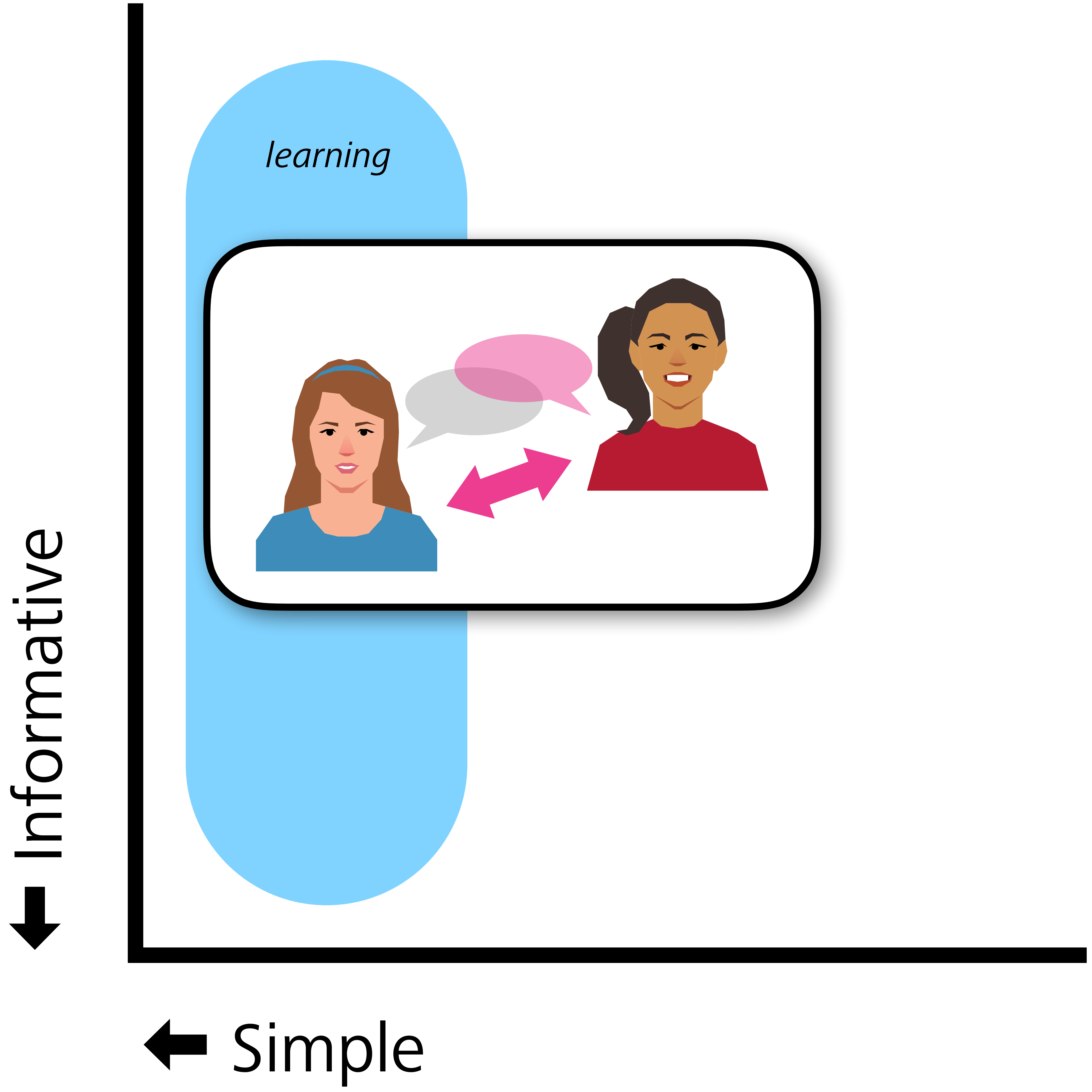
Learning and interaction pressures



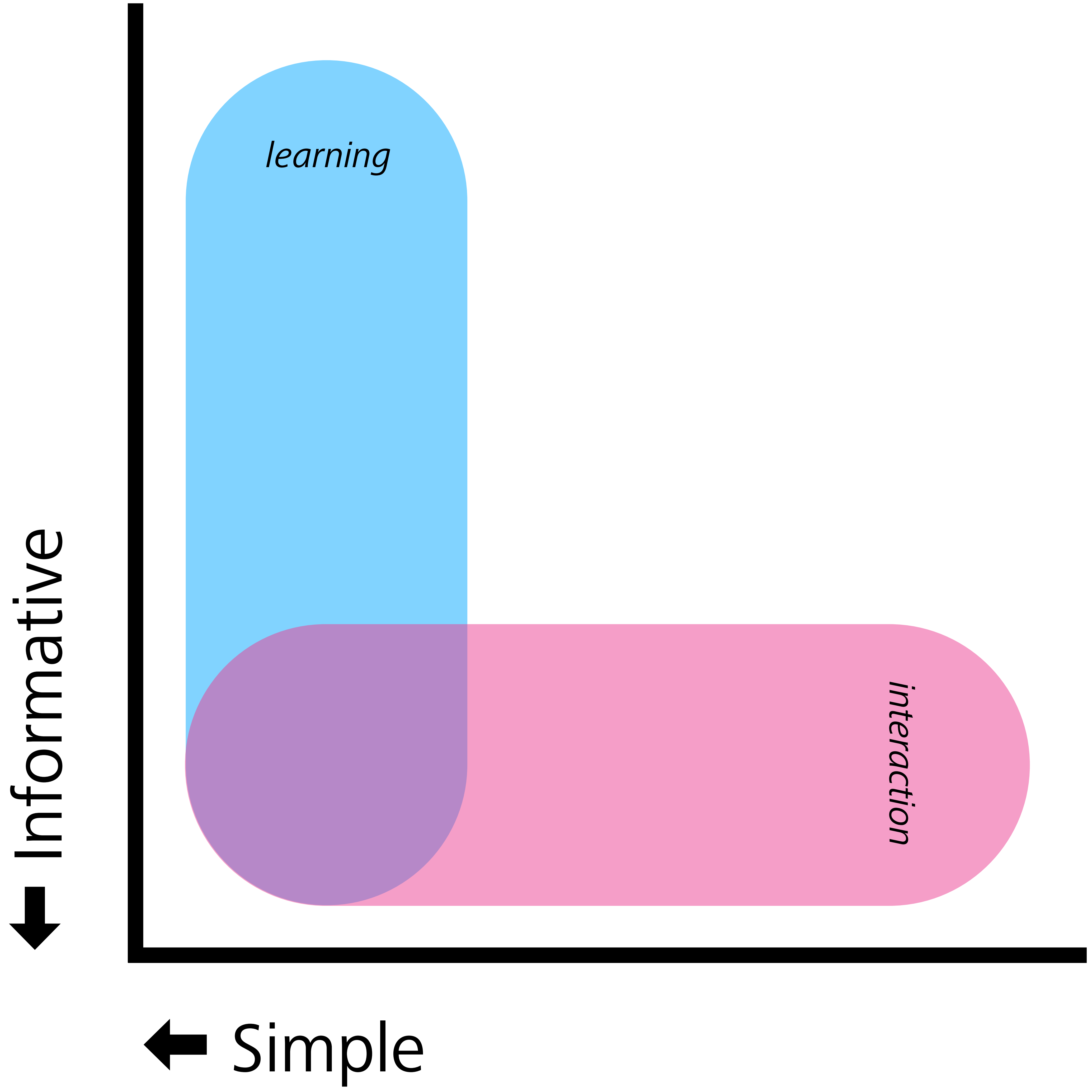
Learning and interaction pressures



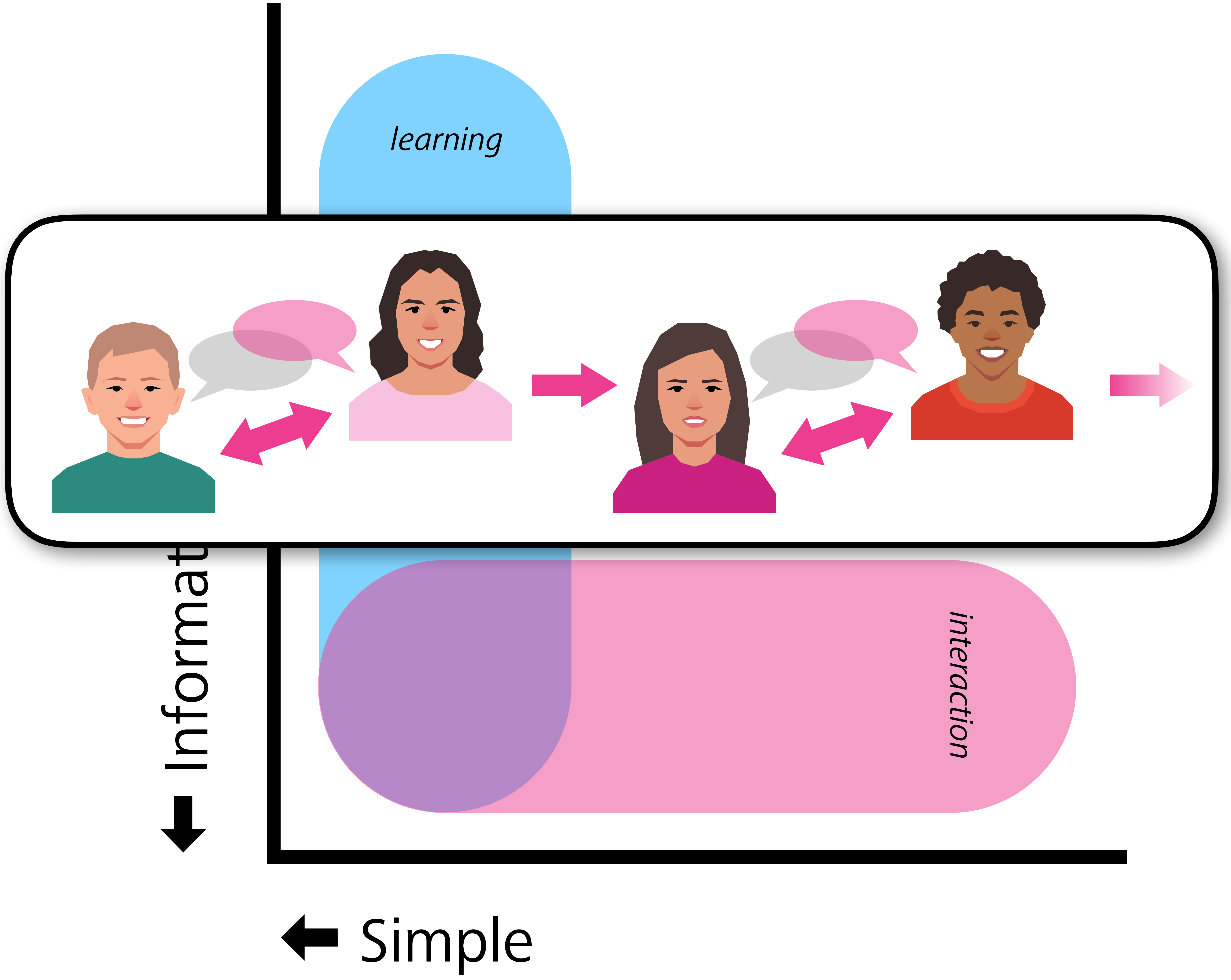
Learning and interaction pressures



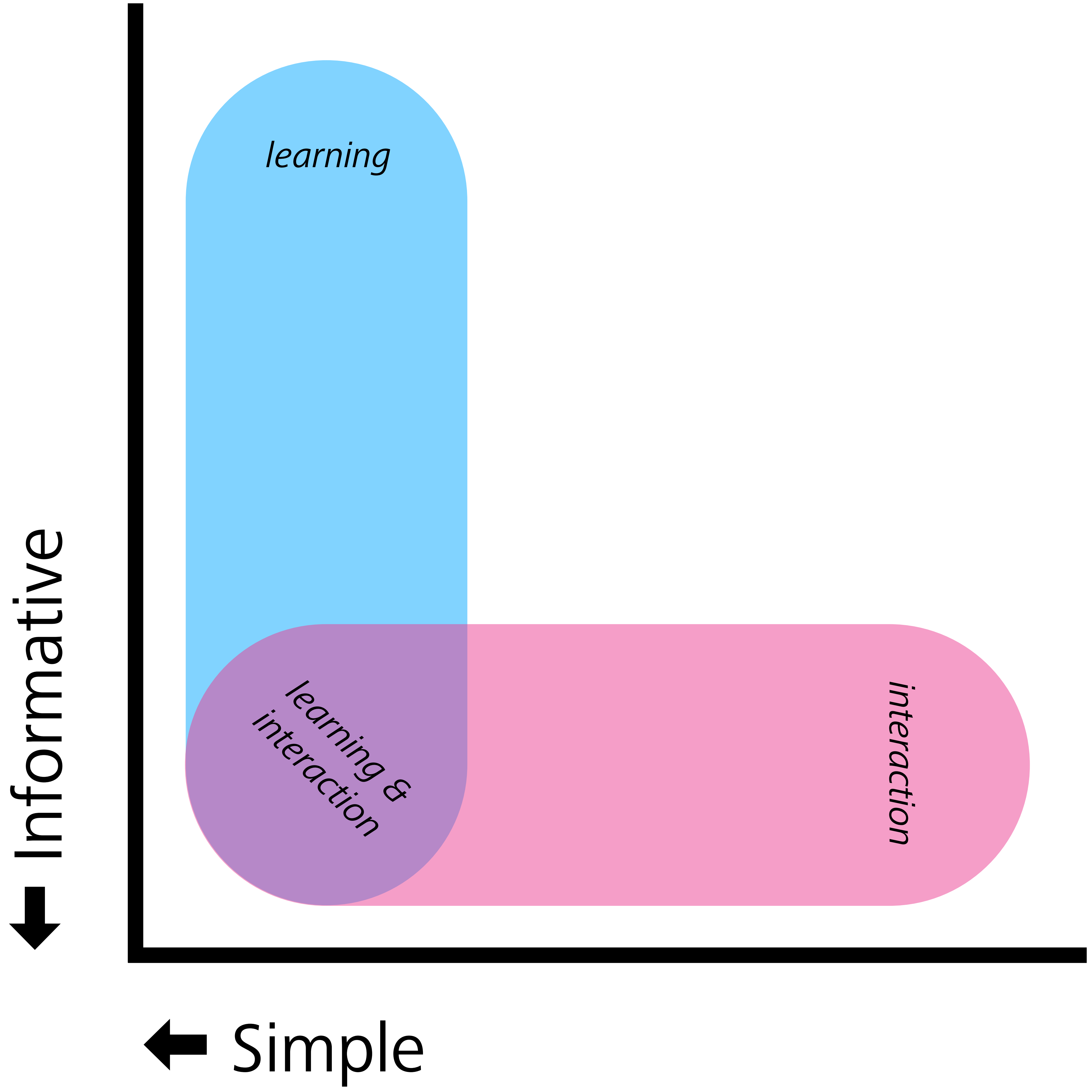
Learning and interaction pressures



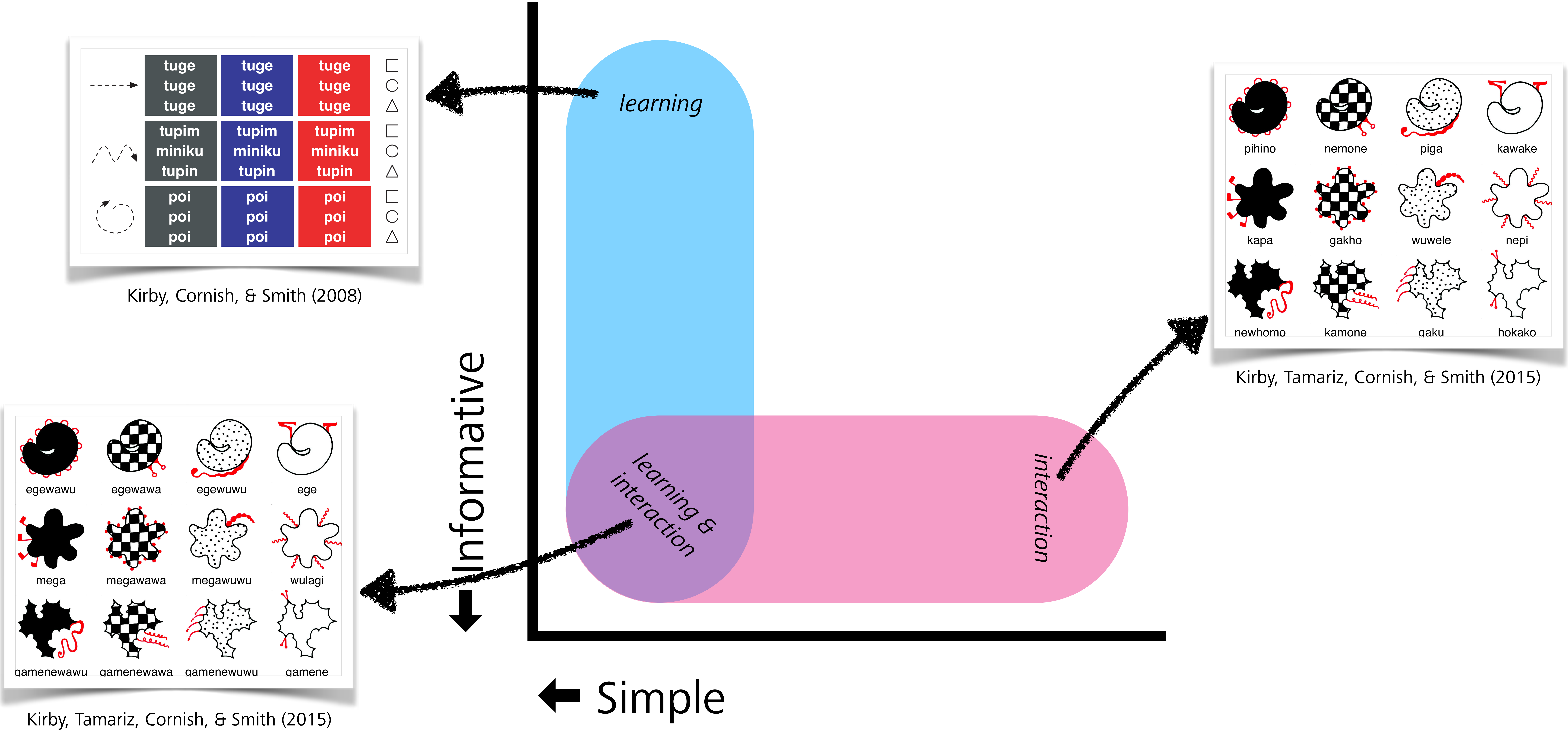
Learning and interaction pressures



Learning and interaction pressures



Learning and interaction pressures

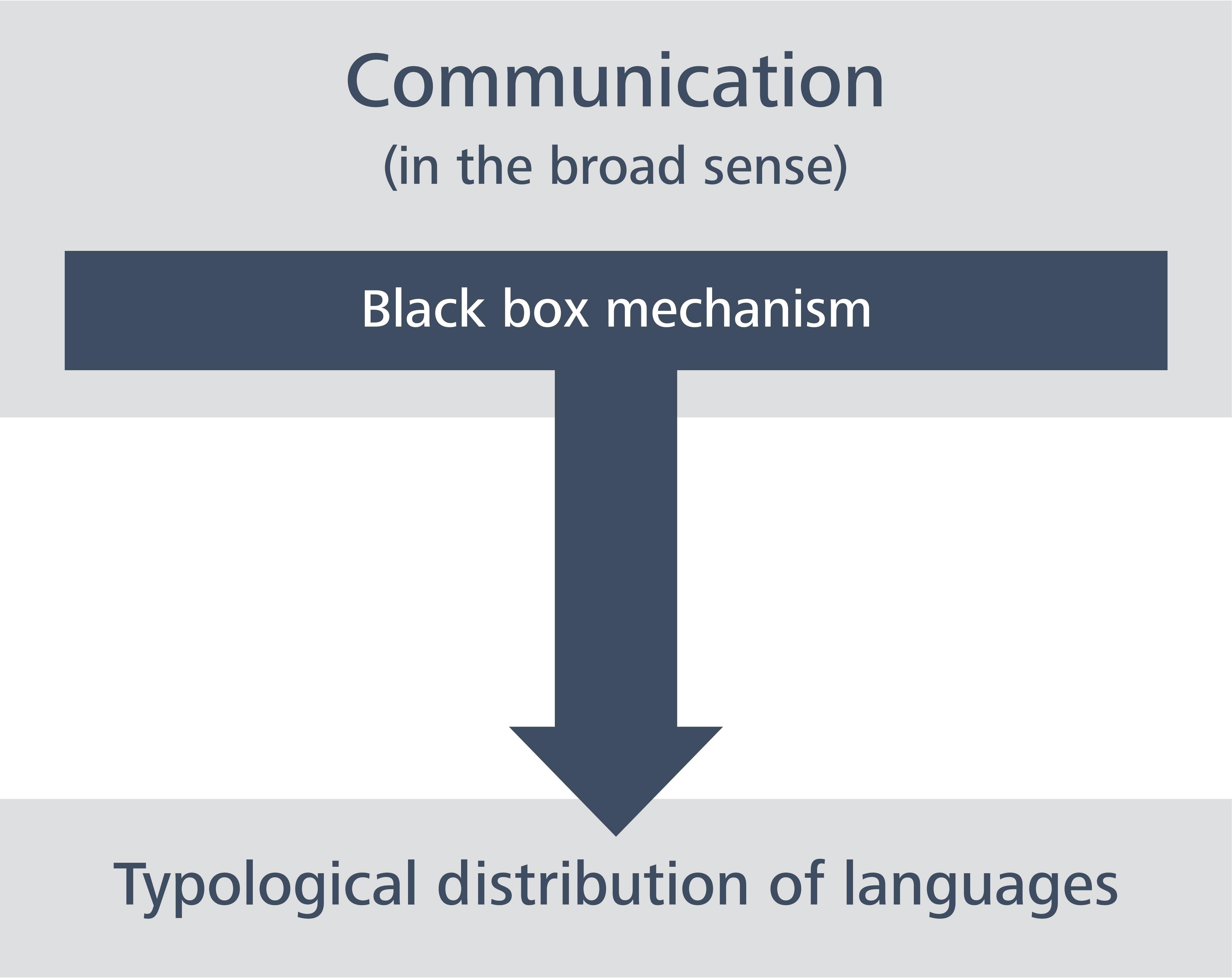


The problem of linkage

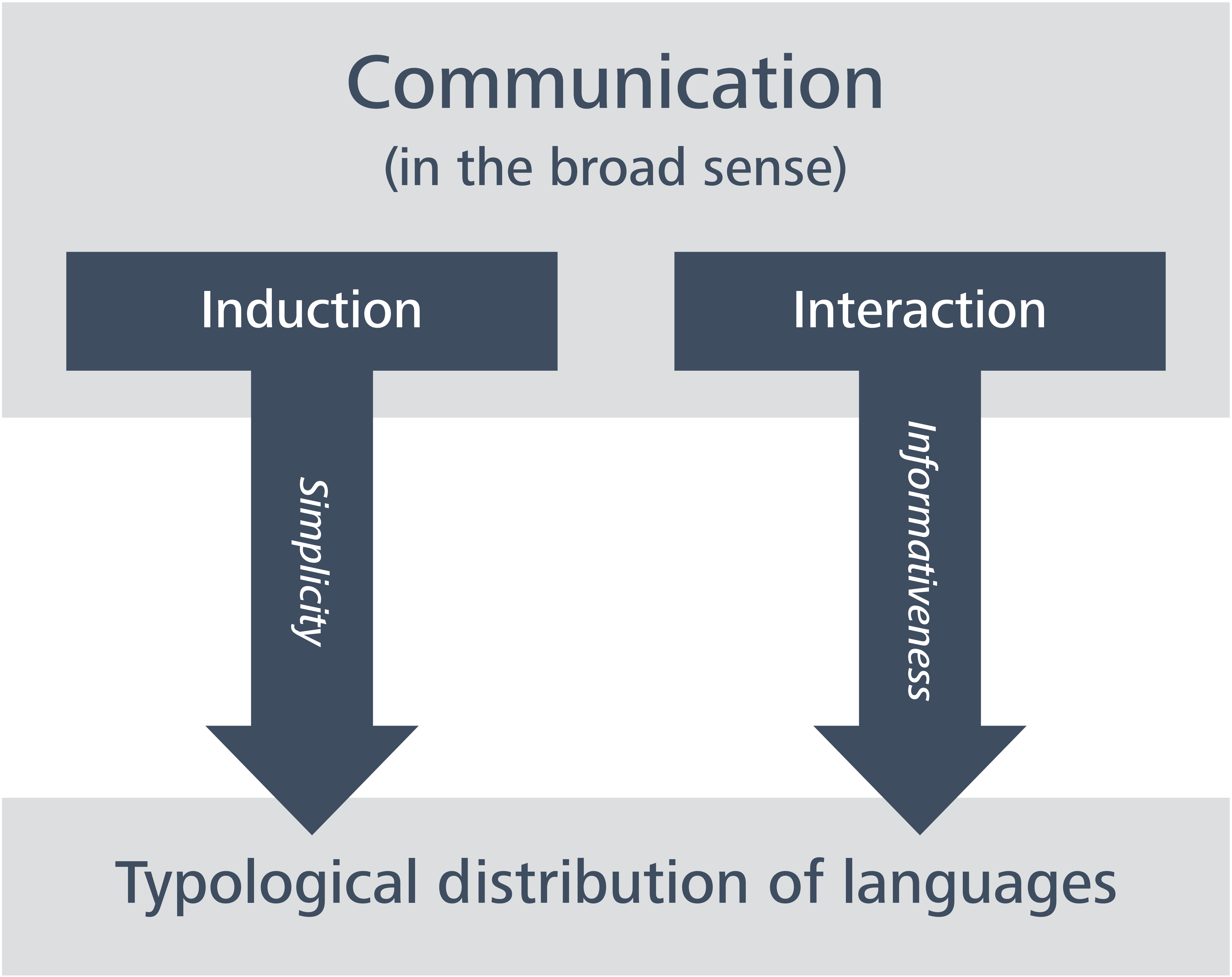
Communication
(in the broad sense)

Typological distribution of languages

The problem of linkage



The problem of linkage



Induction

as the pressure for simplicity

The Minimum Description Length principle

$$\mathrm{DL}(H|D) = \mathrm{DL}(D|H) + \mathrm{DL}(H)$$

The Minimum Description Length principle

$$\mathrm{DL}(H|D) = \mathrm{DL}(D|H) + \mathrm{DL}(H)$$

$$\mathrm{posterior}(H|D) \propto \mathrm{likelihood}(D|H) \times \mathrm{prior}(H)$$

The Minimum Description Length principle

$$\mathrm{DL}(H|D) = \mathrm{DL}(D|H) + \mathrm{DL}(H)$$

$$\mathrm{posterior}(H|D) \propto \mathrm{likelihood}(D|H) \times 2^{-\mathrm{DL}(H)}$$

The Minimum Description Length principle

$$\mathrm{DL}(H|D) = \mathrm{DL}(D|H) + \mathrm{DL}(H)$$

$$\mathrm{posterior}(H|D) \propto \mathrm{likelihood}(D|H) \times 2^{-\mathrm{DL}(H)}$$

Any regularities in data can be used to compress that data

The Minimum Description Length principle

$$\mathrm{DL}(H|D) = \mathrm{DL}(D|H) + \mathrm{DL}(H)$$

$$\mathrm{posterior}(H|D) \propto \mathrm{likelihood}(D|H) \times 2^{-\mathrm{DL}(H)}$$

Any regularities in data can be used to compress that data

The more regularities there are, the more the data can be compressed

The Minimum Description Length principle

$$\mathrm{DL}(H|D) = \mathrm{DL}(D|H) + \mathrm{DL}(H)$$

Any regular

The more re compressed

For example…

010010111110010000110001000101101100001111010001

```
print('010010111110010000110001000101101100001111010001')
```

010101010101010101010101010101010101010101010101

```
print('0101'*12)
```

or

```
print('01'*24)
```

The Minimum Description Length principle

$$\mathrm{DL}(H|D) = \mathrm{DL}(D|H) + \mathrm{DL}(H)$$

$$\mathrm{posterior}(H|D) \propto \mathrm{likelihood}(D|H) \times 2^{-\mathrm{DL}(H)}$$

Any regularities in data can be used to compress that data

The more regularities there are, the more the data can be compressed

We equate **learning** with **compression**: The more the data can be compressed, the more insight we gain from that data

The Minimum Description Length principle

$$\mathrm{DL}(H|D) = \mathrm{DL}(D|H) + \mathrm{DL}(H)$$

$$\mathrm{posterior}(H|D) \propto \mathrm{likelihood}(D|H) \times 2^{-\mathrm{DL}(H)}$$

Any regularities in data can be used to compress that data

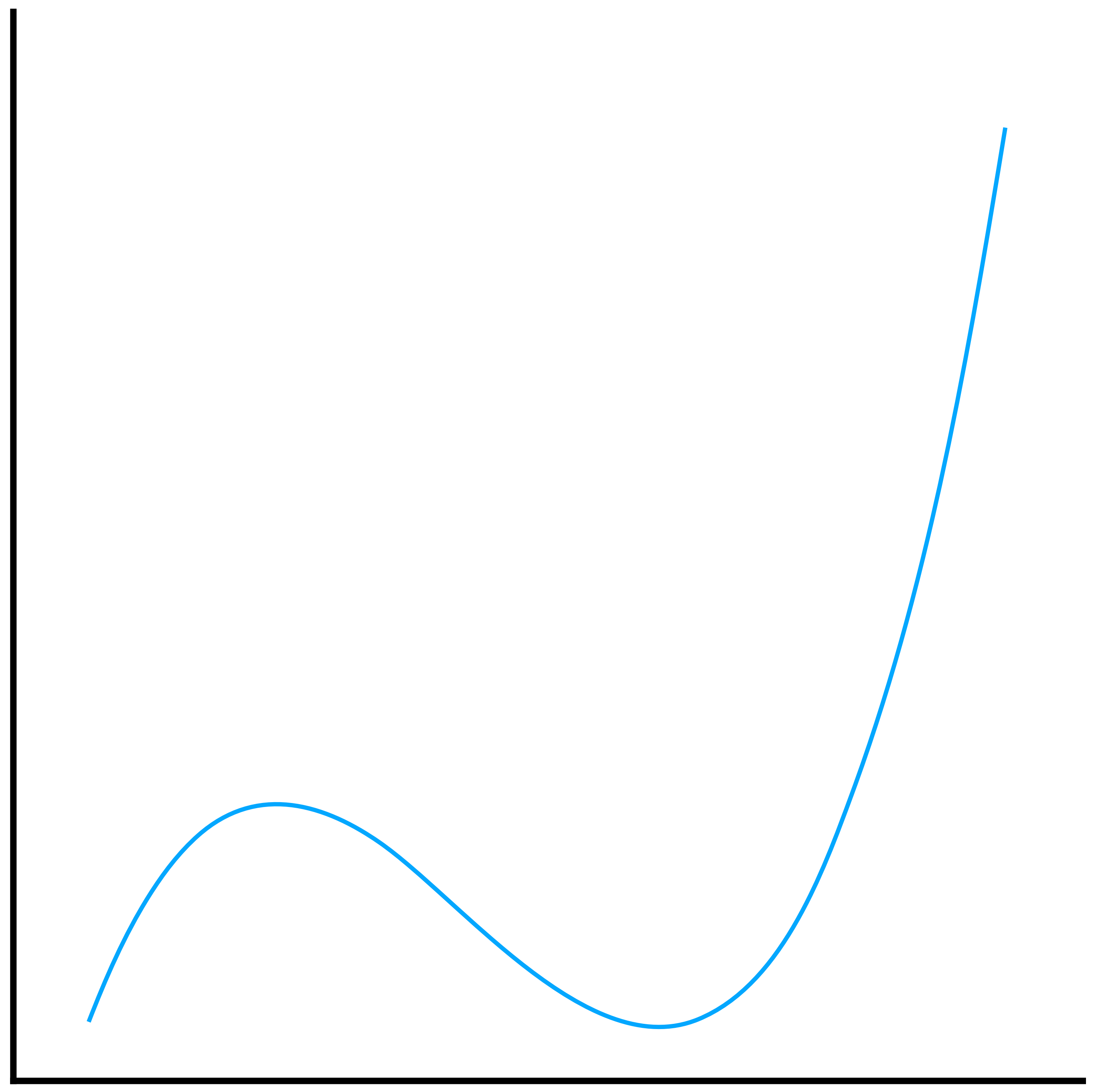
The more regularities there are, the more the data can be compressed

We equate **learning** with **compression**: The more the data can be compressed, the more insight we gain from that data

In other words, the more regularity we can identify, the more we can predict what the generating process will do next

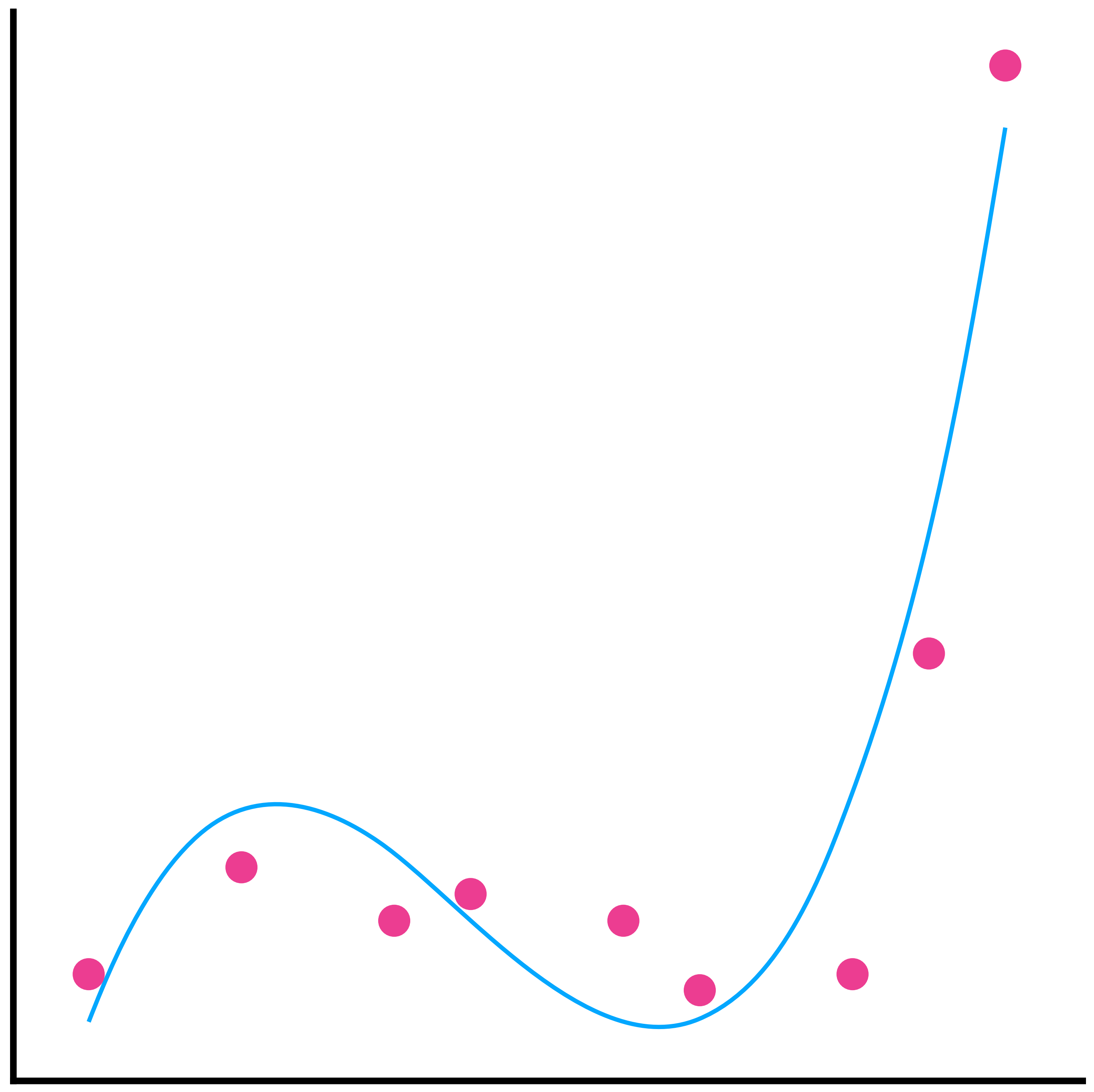
The Minimum Description Length principle

$$\mathrm{DL}(H|D) = \mathrm{DL}(D|H) + \mathrm{DL}(H)$$



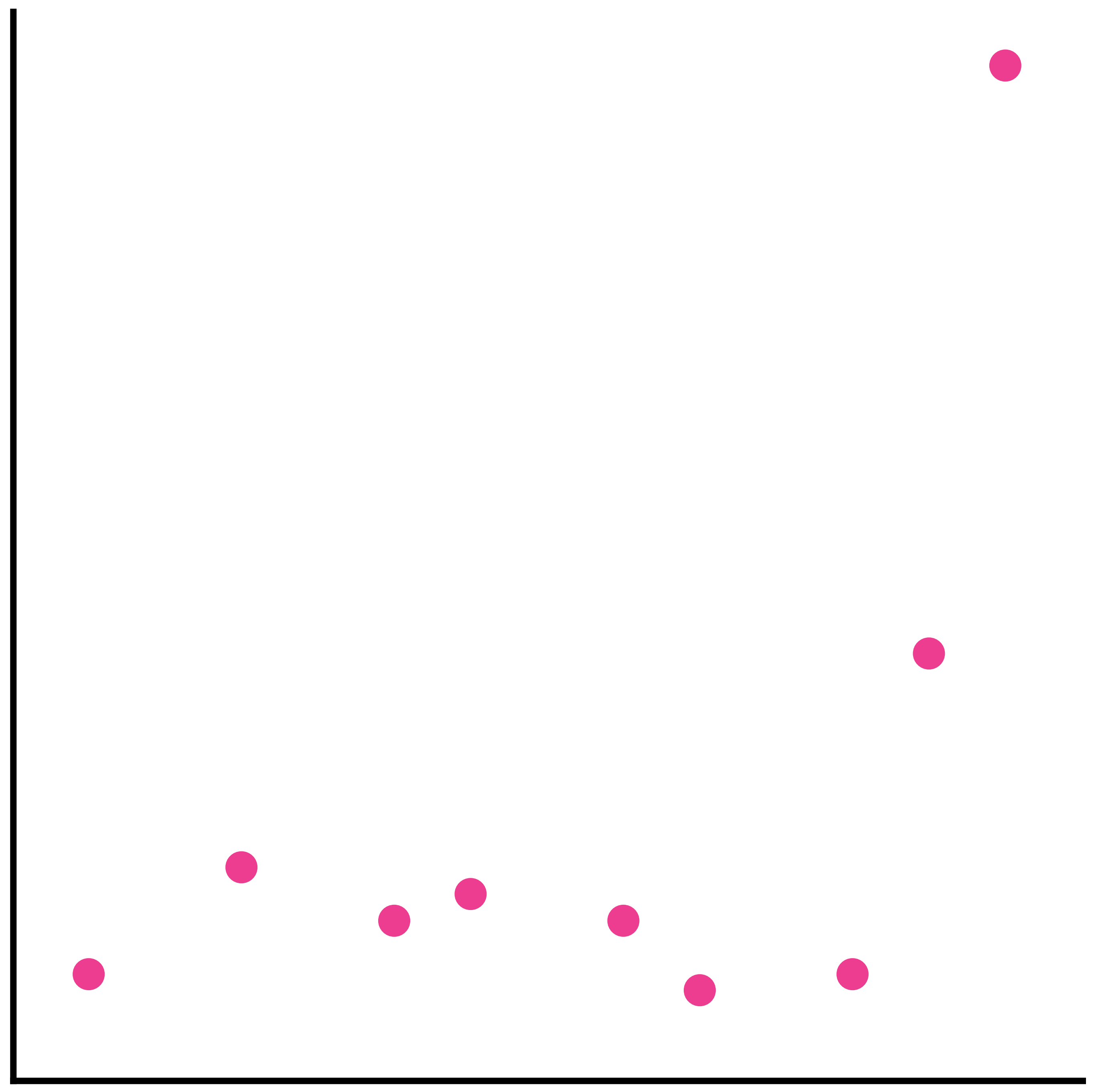
The Minimum Description Length principle

$$\mathrm{DL}(H|D) = \mathrm{DL}(D|H) + \mathrm{DL}(H)$$



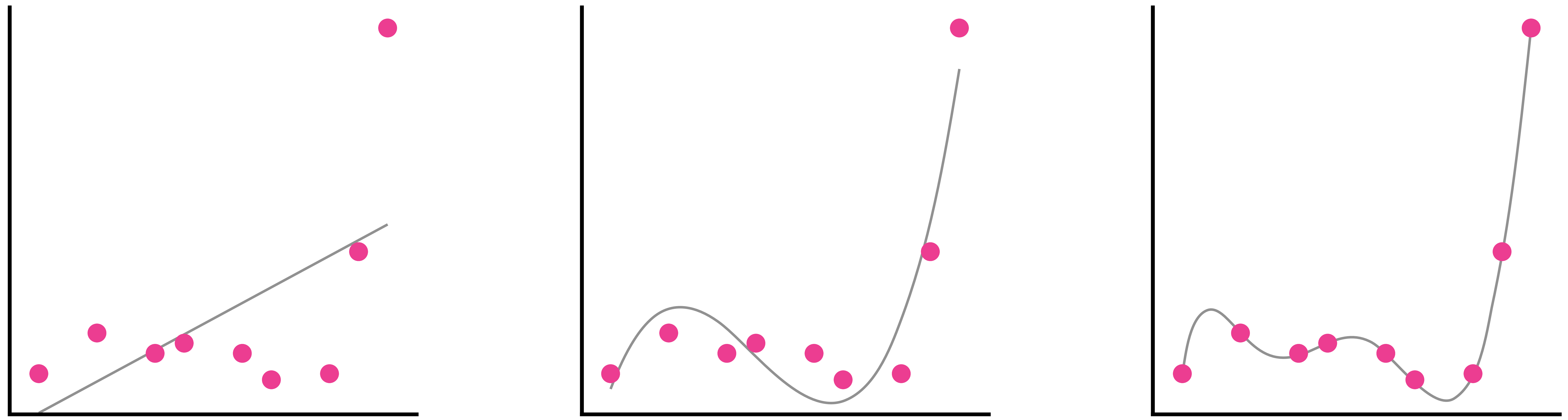
The Minimum Description Length principle

$$\mathrm{DL}(H|D) = \mathrm{DL}(D|H) + \mathrm{DL}(H)$$



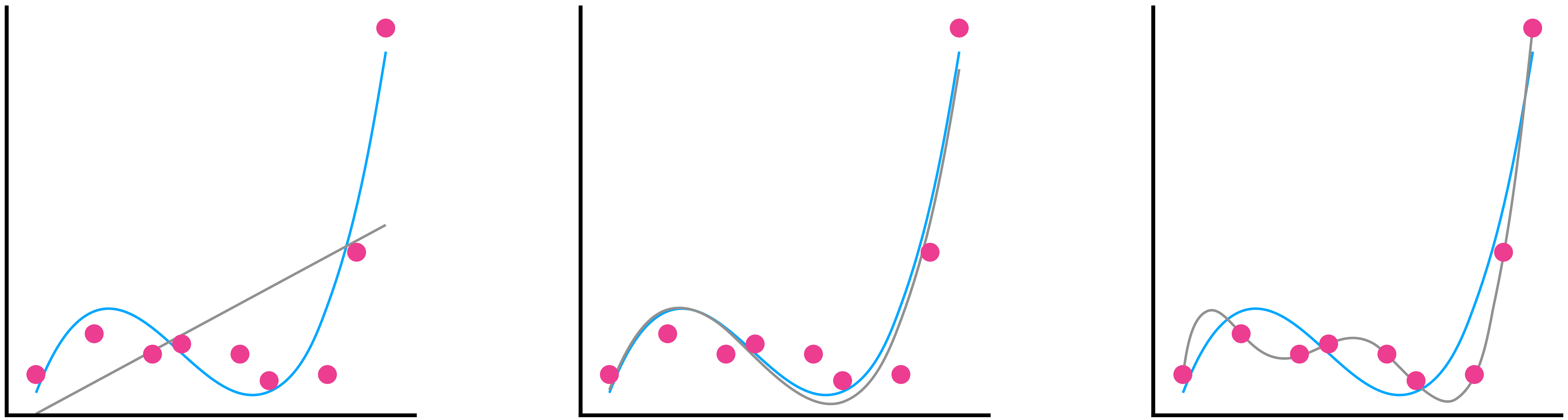
The Minimum Description Length principle

$$\mathrm{DL}(H|D) = \mathrm{DL}(D|H) + \mathrm{DL}(H)$$



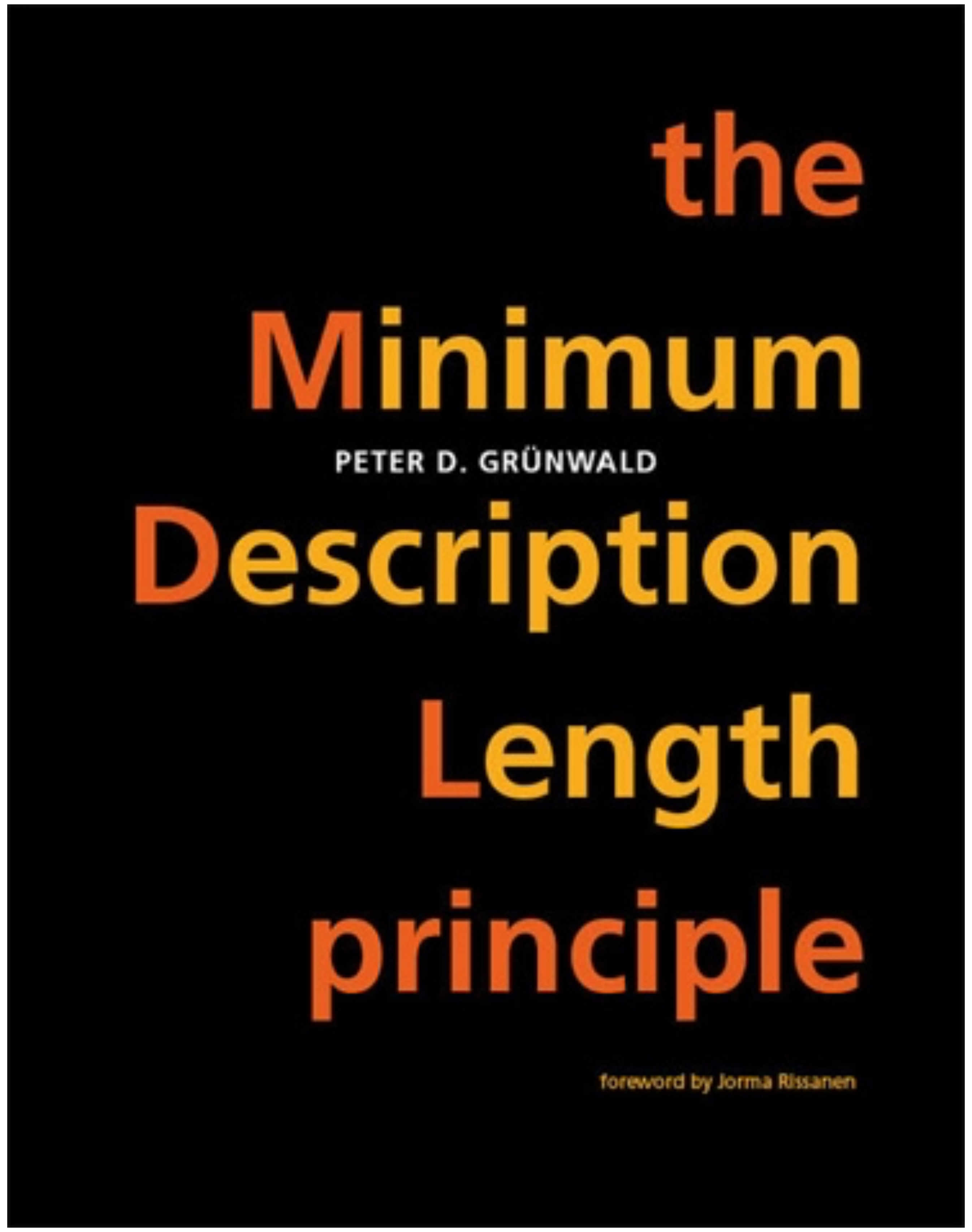
The Minimum Description Length principle

$$\mathrm{DL}(H|D) = \mathrm{DL}(D|H) + \mathrm{DL}(H)$$



$\mathrm{DL}(D\|H)$	10 bits	4 bits	0 bits
$\mathrm{DL}(H)$	1 bit	4 bits	10 bits
$\mathrm{DL}(H\|D)$	11 bits	8 bits ✔	10 bits

The Minimum Description Length principle



Bayesian interpretation: MDL is closely related to Bayesian inference

Occam's razor: MDL trades-off *goodness-of-fit* with *model complexity*, embodying Occam's razor

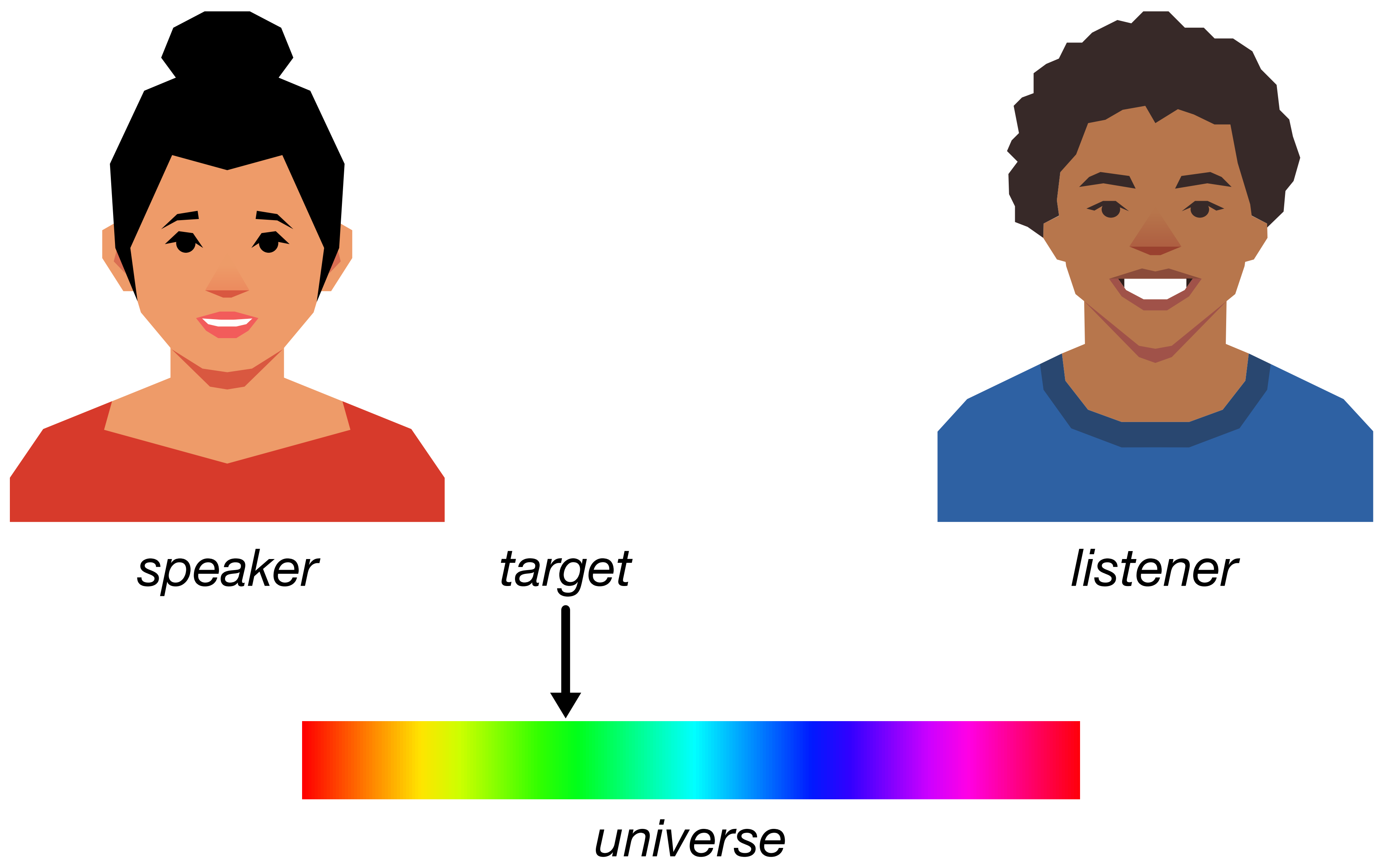
No overfitting: MDL automatically guards against *overfitting noise* in data

Predictive performance: Since data compression is formally equivalent to probabilistic prediction, MDL finds models offering *good predictive performance on unseen data*

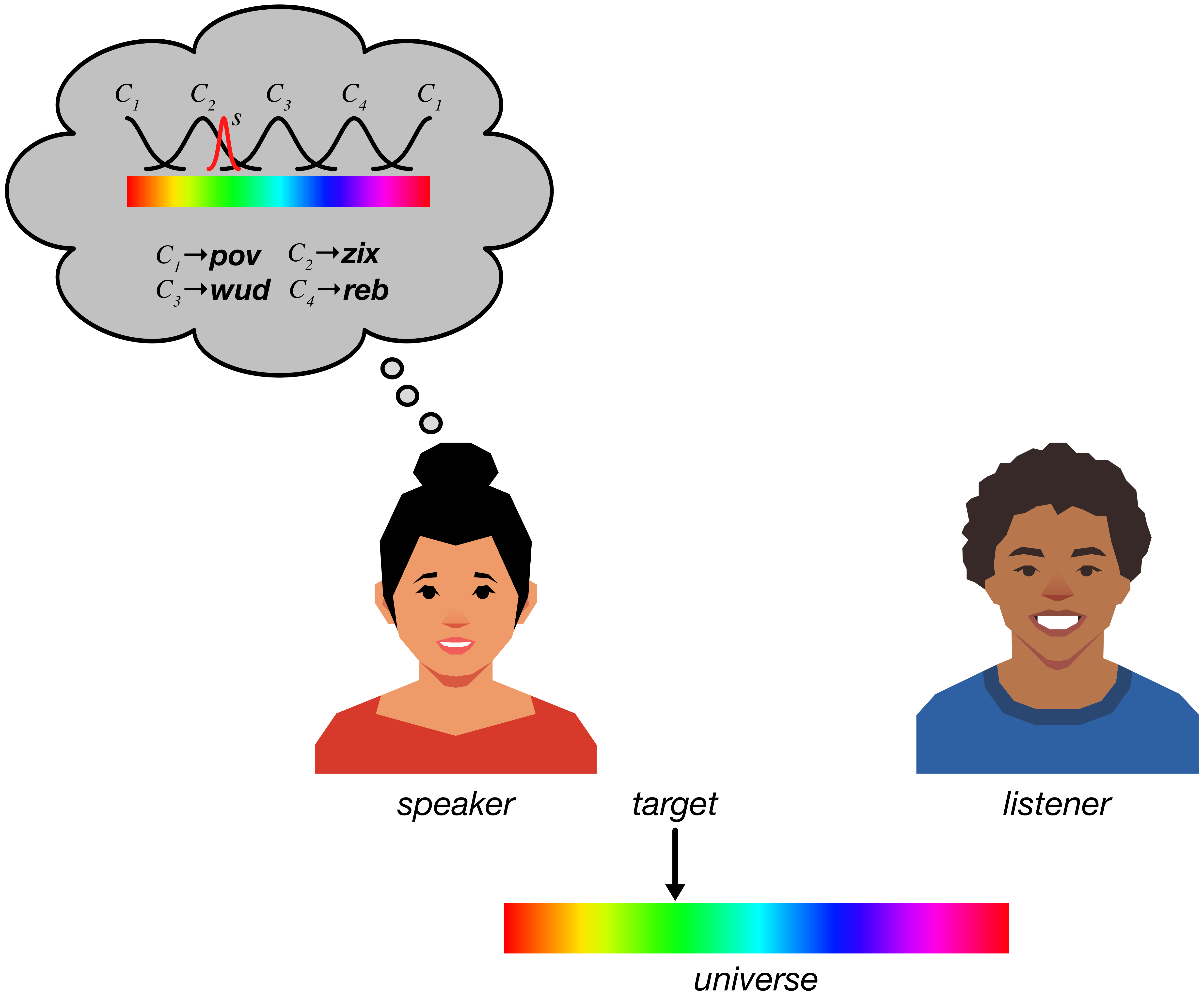
Interaction

as the pressure for informativeness

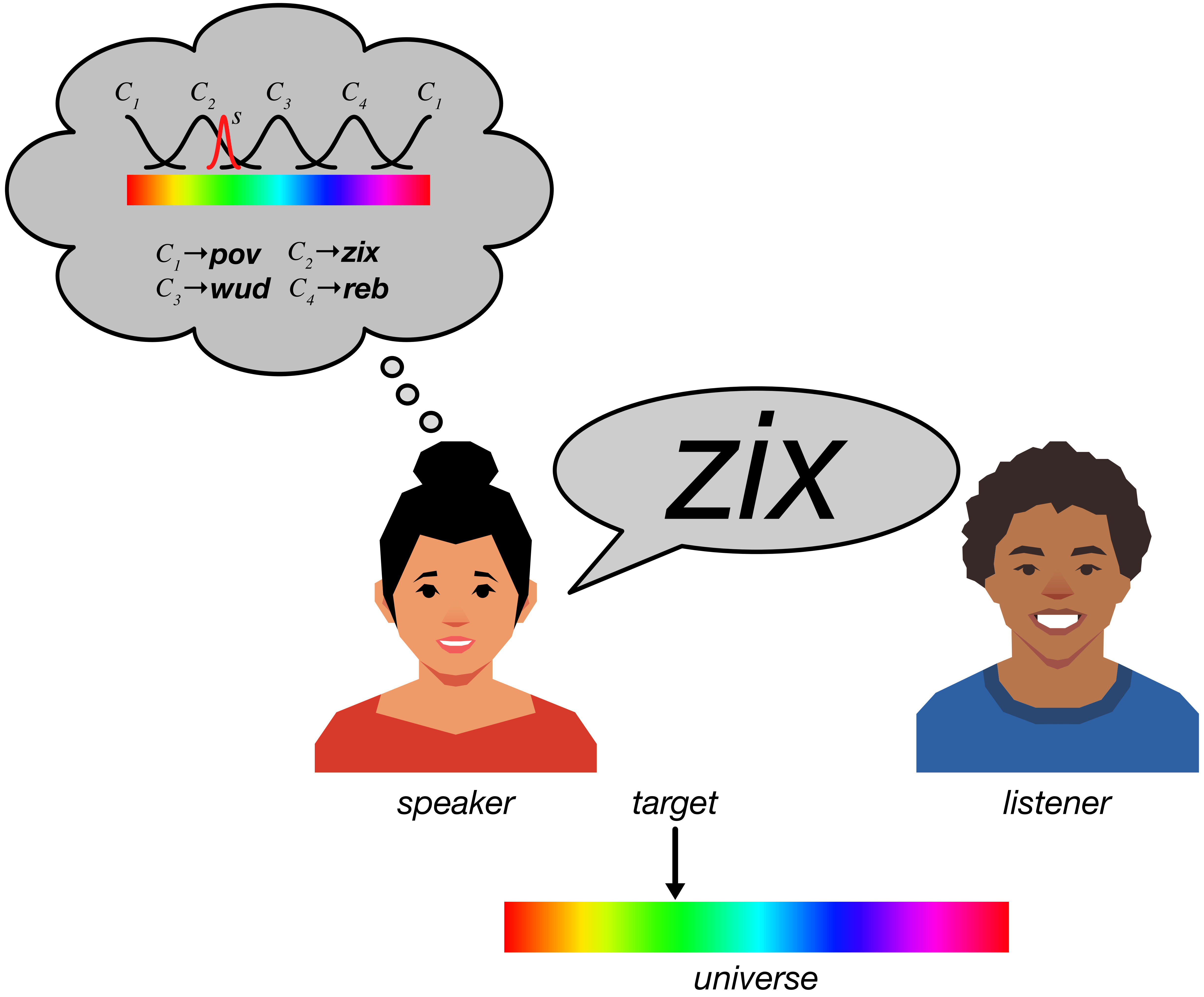
Regier et al.'s informativeness model



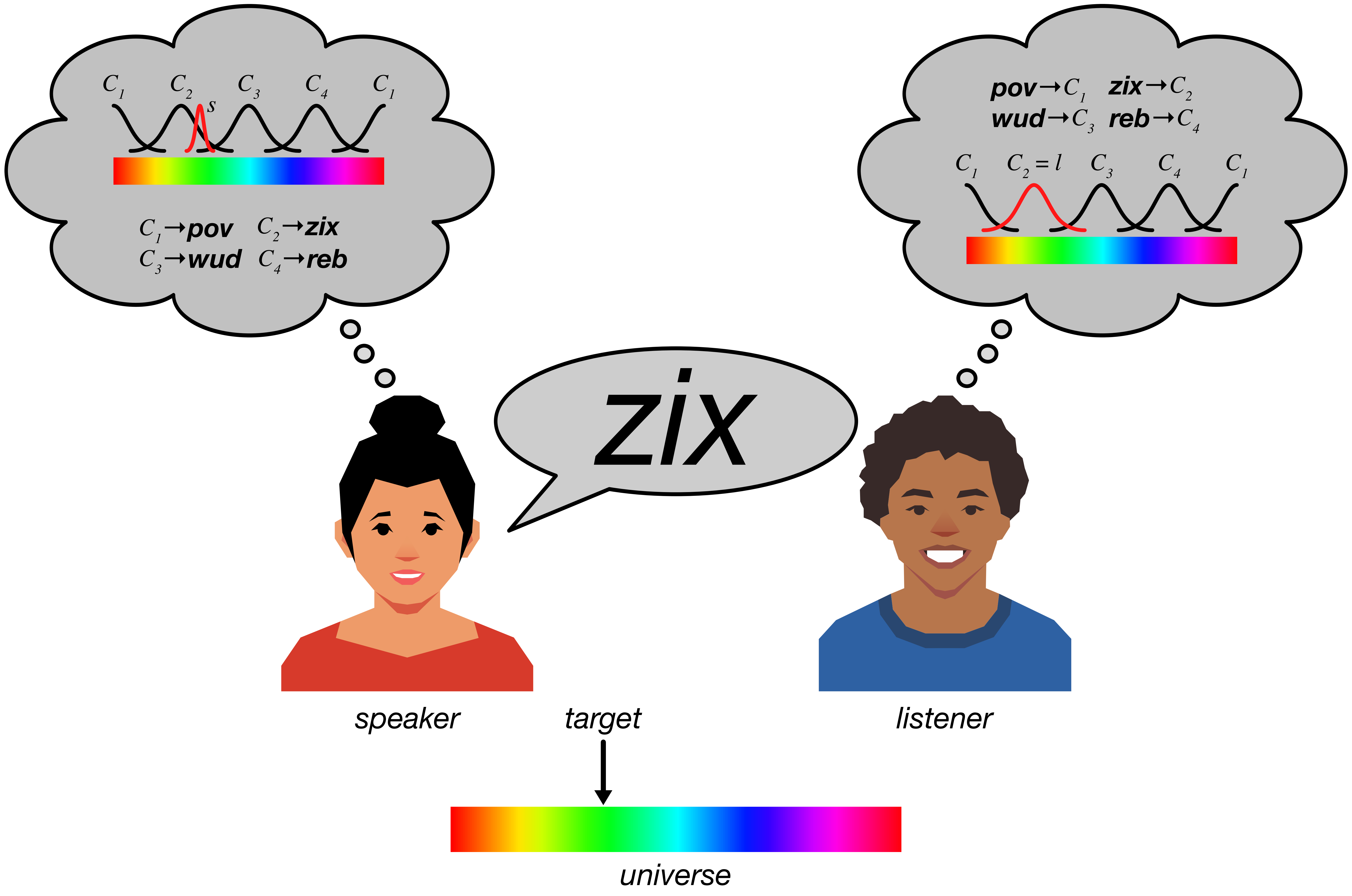
Regier et al.'s informativeness model



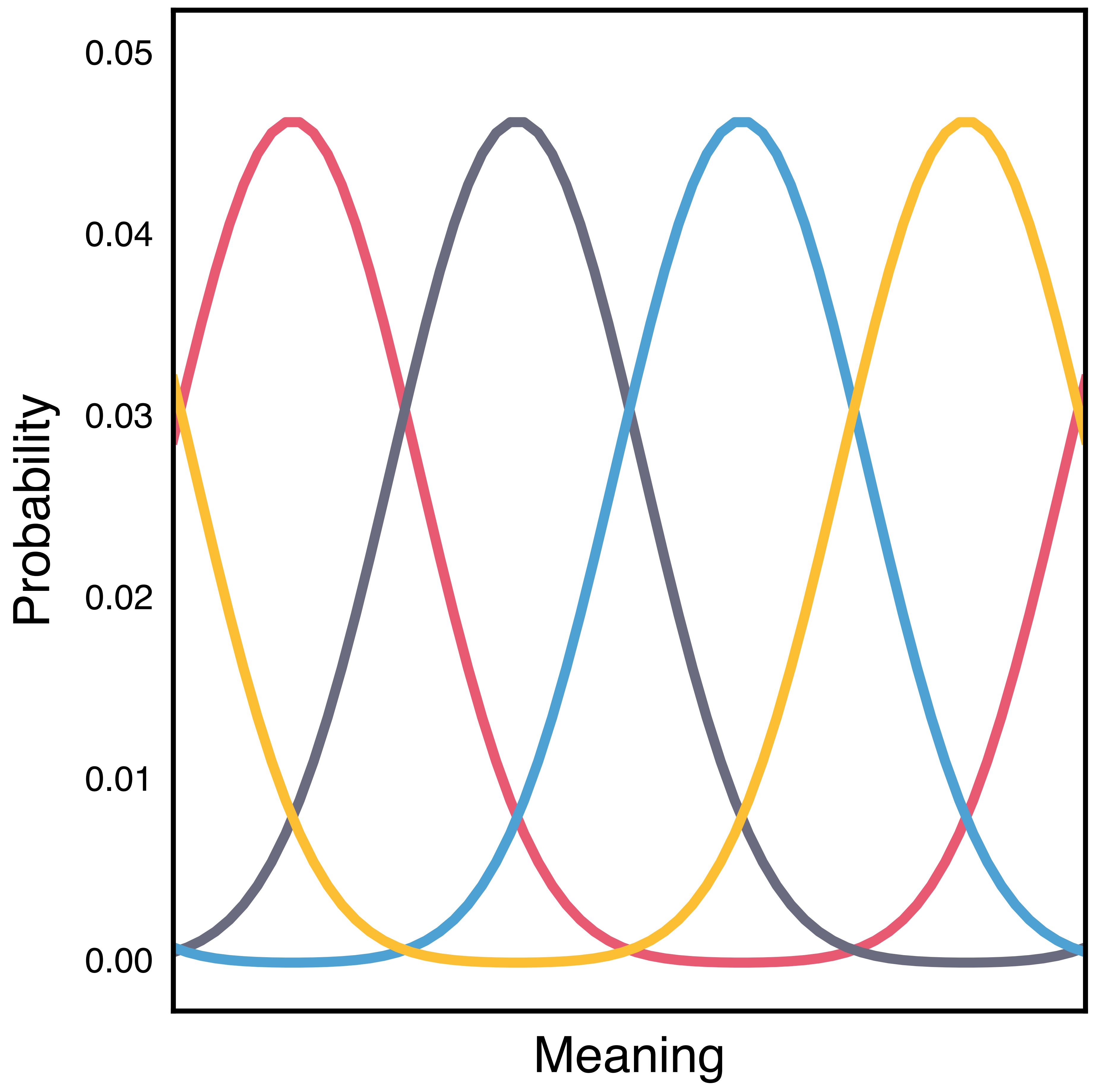
Regier et al.'s informativeness model



Regier et al.'s informativeness model



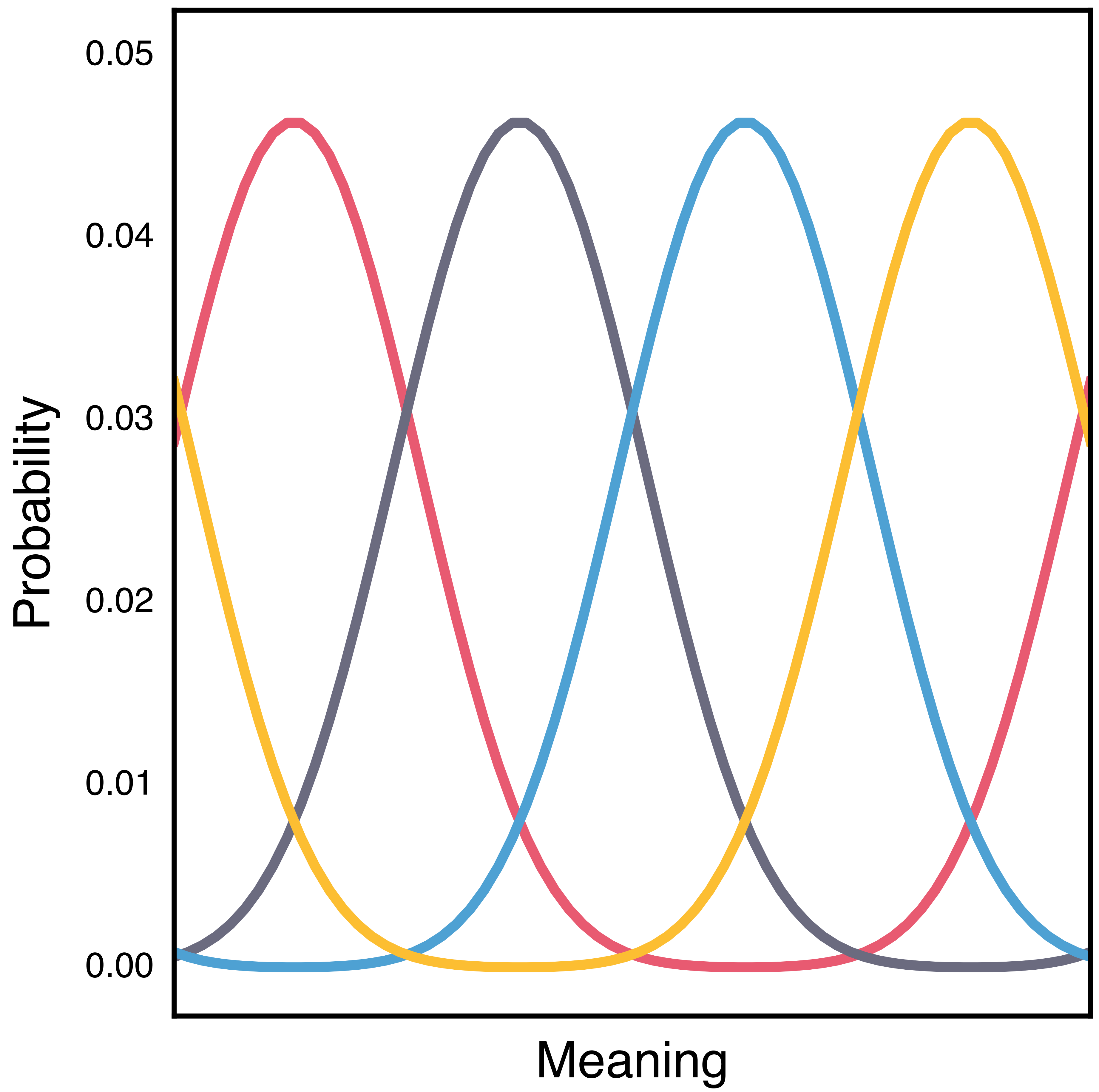
Communicative cost



$$C_j(i) \propto \sum_{c \in C_j} e^{-\gamma d(i,c)^2}$$

$$K(L) := \sum_{i \in U} P(i) \cdot -\log C(i)$$

Communicative cost

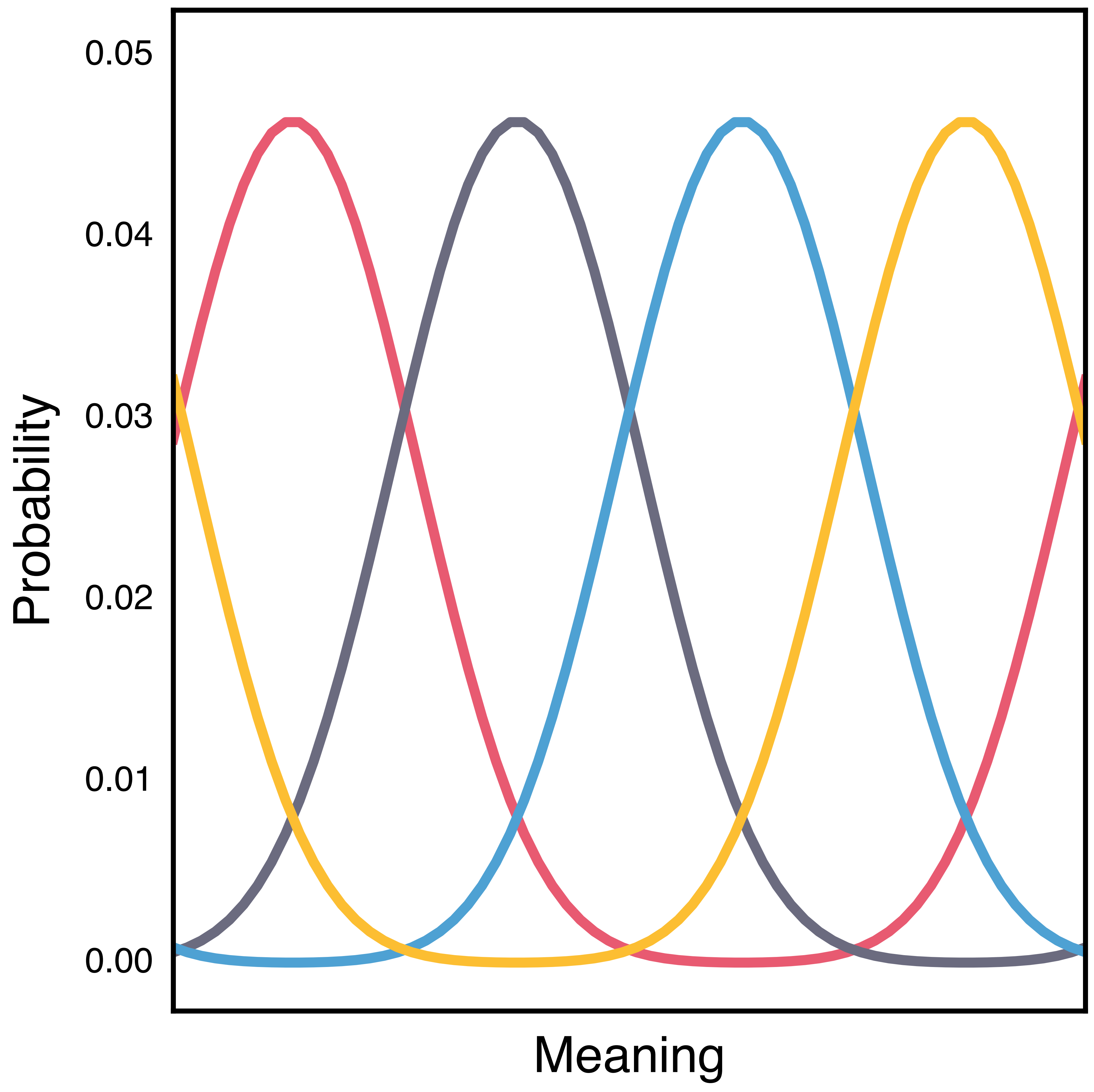


$$C_j(i) \propto \sum_{c \in C_j} e^{-\gamma d(i,c)^2}$$

$$K(L) := \sum_{i \in U} P(i) \cdot - \log C(i)$$

Expressivity A system of many categories is more informative than a system of few categories

Communicative cost



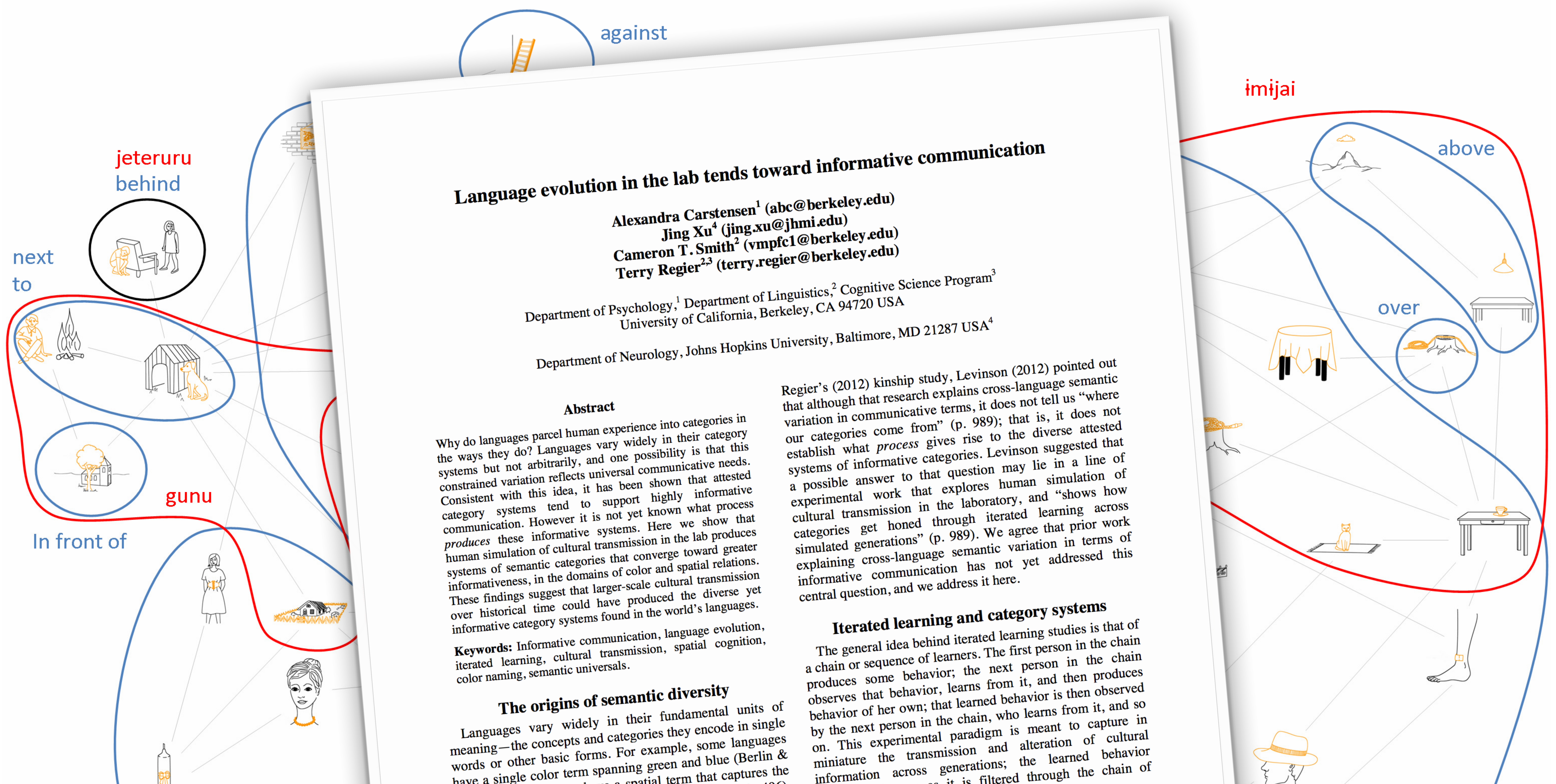
$$C_j(i) \propto \sum_{c \in C_j} e^{-\gamma d(i,c)^2}$$

$$K(L) := \sum_{i \in U} P(i) \cdot -\log C(i)$$

Expressivity A system of many categories is more informative than a system of few categories

Compactness A system of compact categories is more informative than a system of noncompact categories

Could humans have a learning bias for informativeness?



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

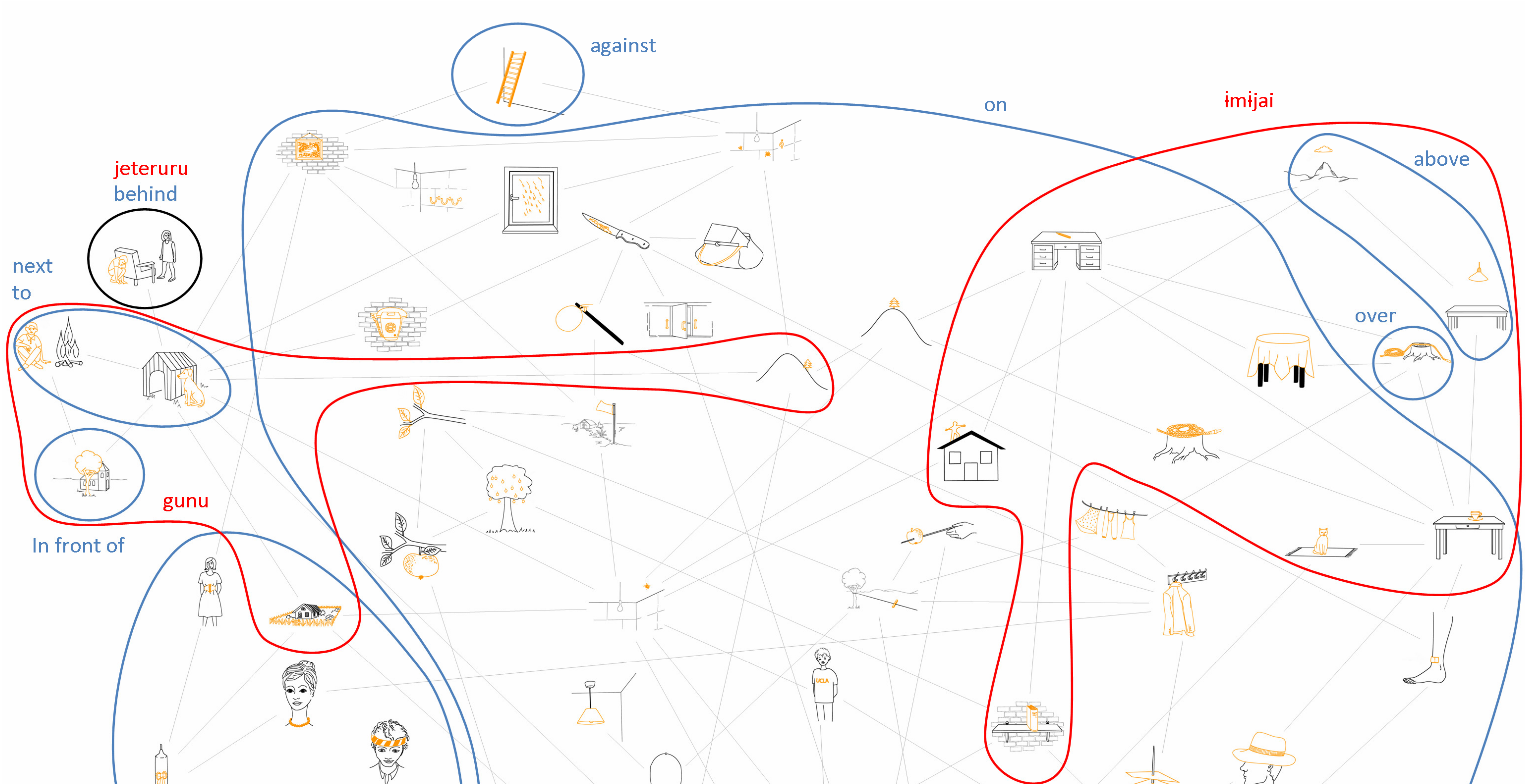
Languages vary widely in their fundamental units of meaning—the concepts and categories they encode in single words or other basic forms. For example, some languages have a single color term spanning green and blue (Berlin &

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.

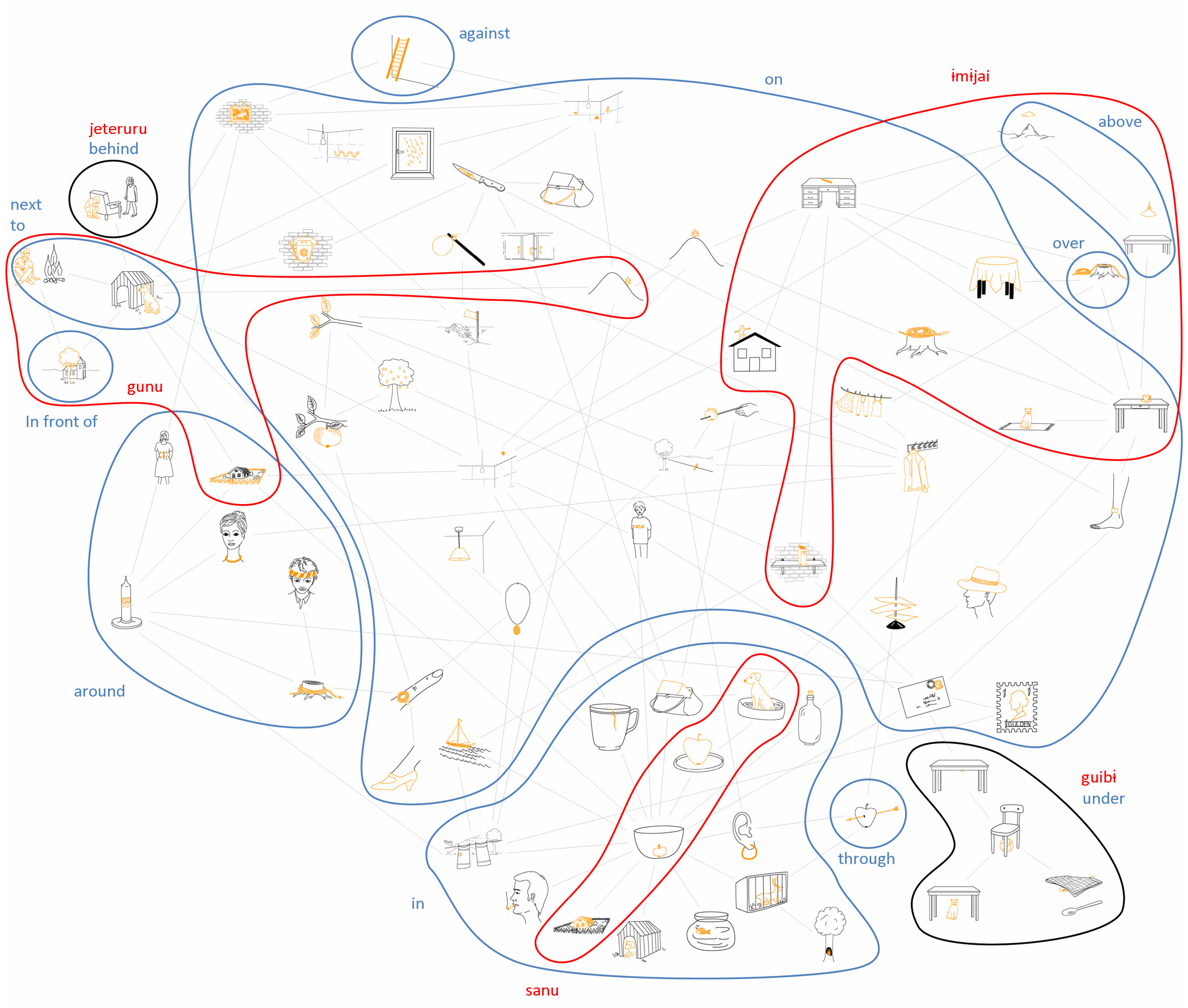
Iterated learning and category systems

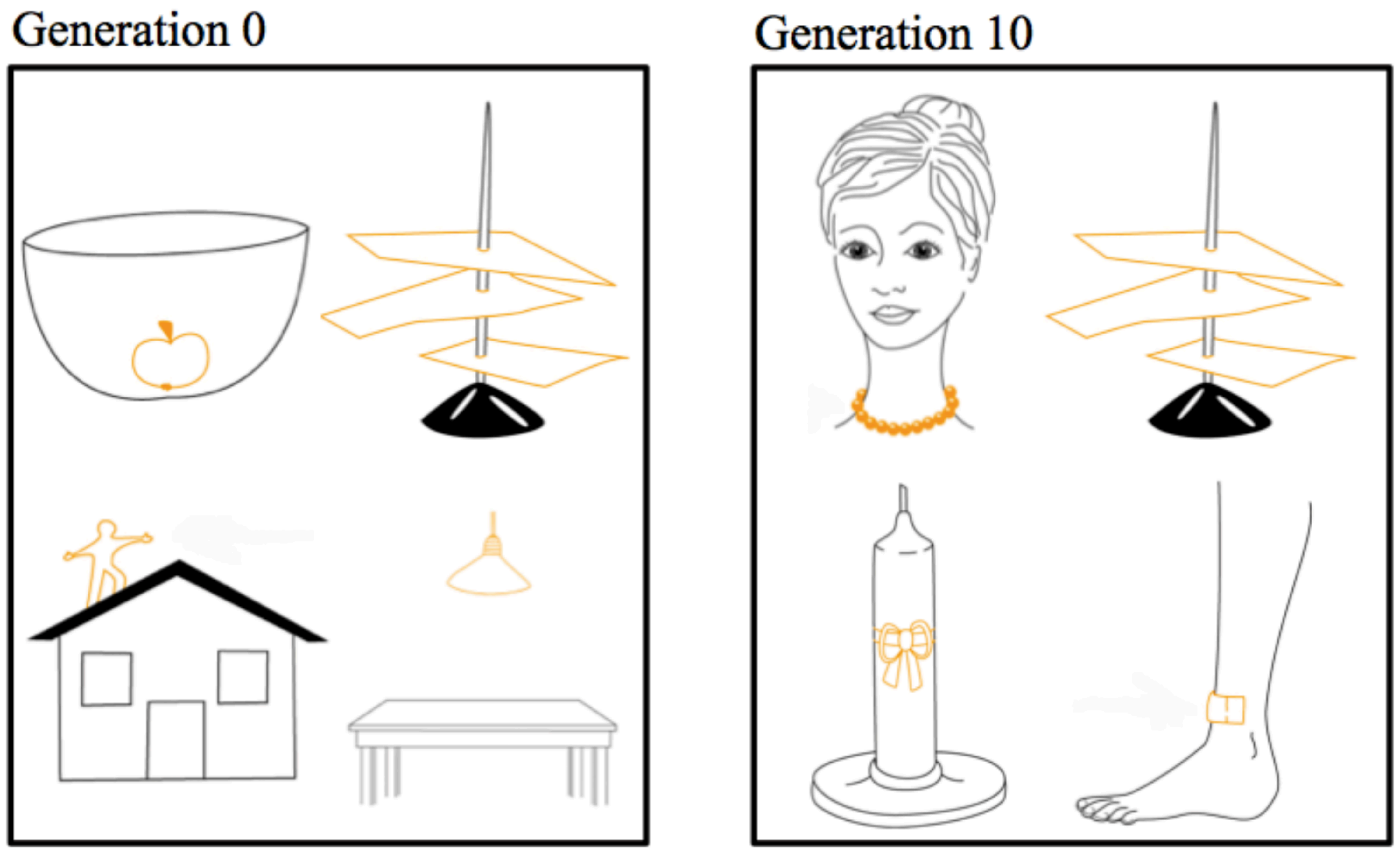
The general idea behind iterated learning studies is that of a chain or sequence of learners. The first person in the chain produces some behavior; the next person in the chain observes that behavior, learns from it, and then produces behavior of her own; that learned behavior is then observed by the 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

Could humans have a learning bias for informativeness?



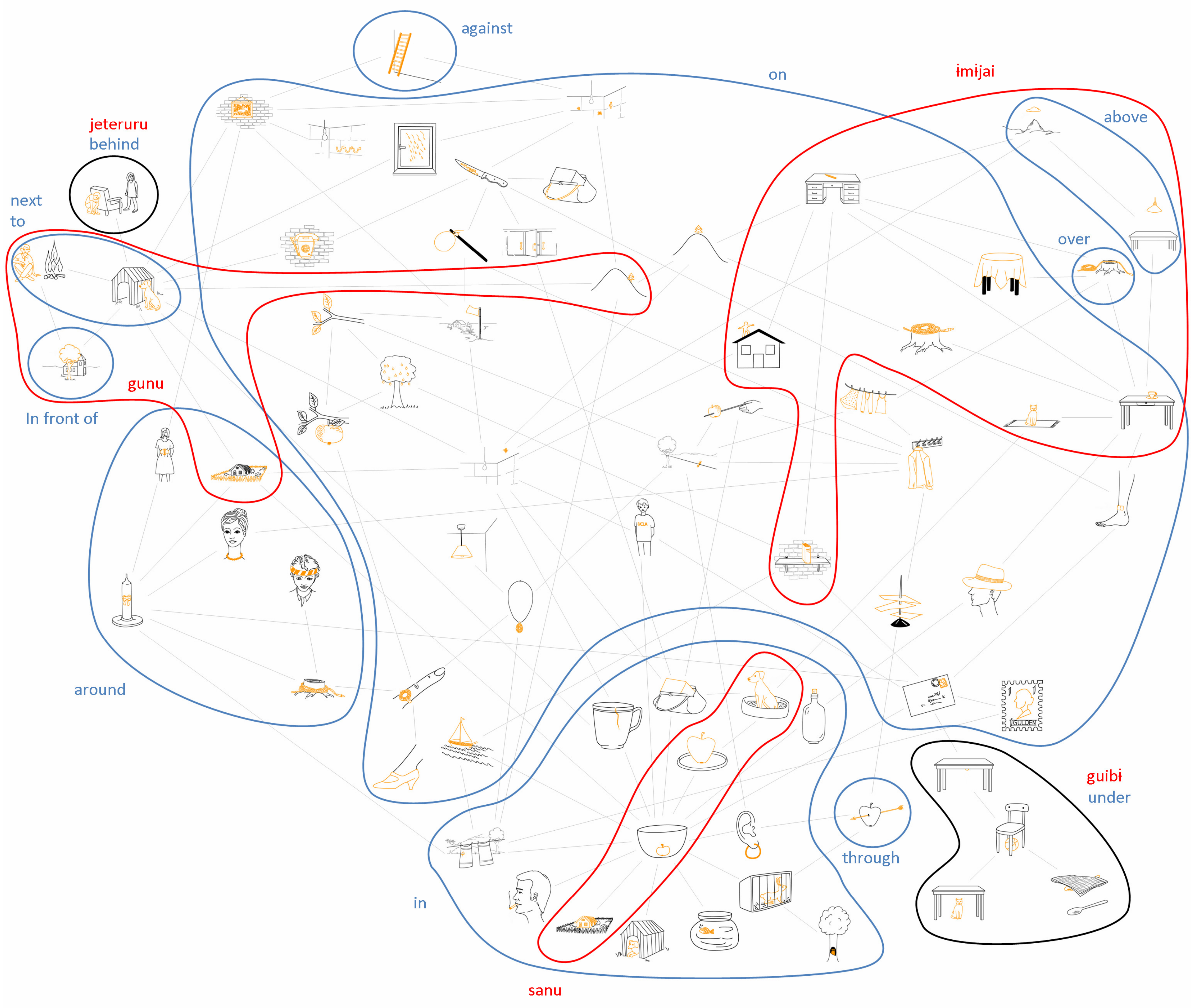
Could humans have a learning bias for informativeness?

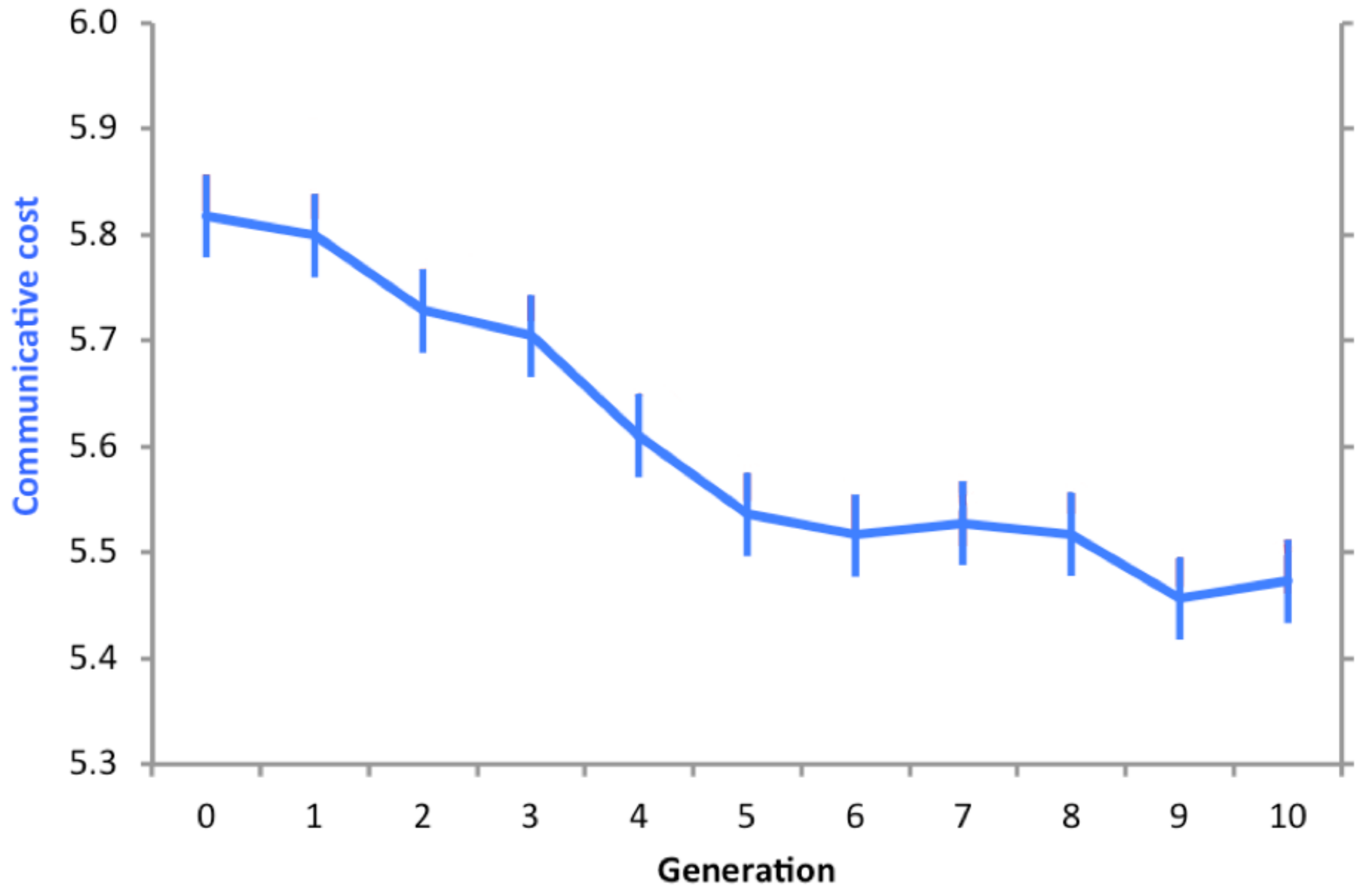




Carstensen, Xu, Smith, Regier (2015)

Could humans have a learning bias for informativeness?

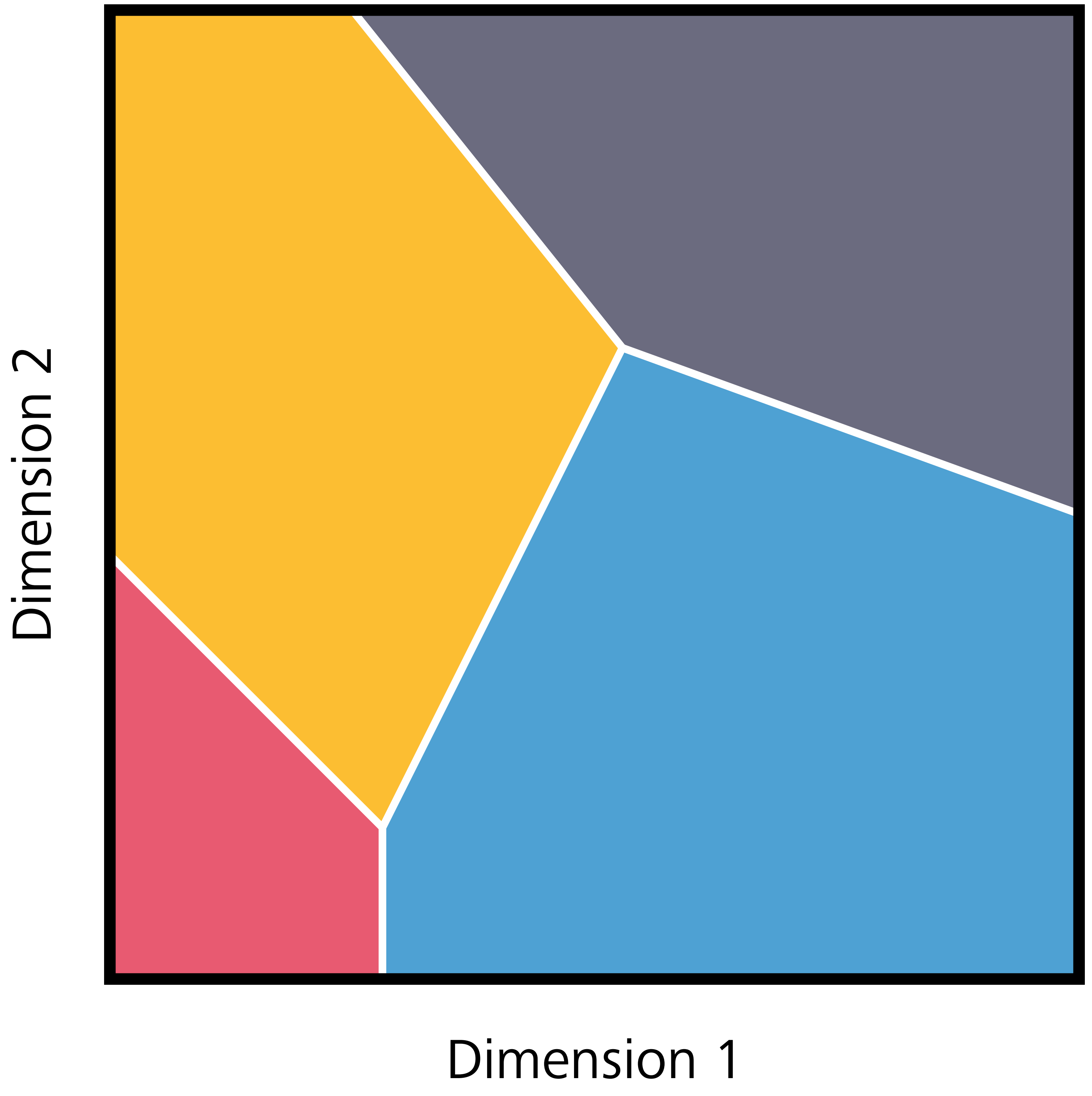




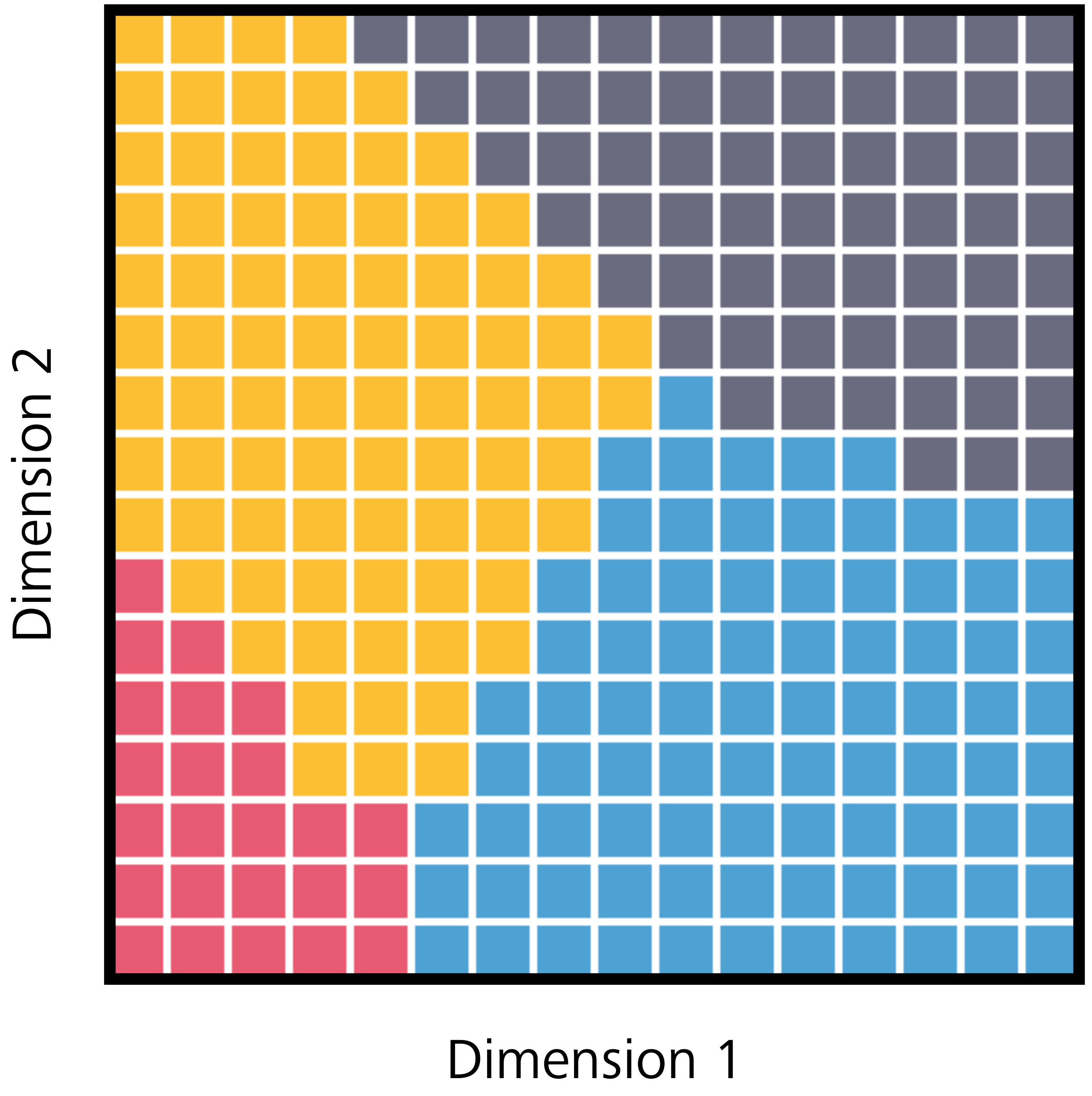
Carstensen, Xu, Smith, Regier (2015)

Bayesian model

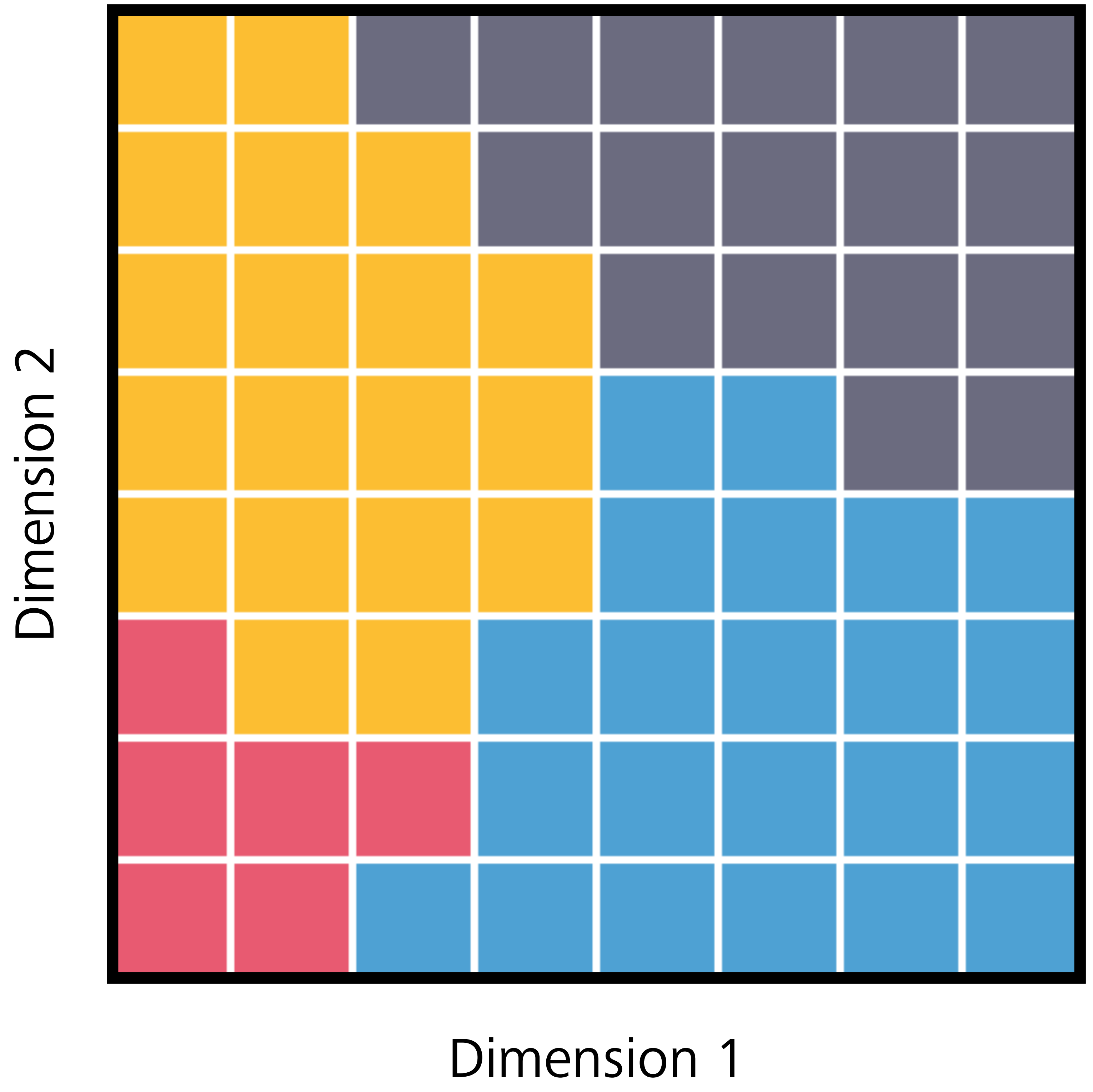
Conceptual spaces and convexity



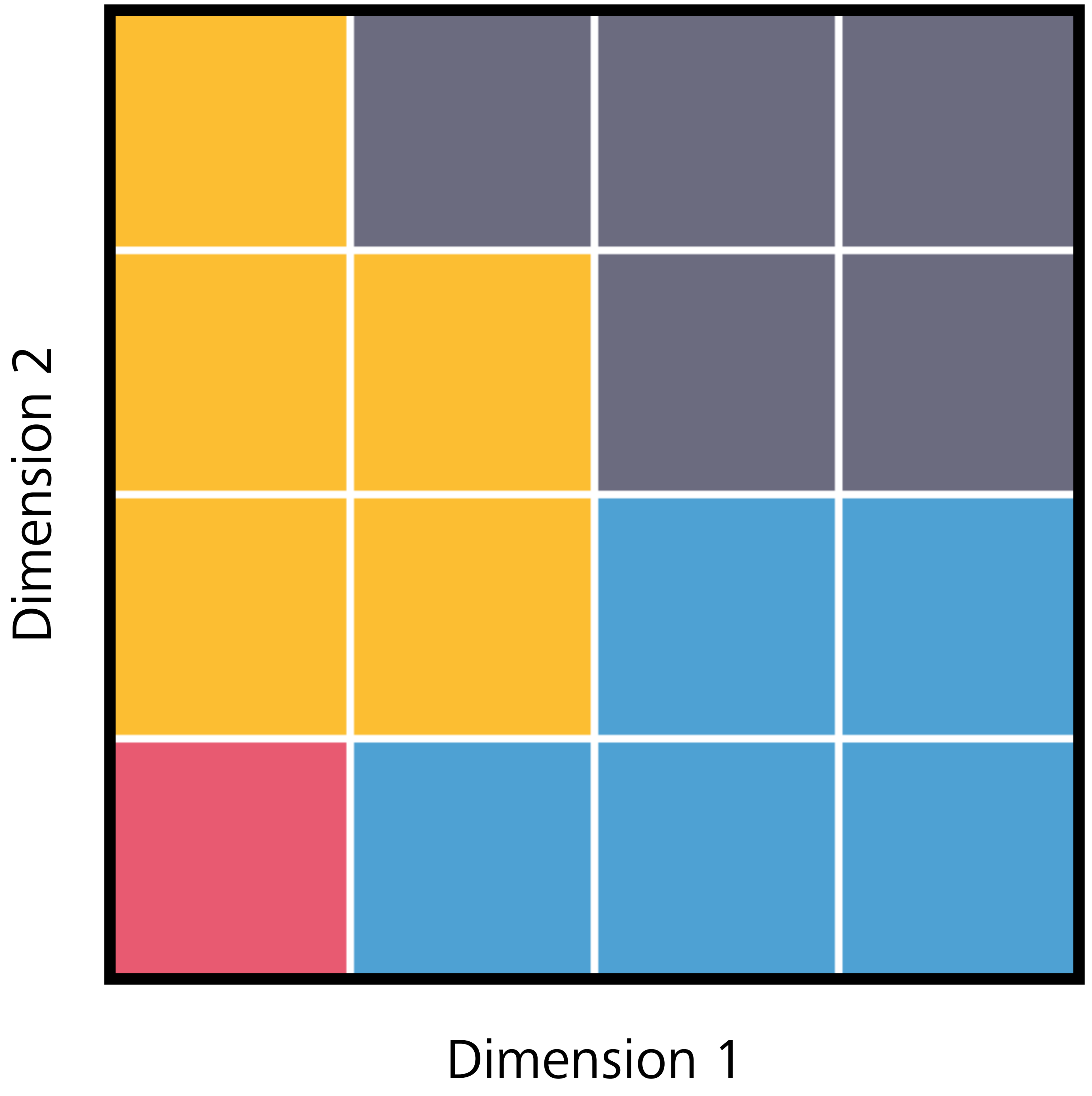
Conceptual spaces and convexity



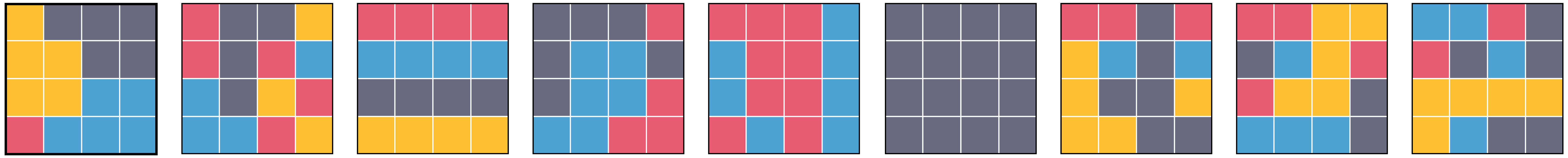
Conceptual spaces and convexity



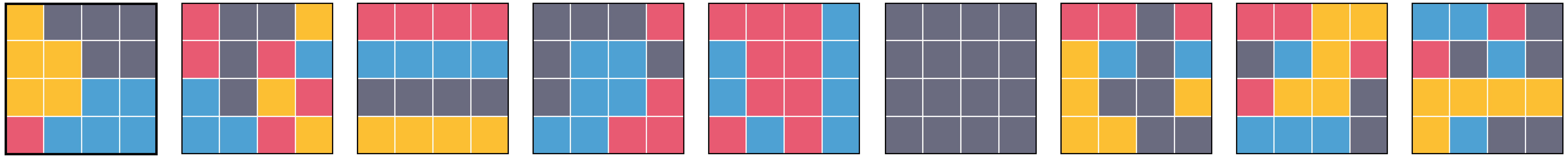
Conceptual spaces and convexity



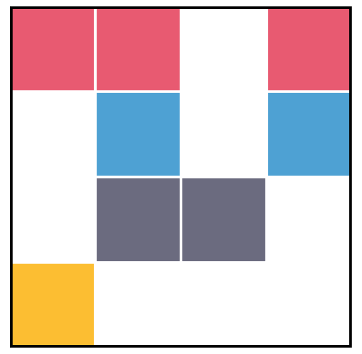
Bayesian inference

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

Bayesian inference

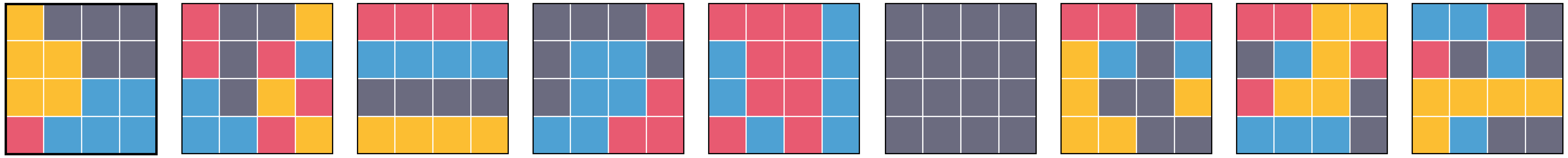
$\mathcal{L} = \{$  $\cdots\}$

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



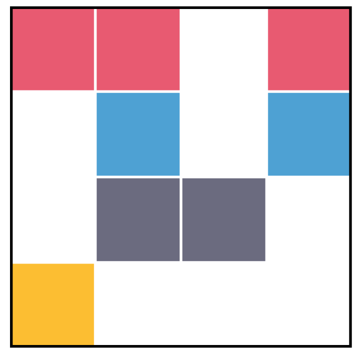
Bayesian inference

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

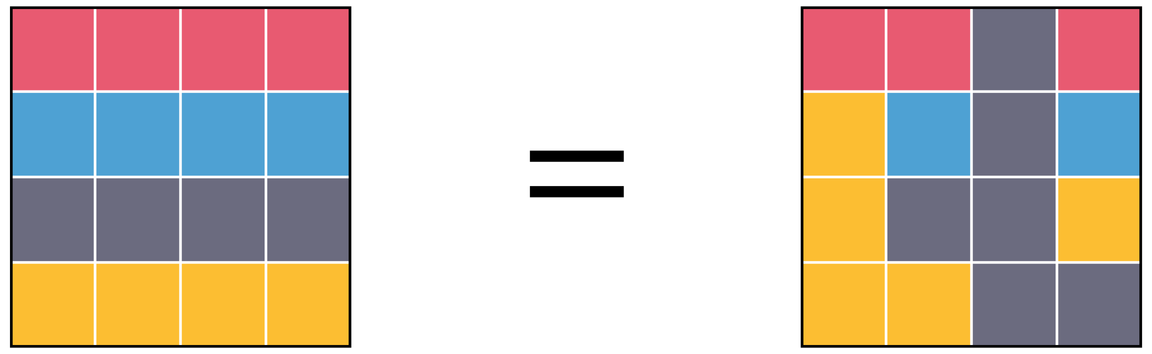


$$\cdots\}$$

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

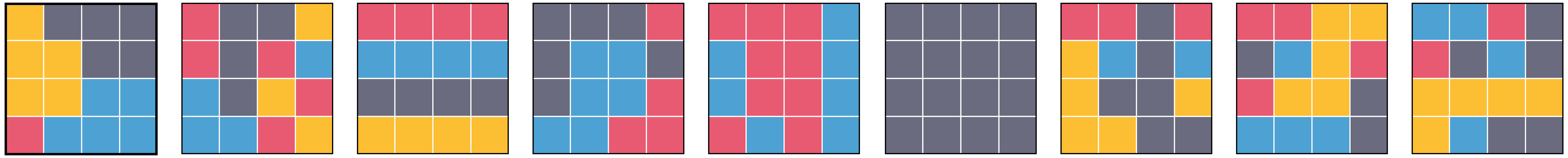


$$\text{likelihood}(D|L) \propto \prod_{\langle m,s\rangle \in D} \frac{1}{|M|} P(s|L,m)$$



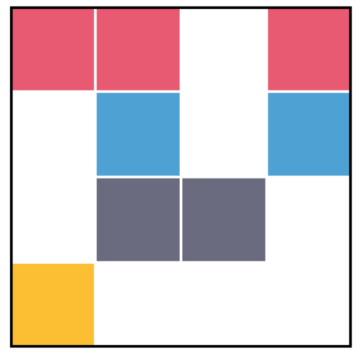
Bayesian inference

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



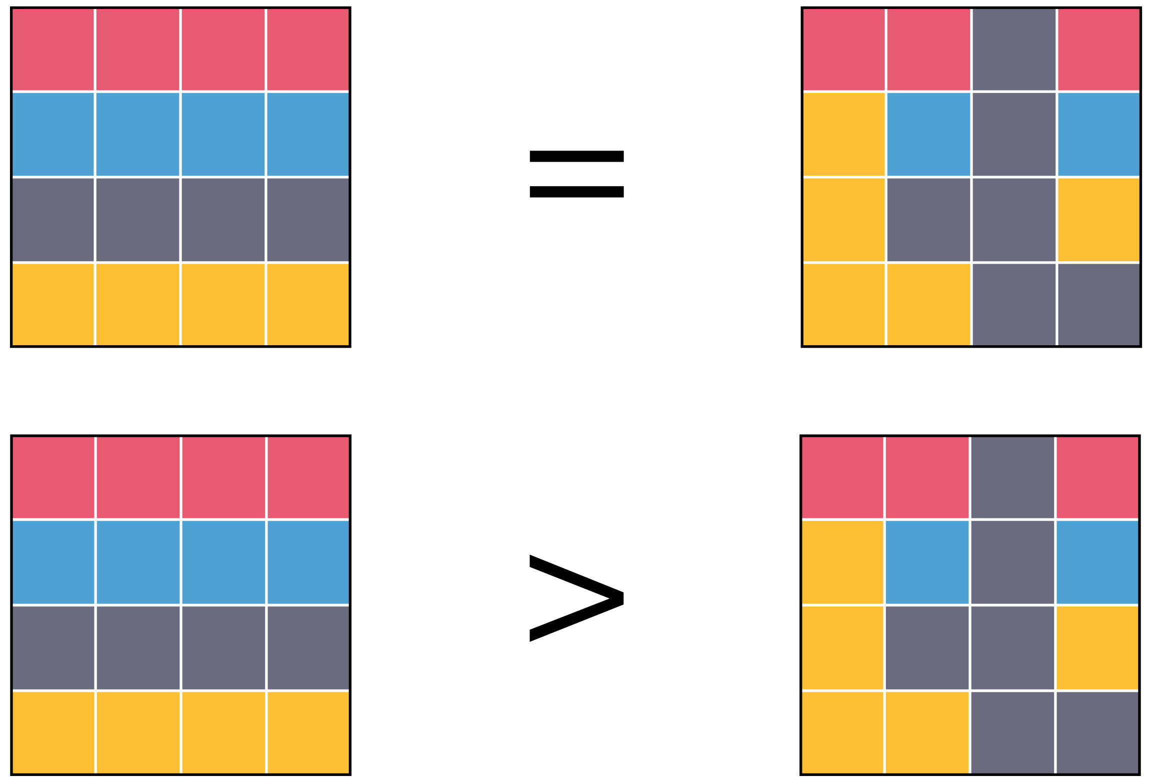
$$\cdots\}$$

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

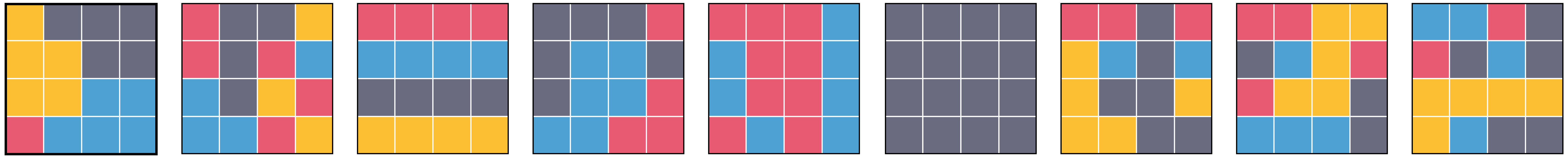


$$\text{likelihood}(D|L) \propto \prod_{\langle m,s \rangle \in D} \frac{1}{|M|} P(s|L, m)$$

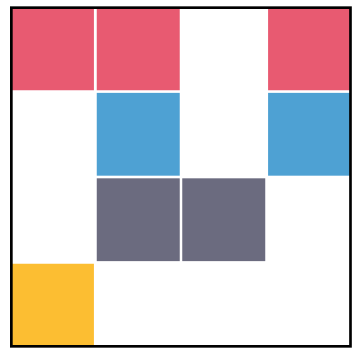
$$\text{prior}(L) \propto 2^{-\text{DL}(L)}$$



Bayesian inference

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

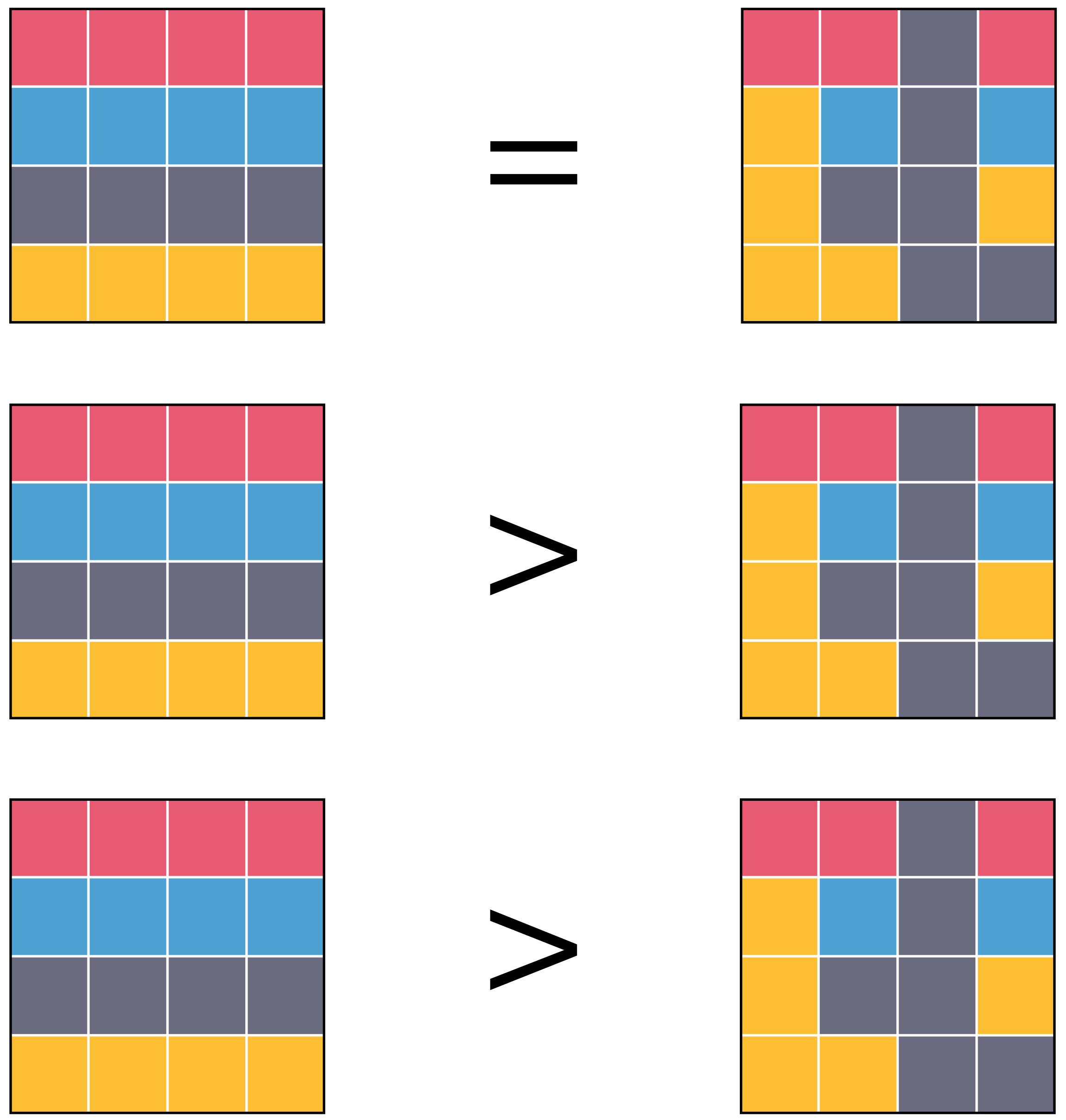
$$D = [\langle m_1, s_1\rangle, \langle m_2, s_2\rangle, \langle m_3, s_3\rangle, \ldots, \langle m_n, s_n\rangle]$$



$$\text{likelihood}(D|L) \propto \prod_{\langle m,s\rangle \in D} \frac{1}{|M|} P(s|L,m)$$

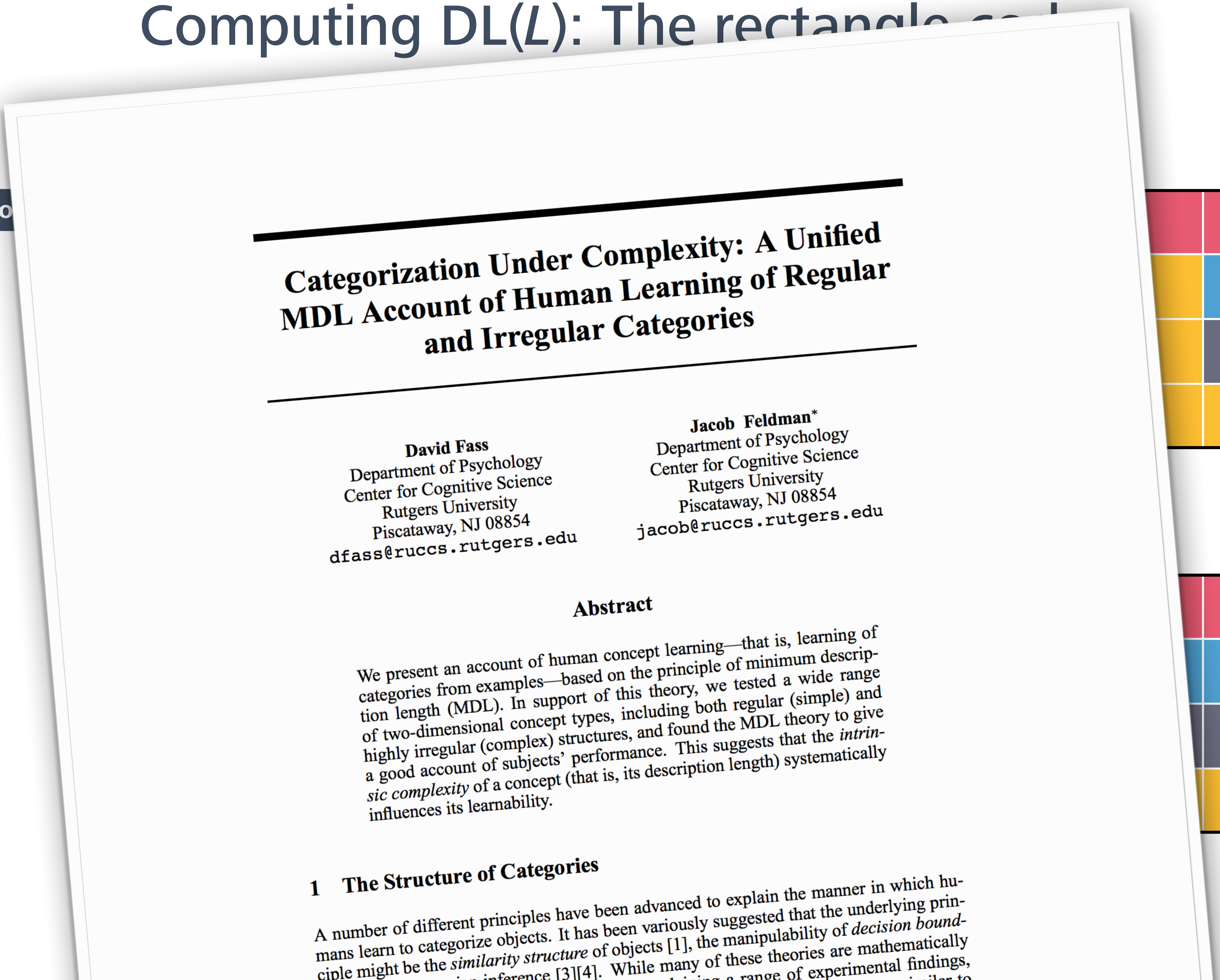
$$\text{prior}(L) \propto 2^{-\text{DL}(L)}$$

$$\text{posterior}(L|D) \propto \text{likelihood}(D|L) \times \text{prior}(L)$$



Computing DL(*L*): The rectangle co…

Class	Positio…
1×1	16
1×2	12
1×3	8
1×4	4
2×1	12
2×2	9
2×3	6
2×4	3
3×1	8
3×2	6
3×3	4
3×4	2
4×1	4
4×2	3
4×3	2
4×4	1



Categorization Under Complexity: A Unified MDL Account of Human Learning of Regular and Irregular Categories

David Fass
Department of Psychology
Center for Cognitive Science
Rutgers University
Piscataway, NJ 08854
dfass@ruccs.rutgers.edu

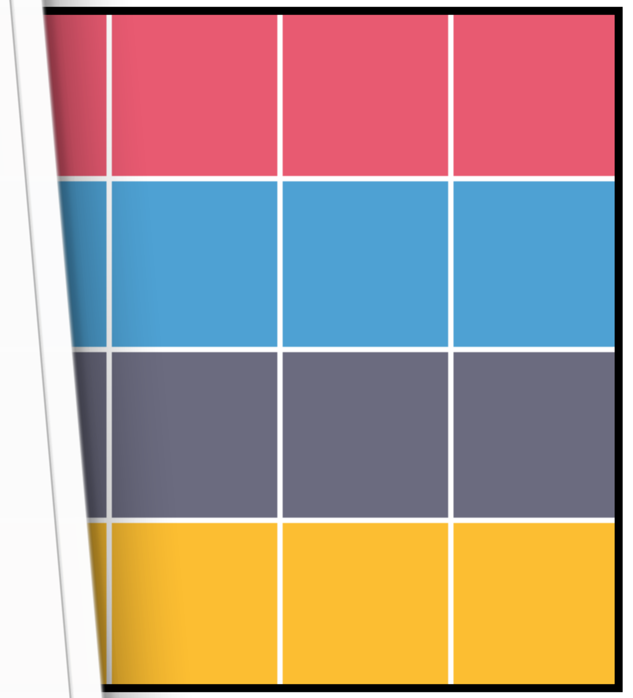
Jacob Feldman*
Department of Psychology
Center for Cognitive Science
Rutgers University
Piscataway, NJ 08854
jacob@ruccs.rutgers.edu

Abstract

We present an account of human concept learning—that is, learning of categories from examples—based on the principle of minimum description length (MDL). In support of this theory, we tested a wide range of two-dimensional concept types, including both regular (simple) and highly irregular (complex) structures, and found the MDL theory to give a good account of subjects' performance. This suggests that the *intrinsic complexity* of a concept (that is, its description length) systematically influences its learnability.

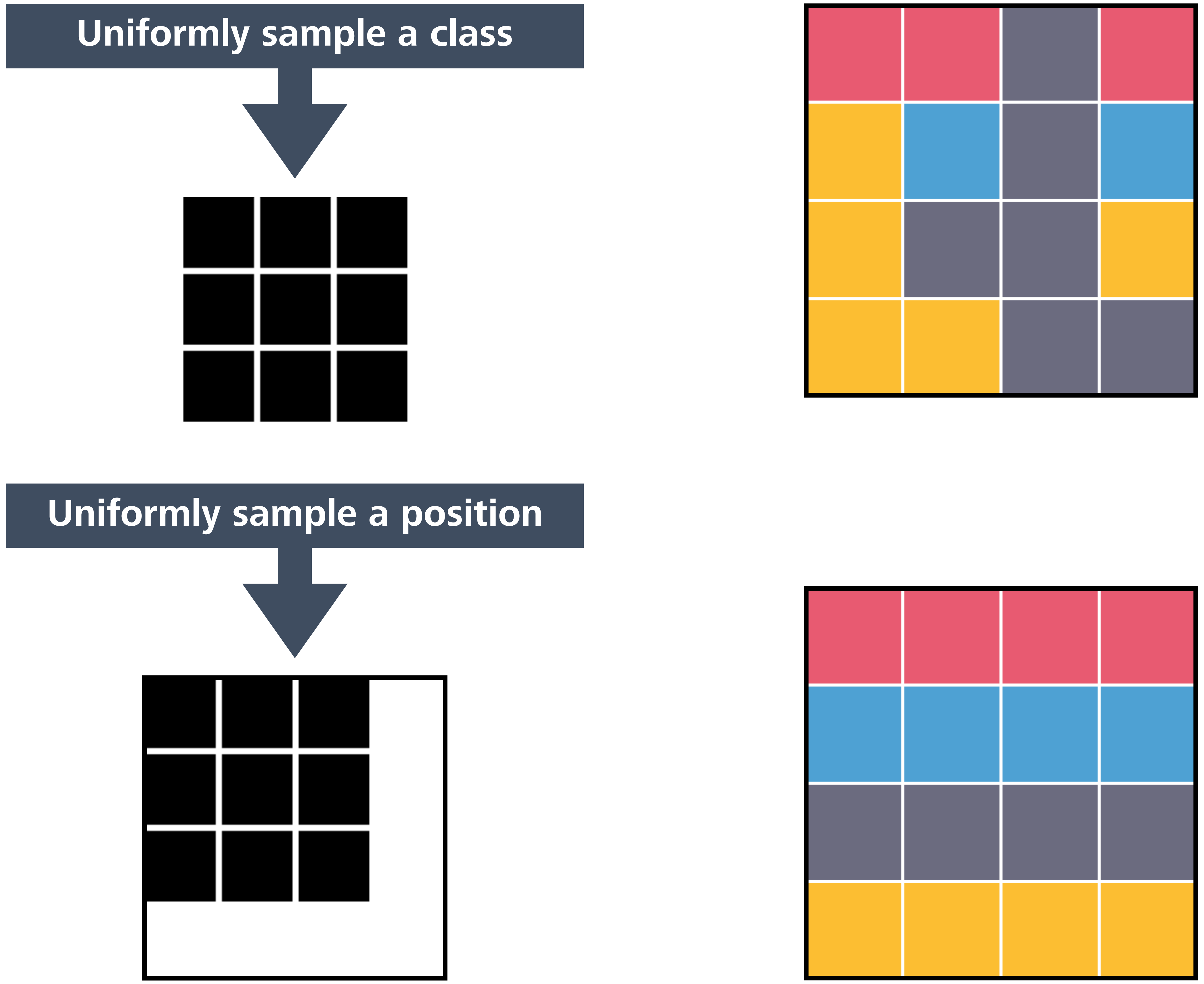
1 The Structure of Categories

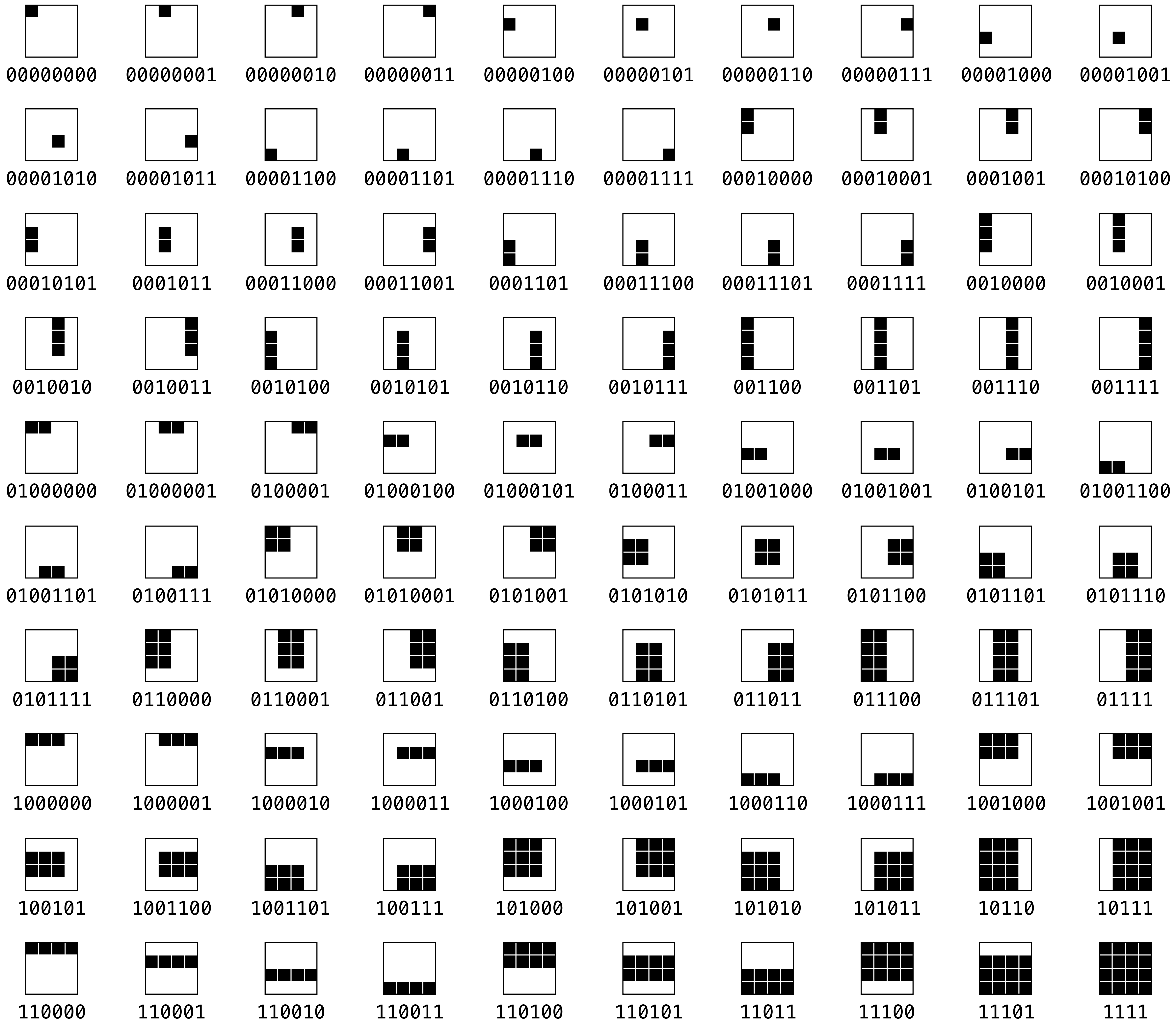
A number of different principles have been advanced to explain the manner in which humans learn to categorize objects. It has been variously suggested that the underlying principle might be the *similarity structure* of objects [1], the manipulability of *decision bound-*… While many of these theories are mathematically … inference [3][4]. … a range of experimental findings, …

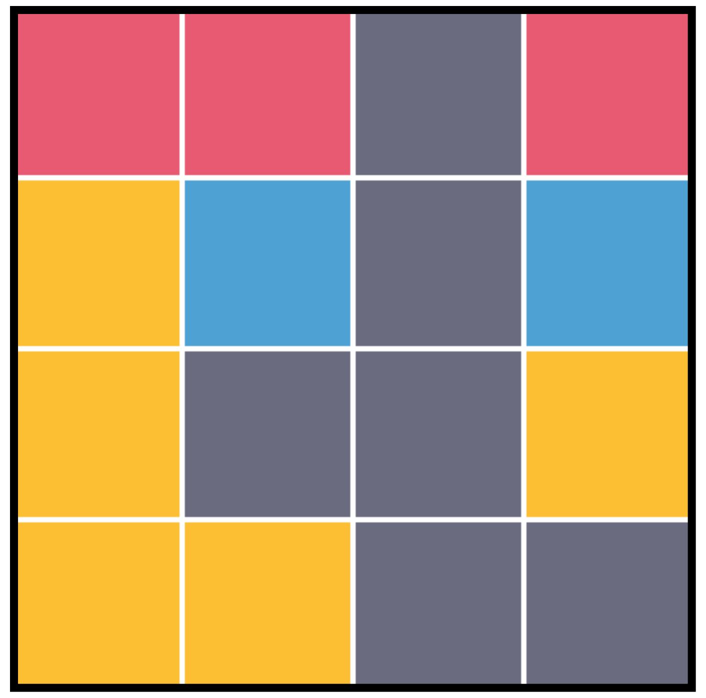


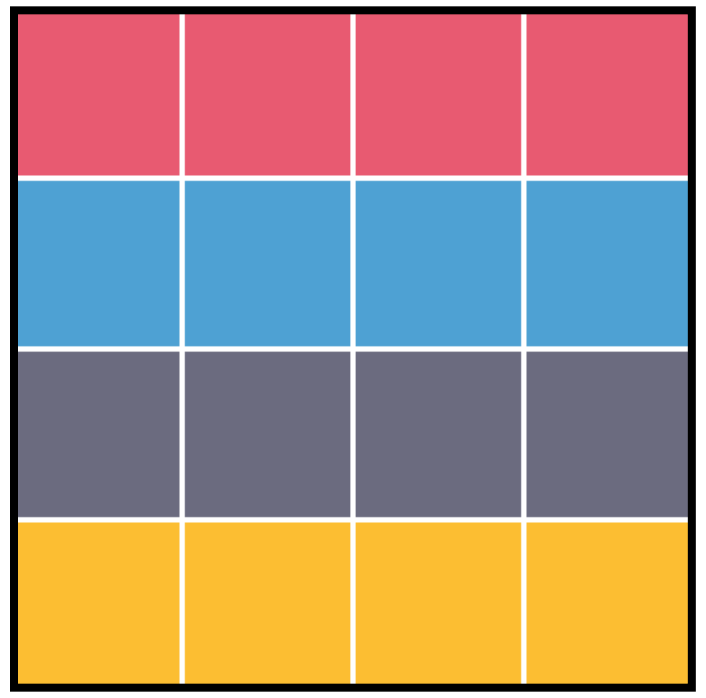
Computing DL(*L*): The rectangle code

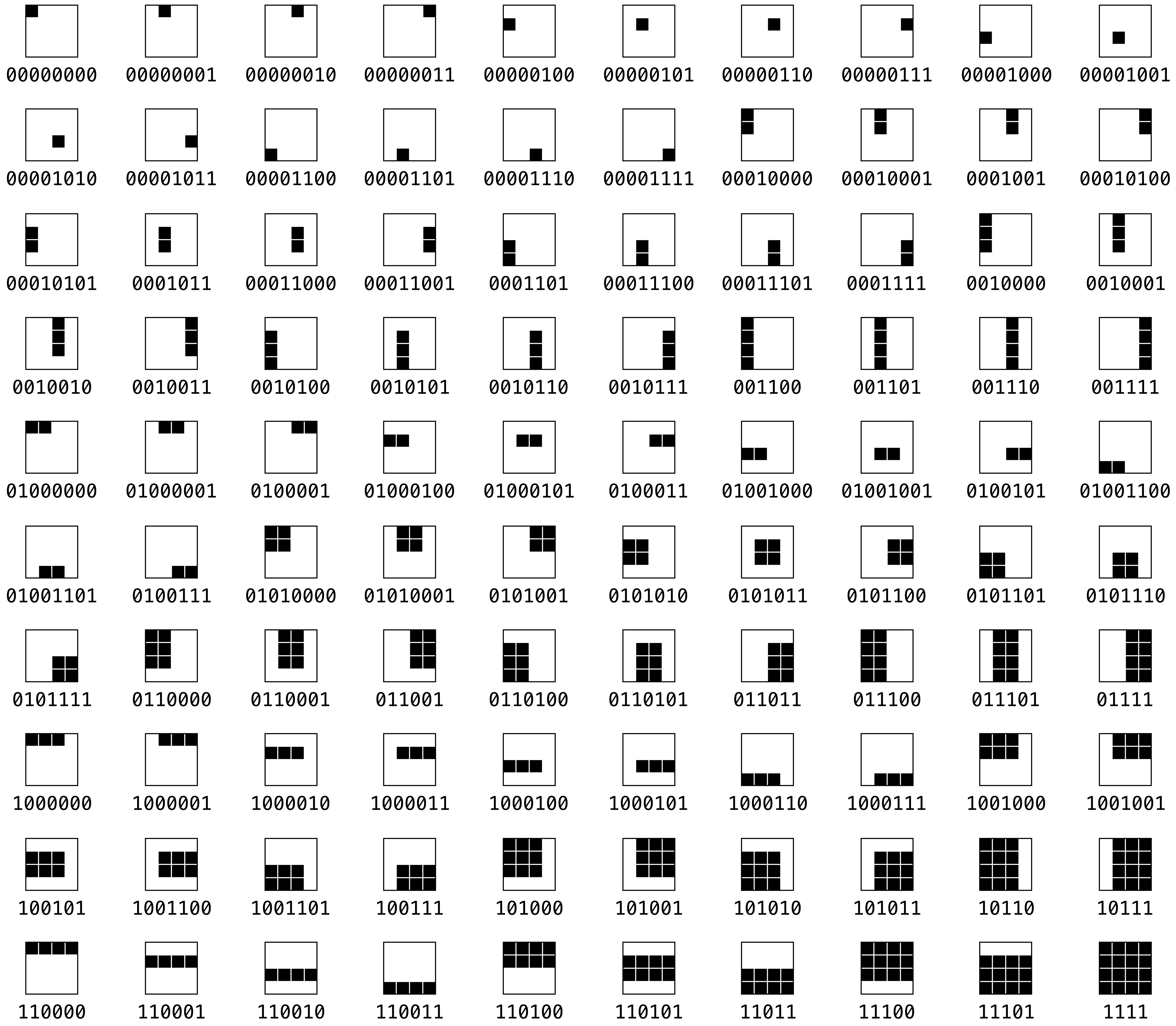
Class	Positions	Probability	Codelength	(bits)
1×1	16	1/16 × 1/16	−log 1/256	8.0
1×2	12	1/16 × 1/12	−log 1/192	7.58
1×3	8	1/16 × 1/8	−log 1/128	7.0
1×4	4	1/16 × 1/4	−log 1/64	6.0
2×1	12	1/16 × 1/12	−log 1/192	7.58
2×2	9	1/16 × 1/9	−log 1/144	7.17
2×3	6	1/16 × 1/6	−log 1/96	6.58
2×4	3	1/16 × 1/3	−log 1/48	5.58
3×1	8	1/16 × 1/8	−log 1/128	7.0
3×2	6	1/16 × 1/6	−log 1/96	6.58
3×3	4	1/16 × 1/4	−log 1/64	6.0
3×4	2	1/16 × 1/2	−log 1/32	5.0
4×1	4	1/16 × 1/4	−log 1/64	6.0
4×2	3	1/16 × 1/3	−log 1/48	5.58
4×3	2	1/16 × 1/2	−log 1/32	5.0
4×4	1	1/16 × 1/1	−log 1/16	4.0

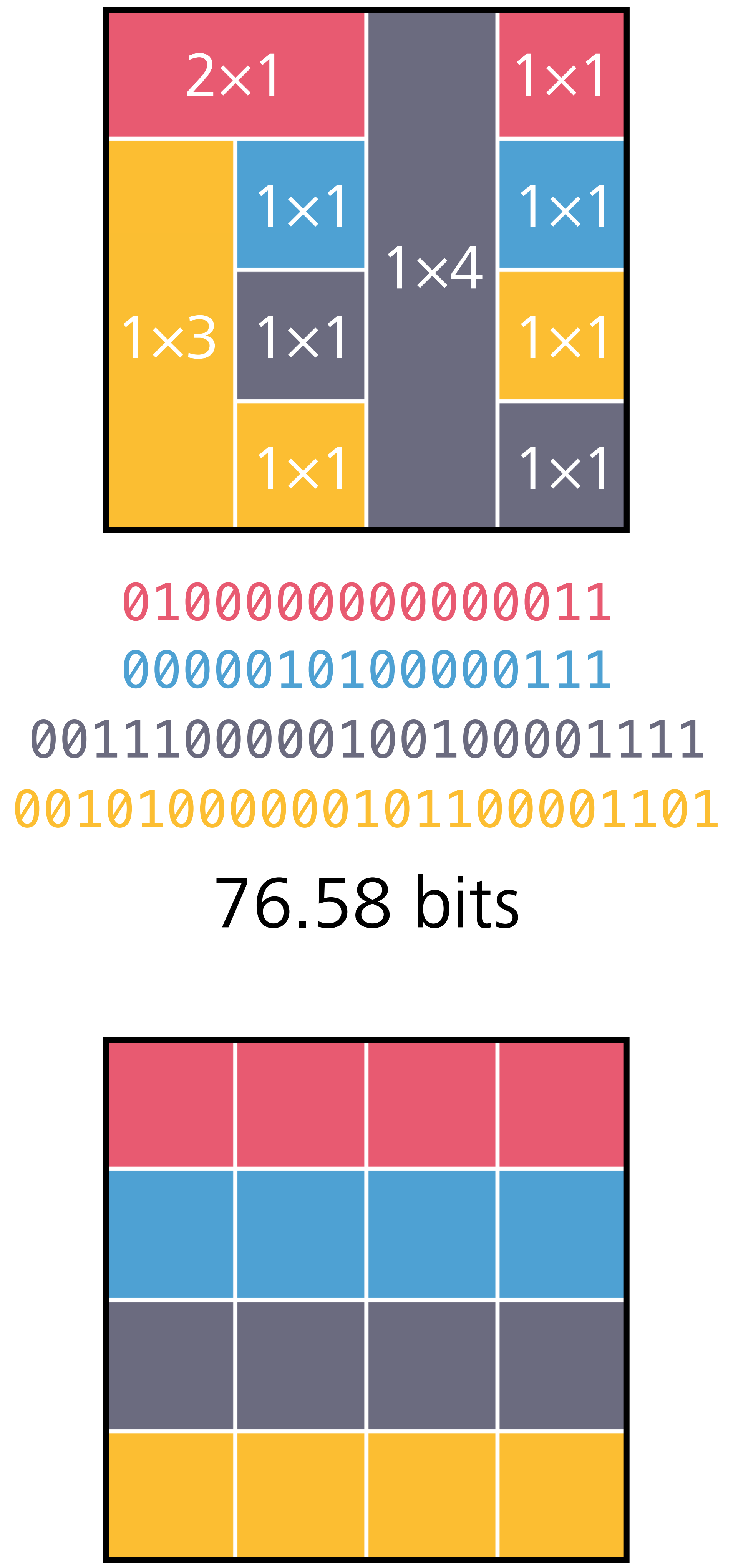


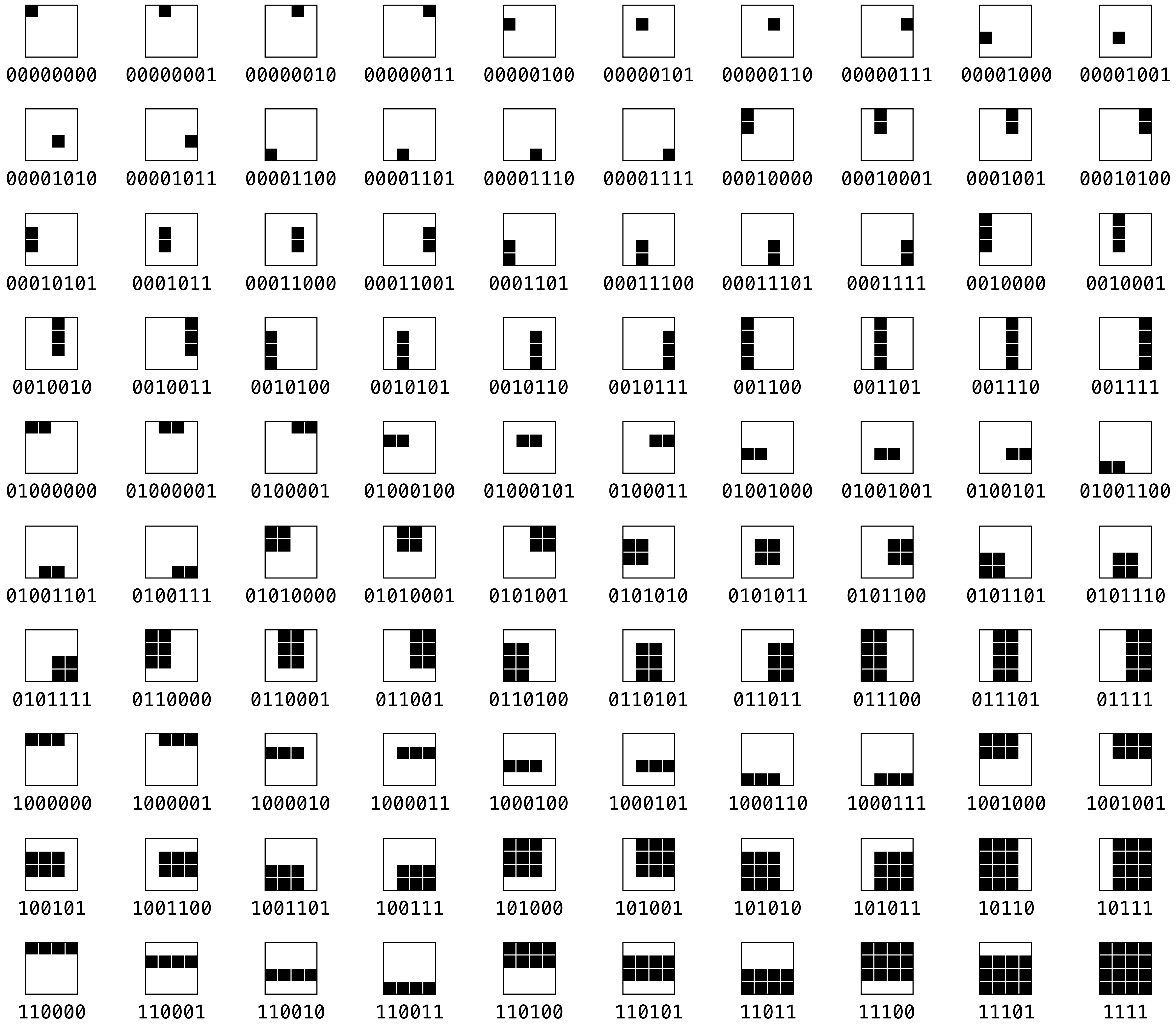

00000000 00000001 00000010 00000011 00000100 00000101 00000110 00000111 00001000 00001001
00001010 00001011 00001100 00001101 00001110 00001111 00010000 00010001 0001001 00010100
00010101 0001011 00011000 00011001 0001101 00011100 00011101 0001111 0010000 0010001
0010010 0010011 0010100 0010101 0010110 0010111 001100 001101 001110 001111
01000000 01000001 0100001 01000100 01000101 0100011 01001000 01001001 0100101 01001100
01001101 0100111 01010000 01010001 0101001 0101010 0101011 0101100 0101101 0101110
0101111 0110000 0110001 011001 0110100 0110101 011011 011100 011101 01111
1000000 1000001 1000010 1000011 1000100 1000101 1000110 1000111 1001000 1001001
100101 1001100 1001101 100111 101000 101001 101010 101011 10110 10111
110000 110001 110010 110011 110100 110101 11011 11100 11101 1111

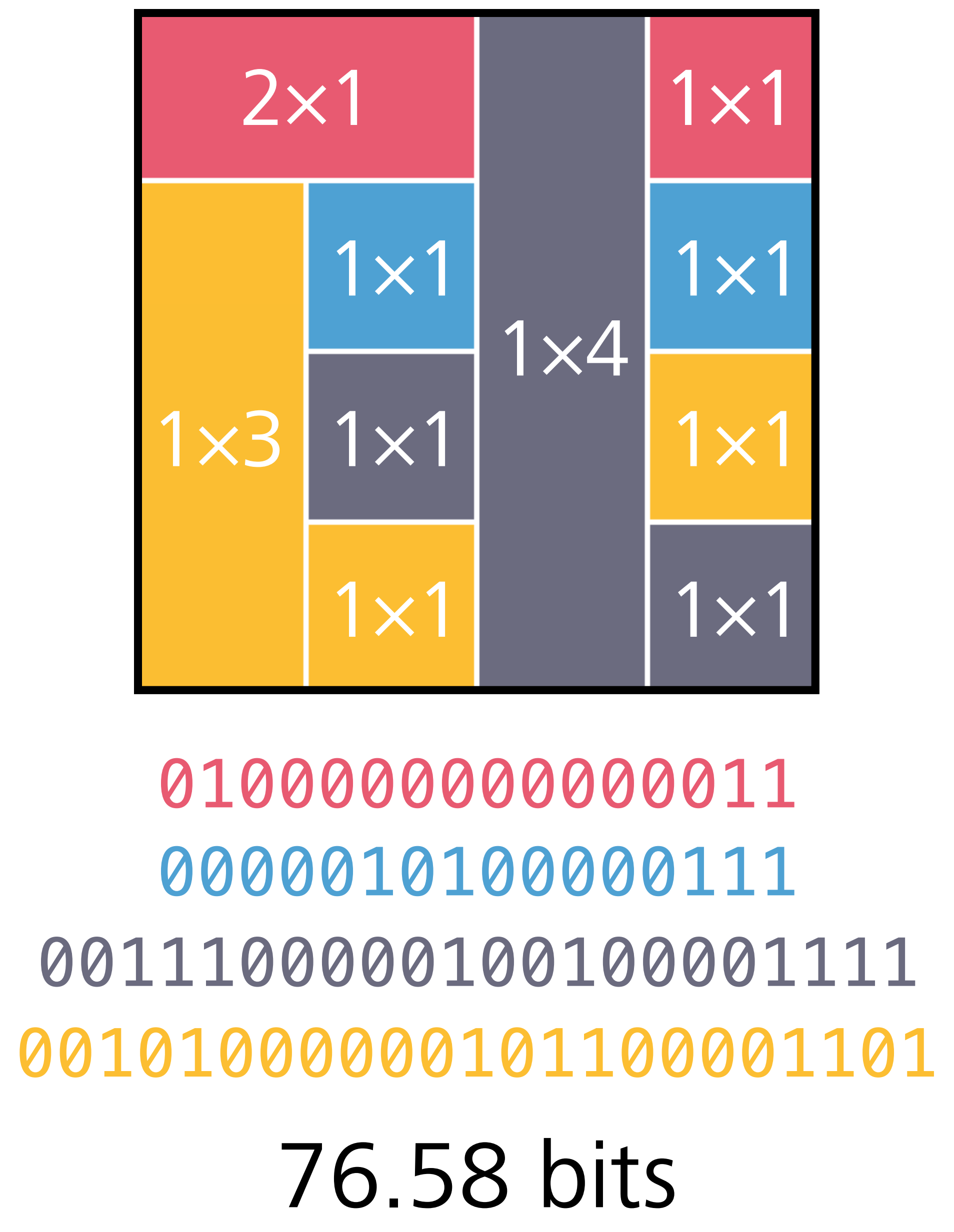


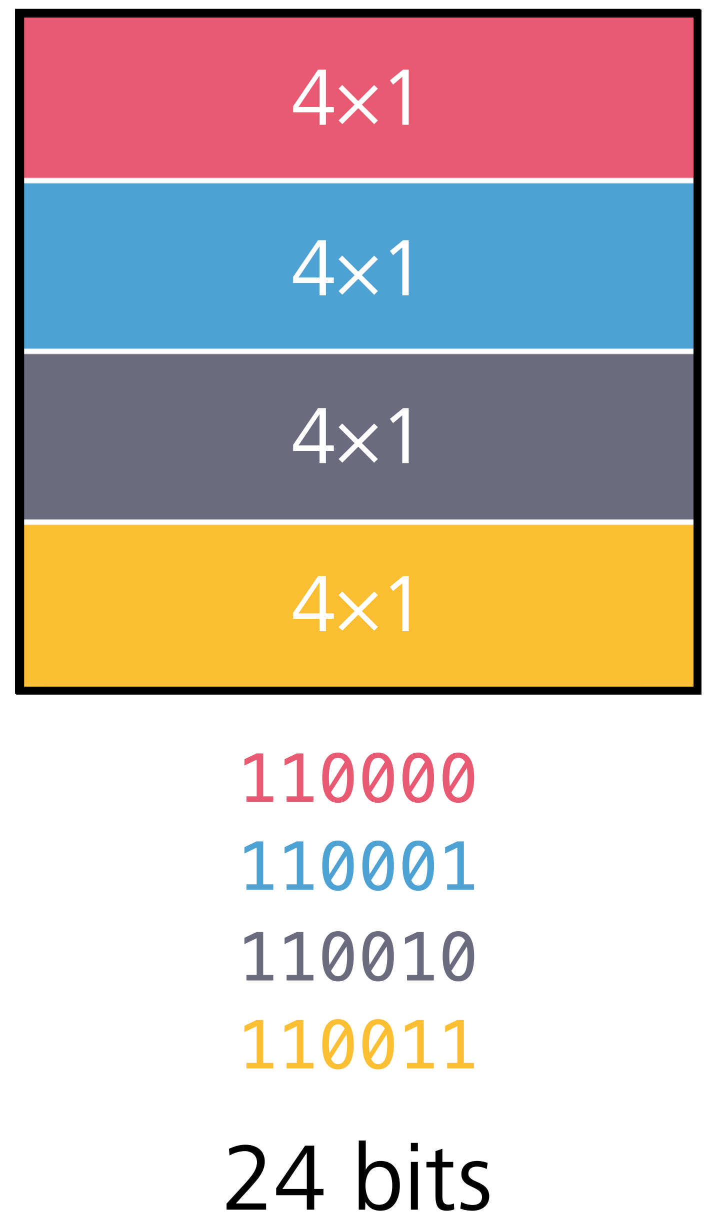



00000000 00000001 00000010 00000011 00000100 00000101 00000110 00000111 00001000 00001001
00001010 00001011 00001100 00001101 00001110 00001111 00010000 00010001 0001001 00010100
00010101 0001011 00011000 00011001 0001101 00011100 00011101 0001111 0010000 0010001
0010010 0010011 0010100 0010101 0010110 0010111 001100 001101 001110 001111
01000000 01000001 0100001 01000100 01000101 0100011 01001000 01001001 0100101 01001100
01001101 0100111 01010000 01010001 0101001 0101010 0101011 0101100 0101101 0101110
0101111 0110000 0110001 011001 0110100 0110101 011011 011100 011101 01111
1000000 1000001 1000010 1000011 1000100 1000101 1000110 1000111 1001000 1001001
100101 1001100 1001101 100111 101000 101001 101010 101011 10110 10111
110000 110001 110010 110011 110100 110101 11011 11100 11101 1111

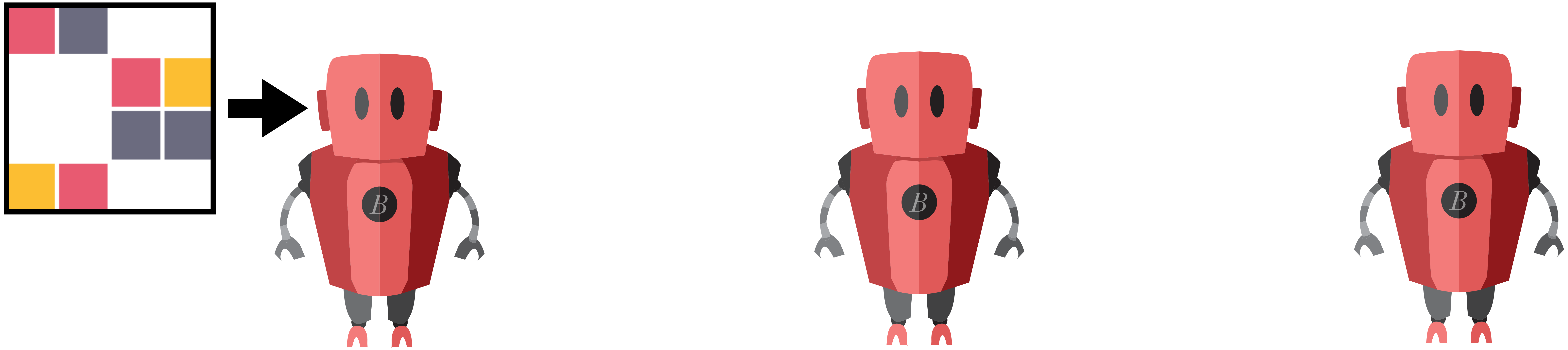

2×1
1×1
1×4
1×3
1×1
1×1
1×1
1×1
1×1
1×1
0100000000000011
0000010100000111
00111000001001000001111
0010100000010110000110 1
76.58 bits


00000000 00000001 00000010 00000011 00000100 00000101 00000110 00000111 00001000 00001001
00001010 00001011 00001100 00001101 00001110 00001111 00010000 00010001 0001001 00010100
00010101 0001011 00011000 00011001 0001101 00011100 00011101 0001111 0010000 0010001
0010010 0010011 0010100 0010101 0010110 0010111 001100 001101 001110 001111
01000000 01000001 0100001 01000100 01000101 0100011 01001000 01001001 0100101 01001100
01001101 0100111 01010000 01010001 0101001 0101010 0101011 0101100 0101101 0101110
0101111 0110000 0110001 011001 0110100 0110101 011011 011100 011101 01111
1000000 1000001 1000010 1000011 1000100 1000101 1000110 1000111 1001000 1001001
100101 1001100 1001101 100111 101000 101001 101010 101011 10110 10111
110000 110001 110010 110011 110100 110101 11011 11100 11101 1111

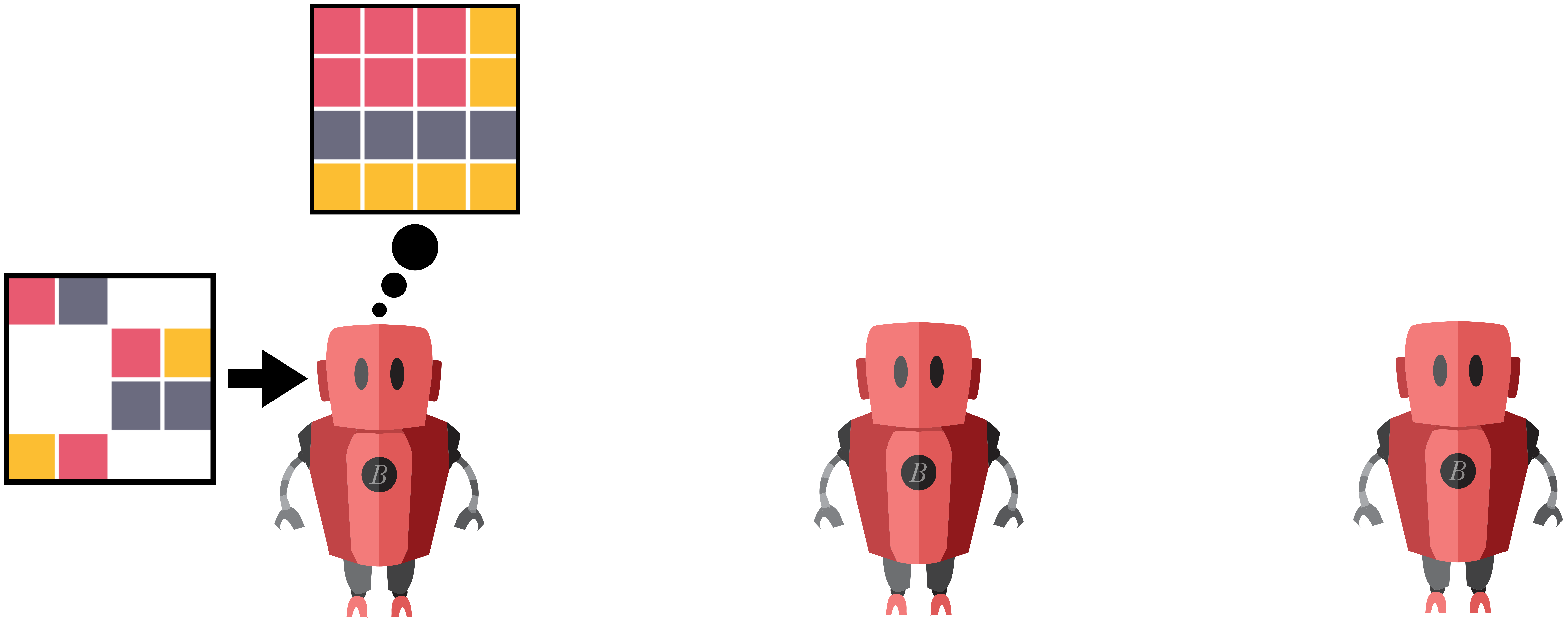

2×1
1×1
1×1
1×4
1×1
1×3
1×1
1×1
1×1
1×1
0100000000000011
0000010100000111
00111000001001000001111
00101000000101100001101
76.58 bits


4×1
4×1
4×1
4×1
110000
110001
110010
110011
24 bits

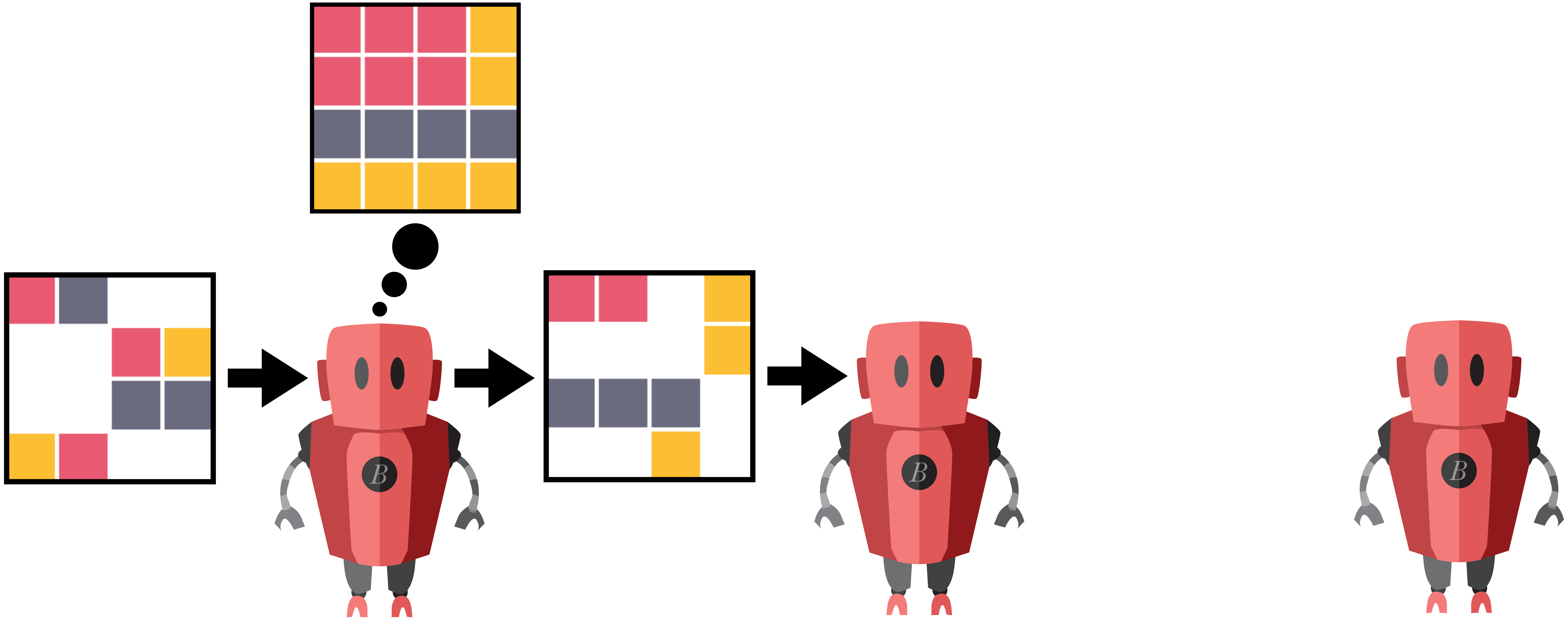
Bayesian iterated learning under a simplicity prior



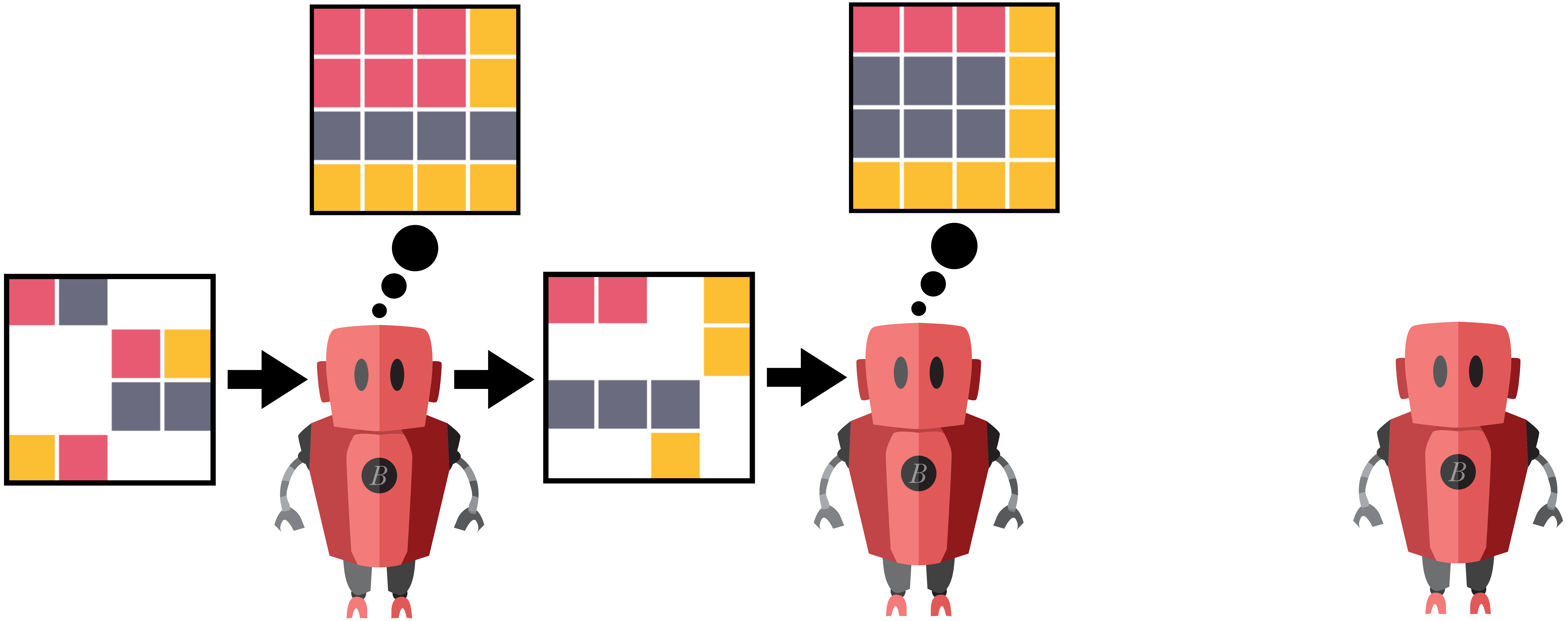
Bayesian iterated learning under a simplicity prior



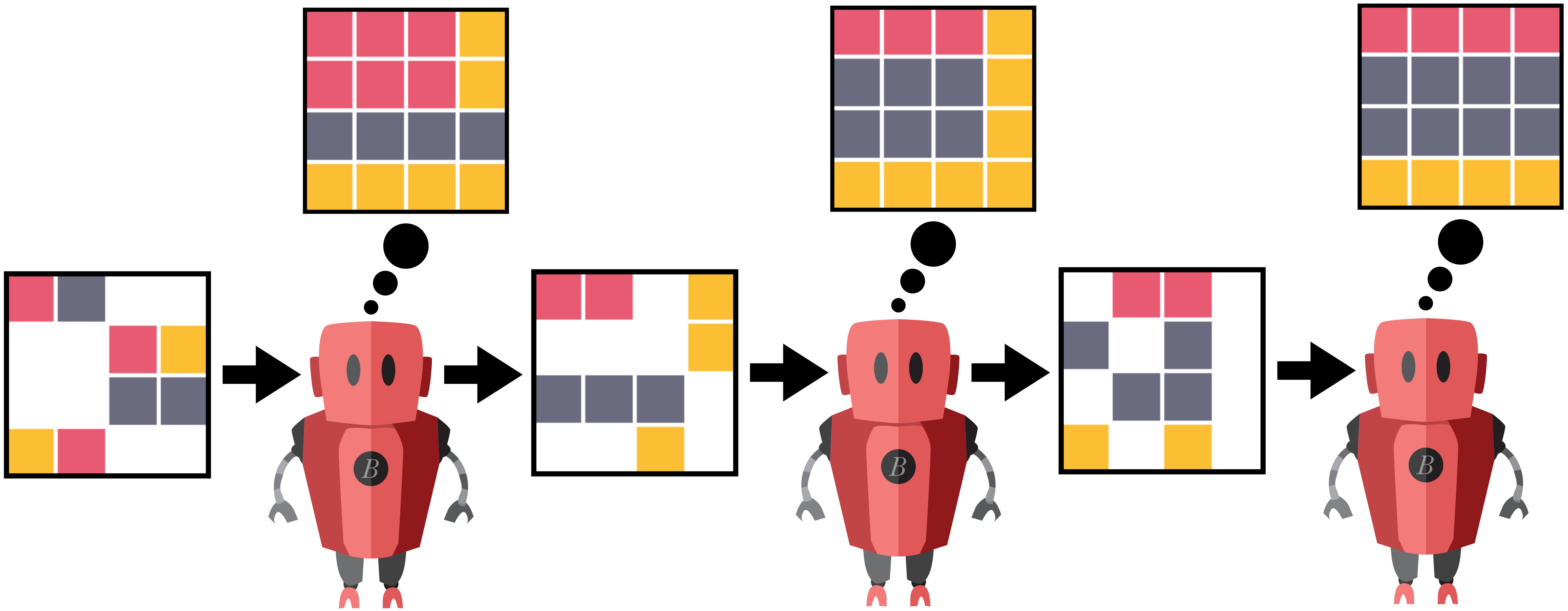
Bayesian iterated learning under a simplicity prior



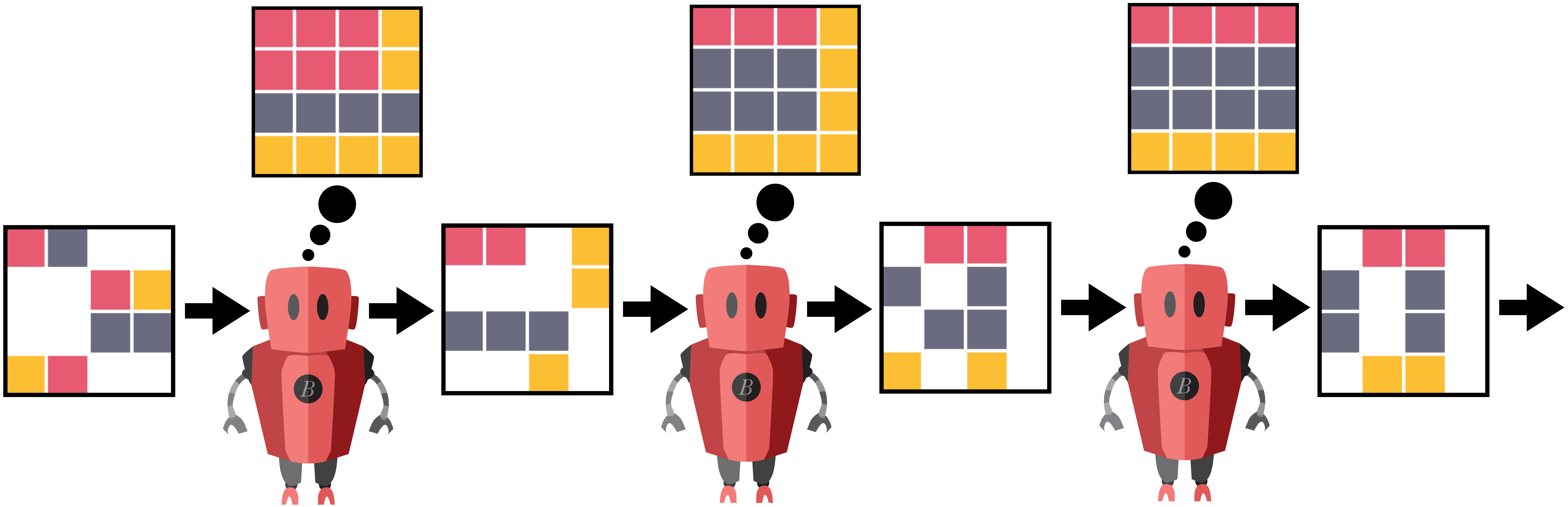
Bayesian iterated learning under a simplicity prior



Bayesian iterated learning under a simplicity prior

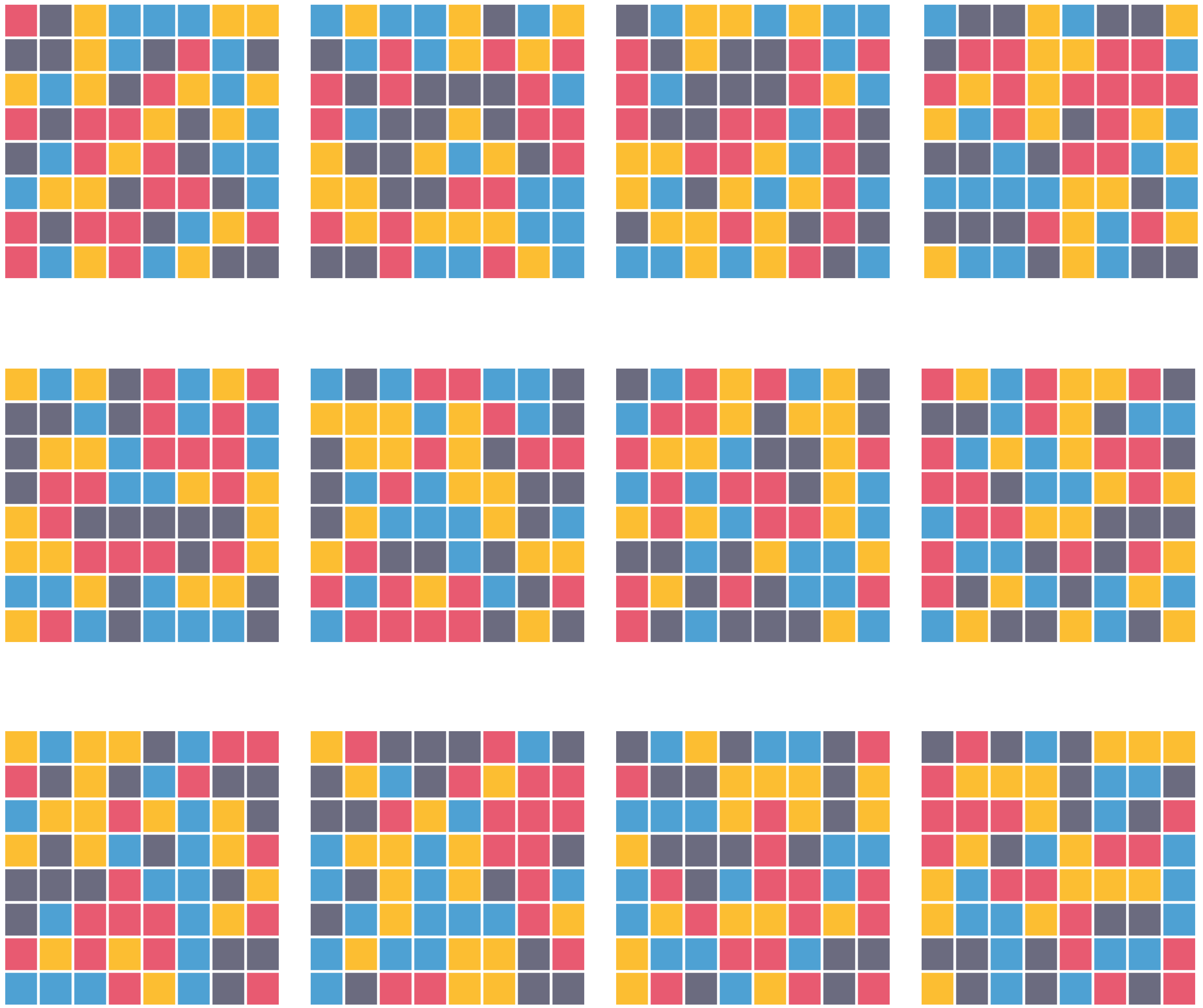


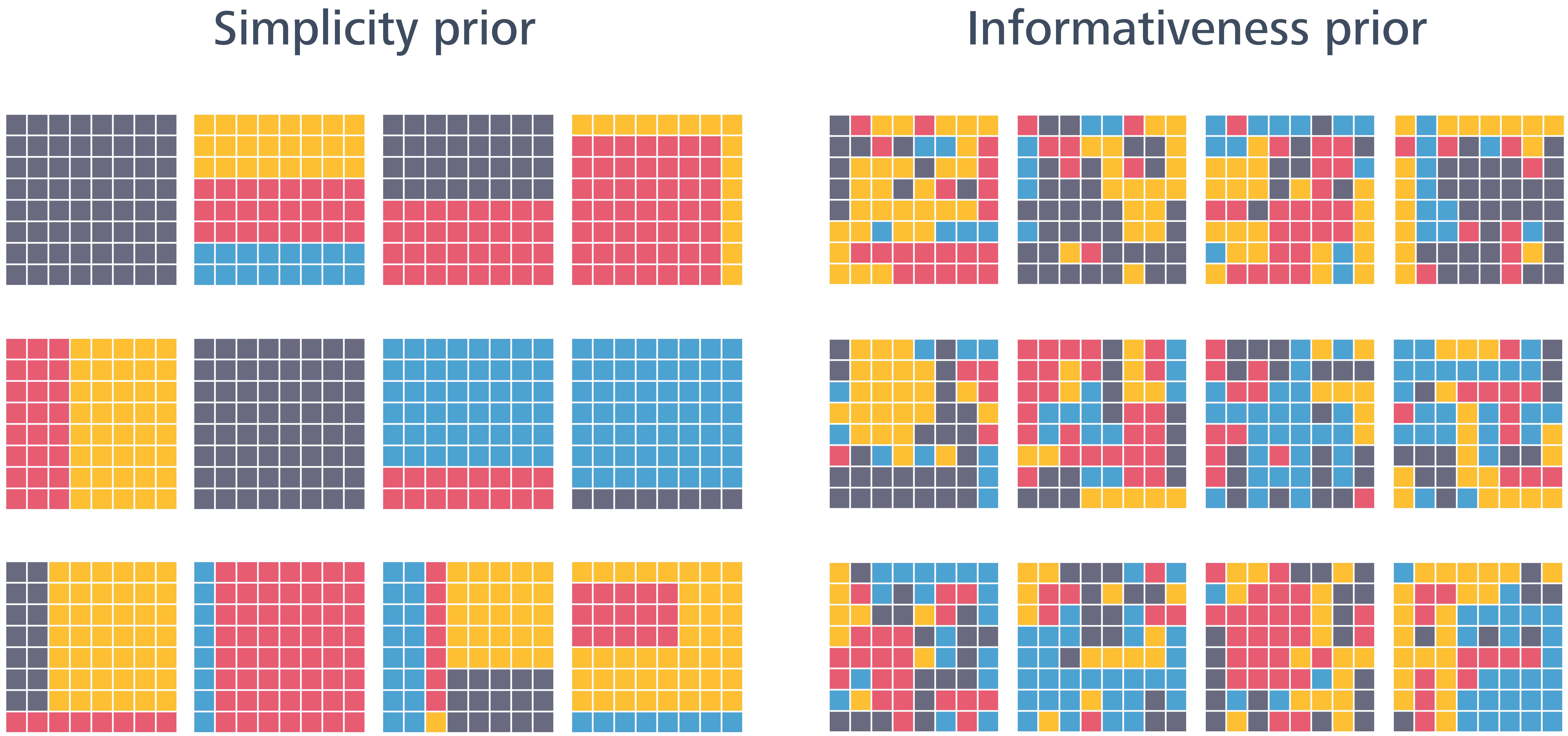
Bayesian iterated learning under a simplicity prior



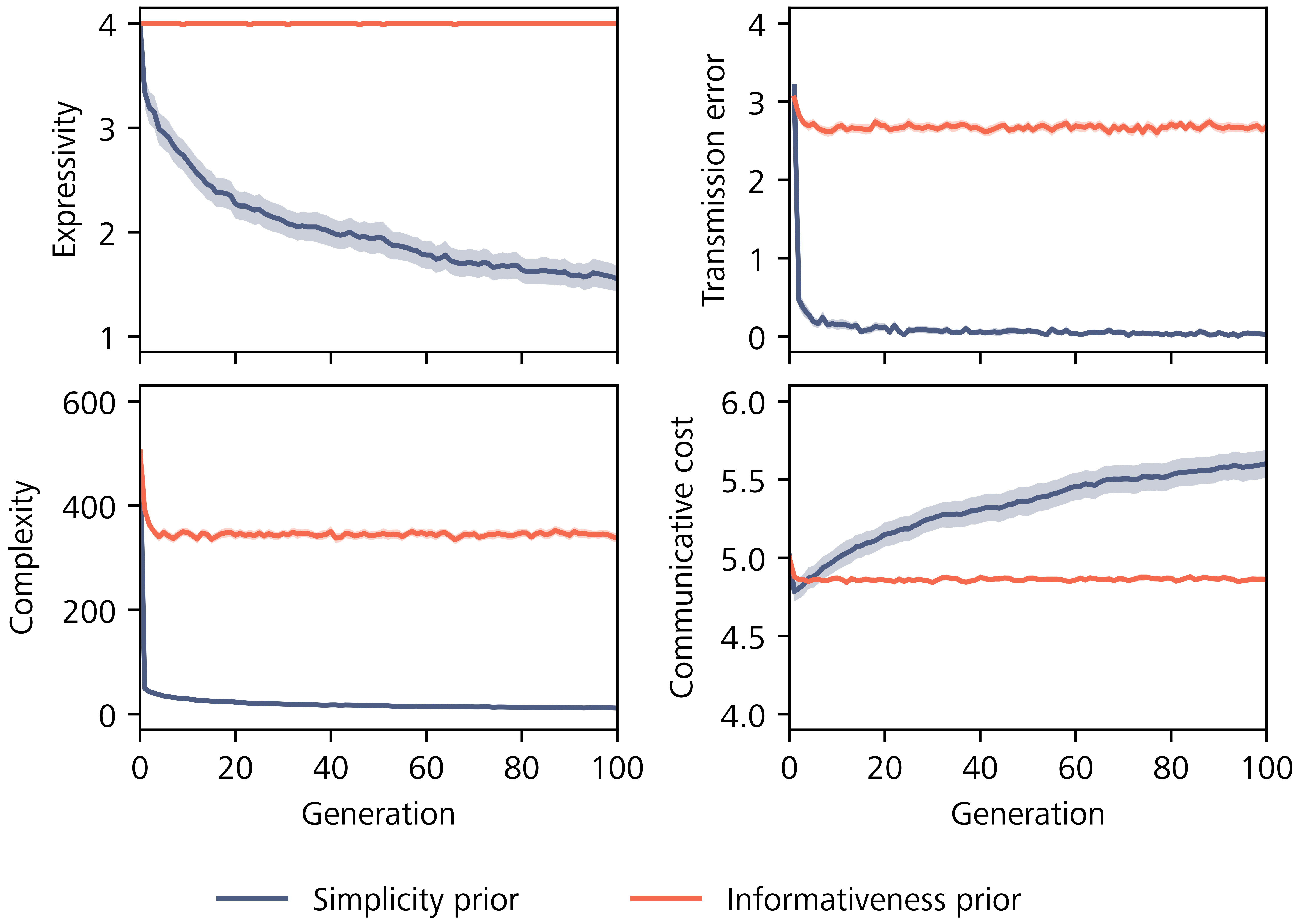
Simplicity prior

Informativeness prior





Model results



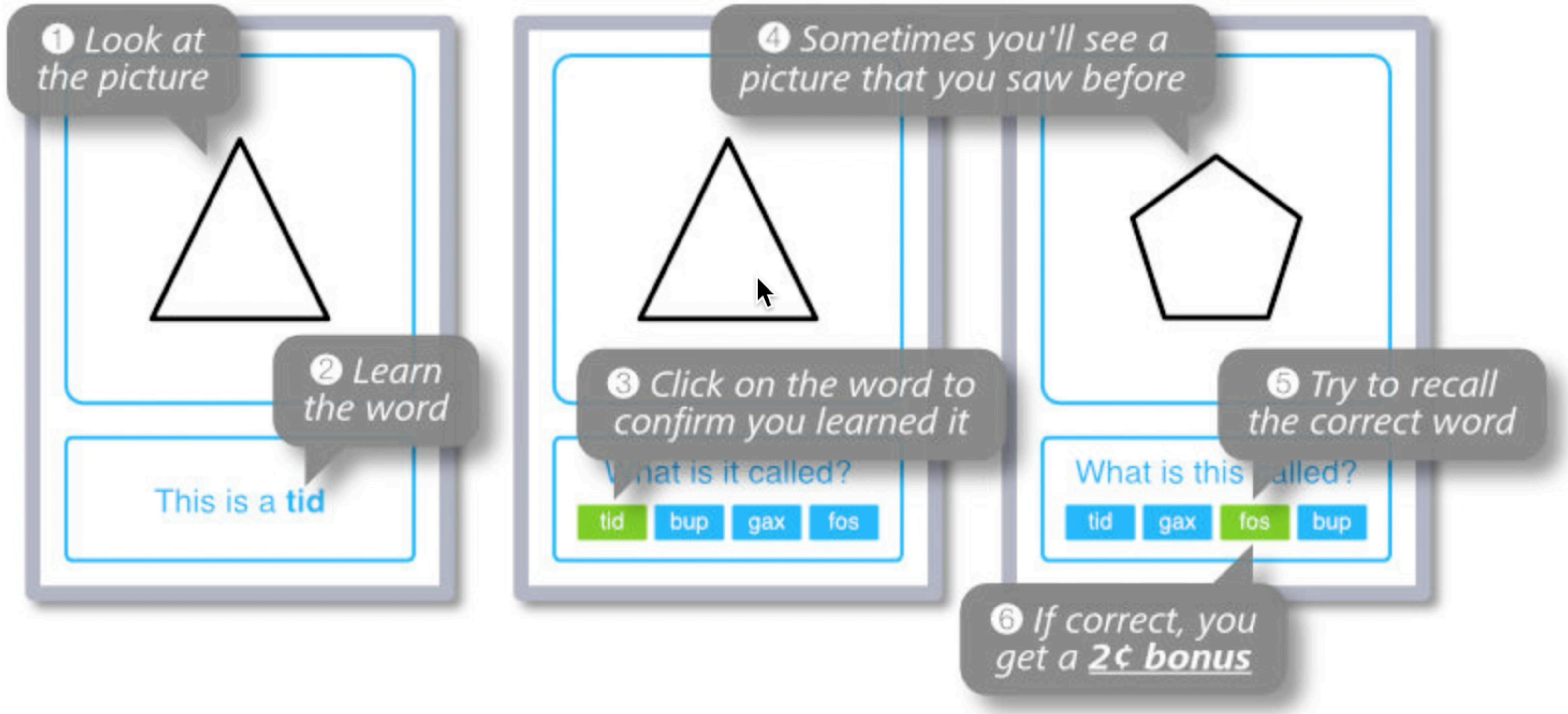
Experiment

localhost/~jon/shepard/

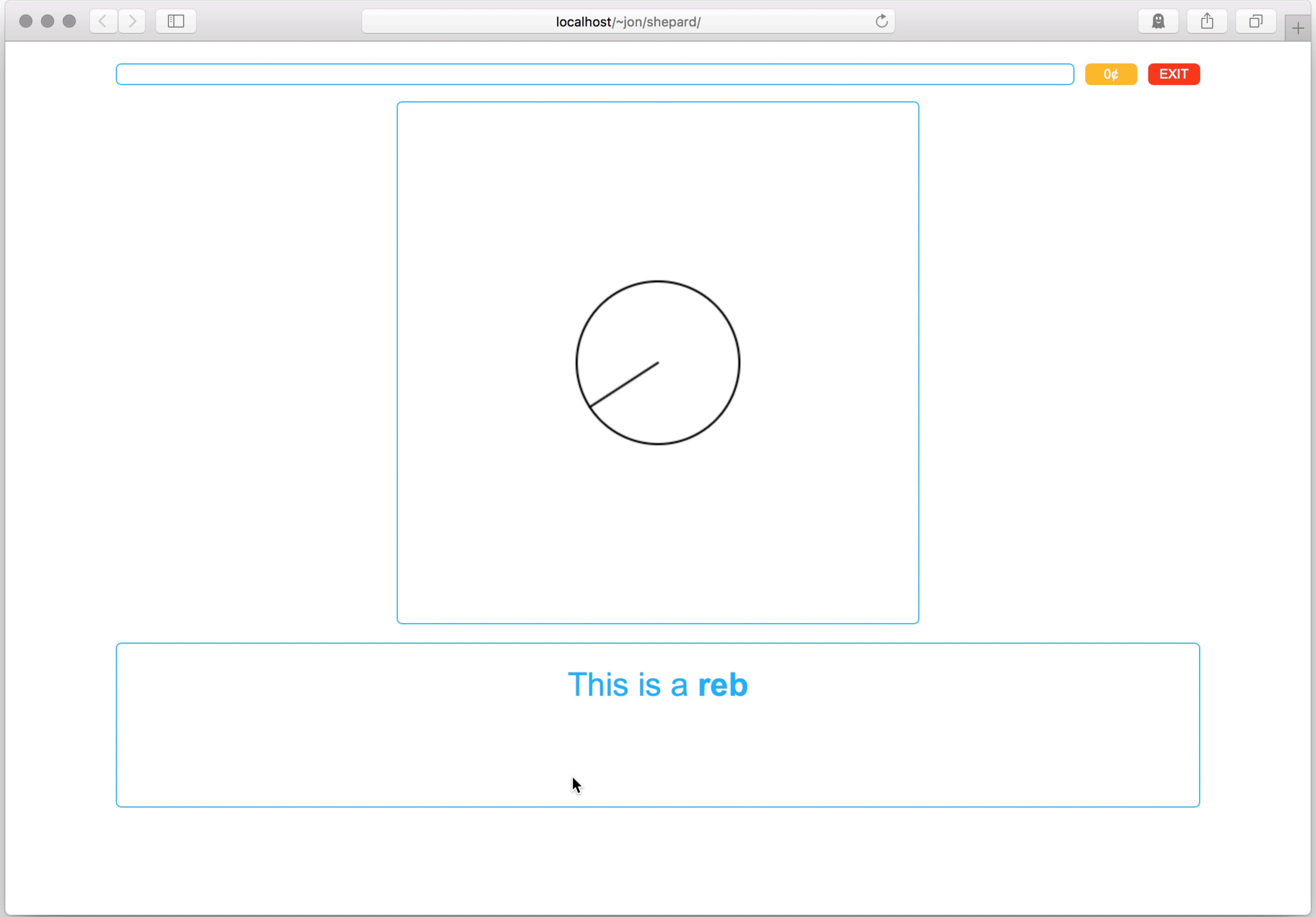
Stage 1: Training

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/
0¢
EXIT
This is a **reb**

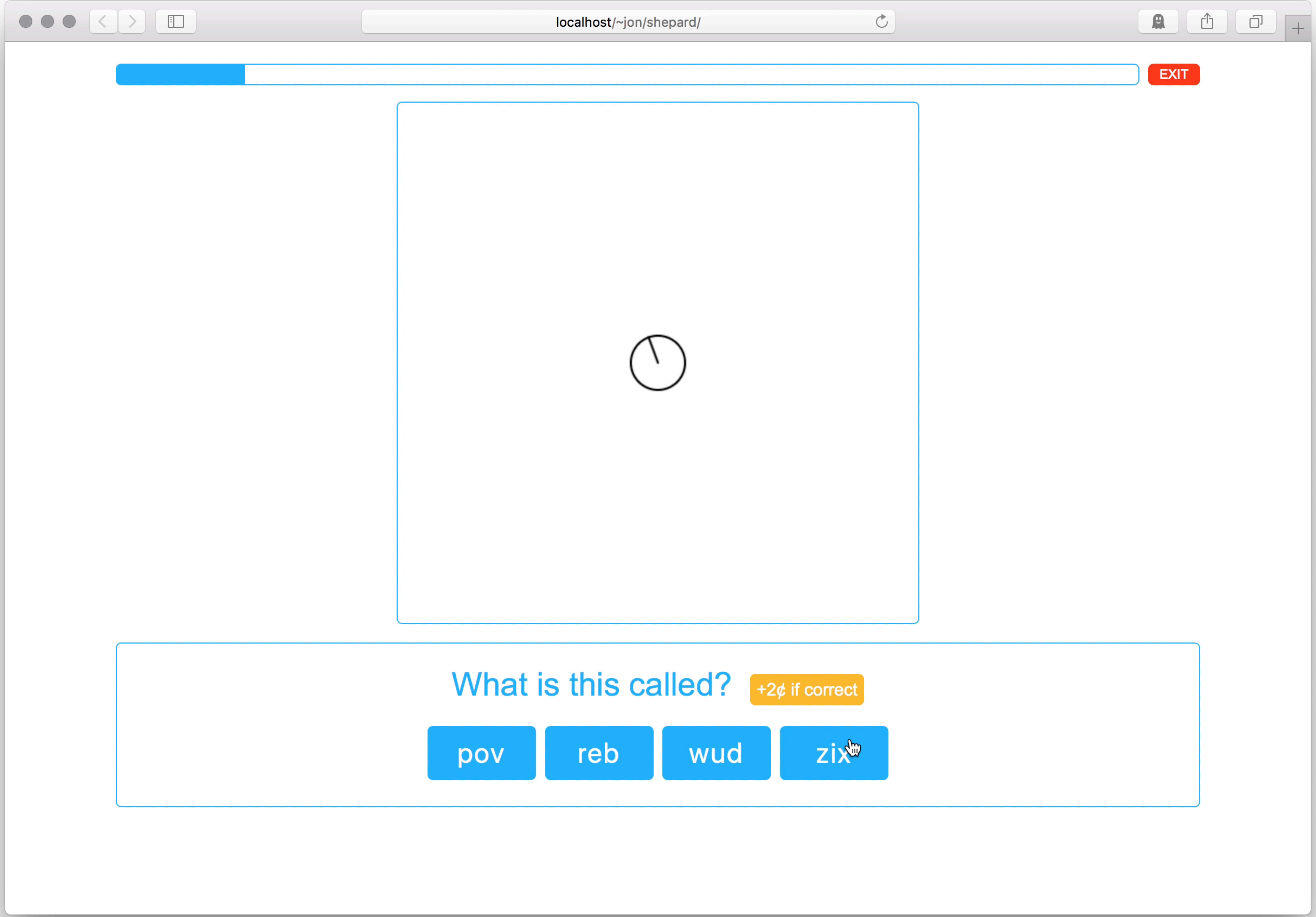
localhost/~jon/shepard/

Stage 2: Test

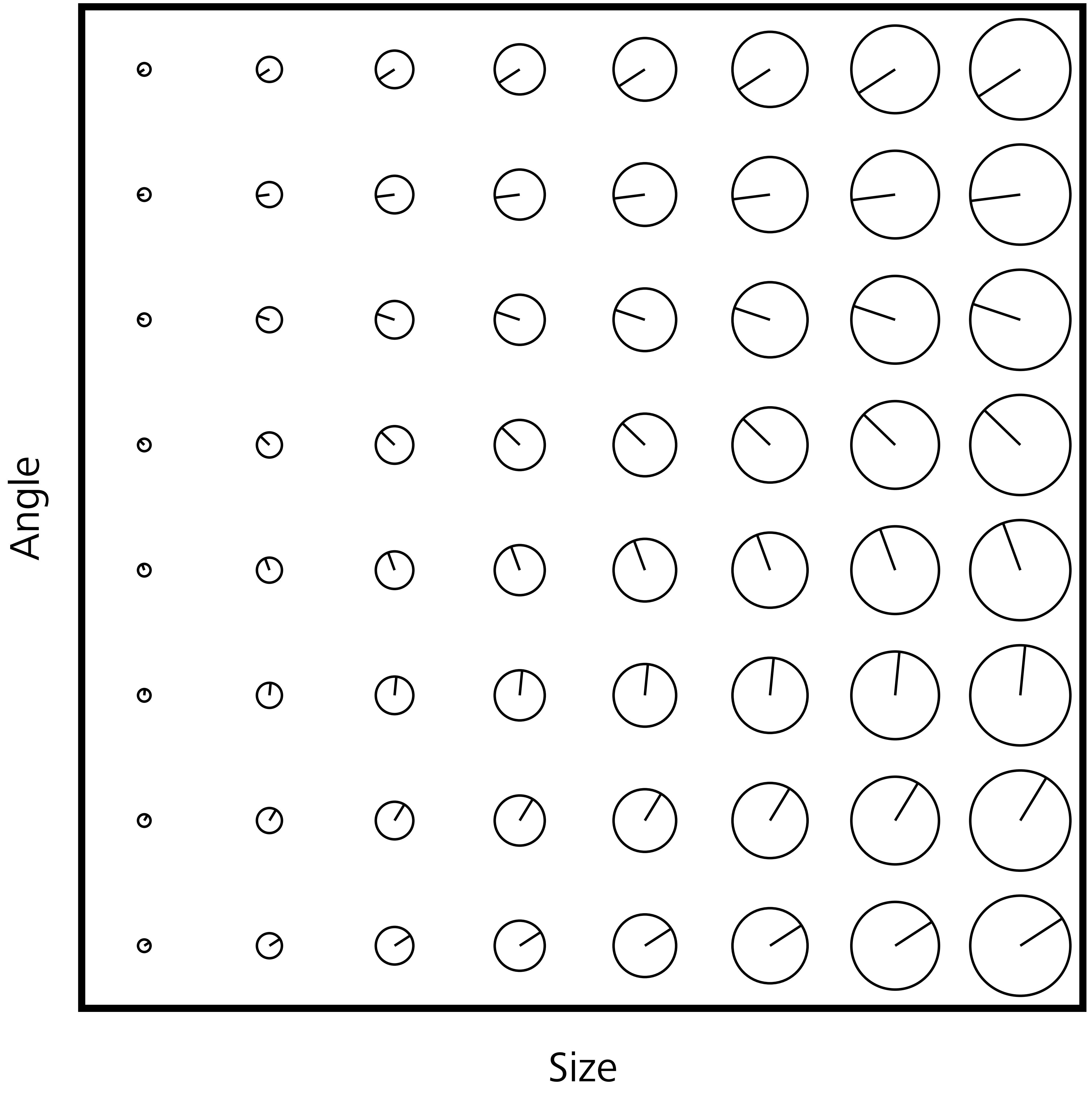
5 minutes

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 not tell you if you are correct or incorrect**. You will find out at the end how many you got correct.

START

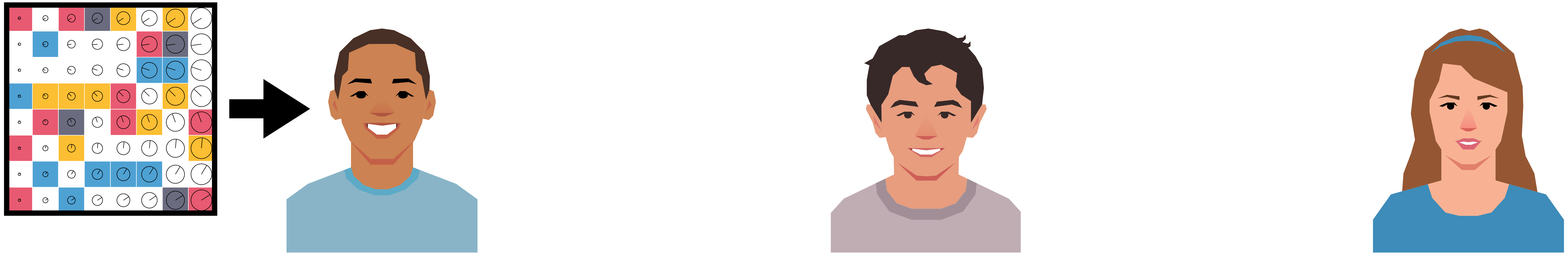

localhost/~jon/shepard/
EXIT
What is this called?
+2¢ if correct
pov
reb
wud
zix

Experimental stimuli

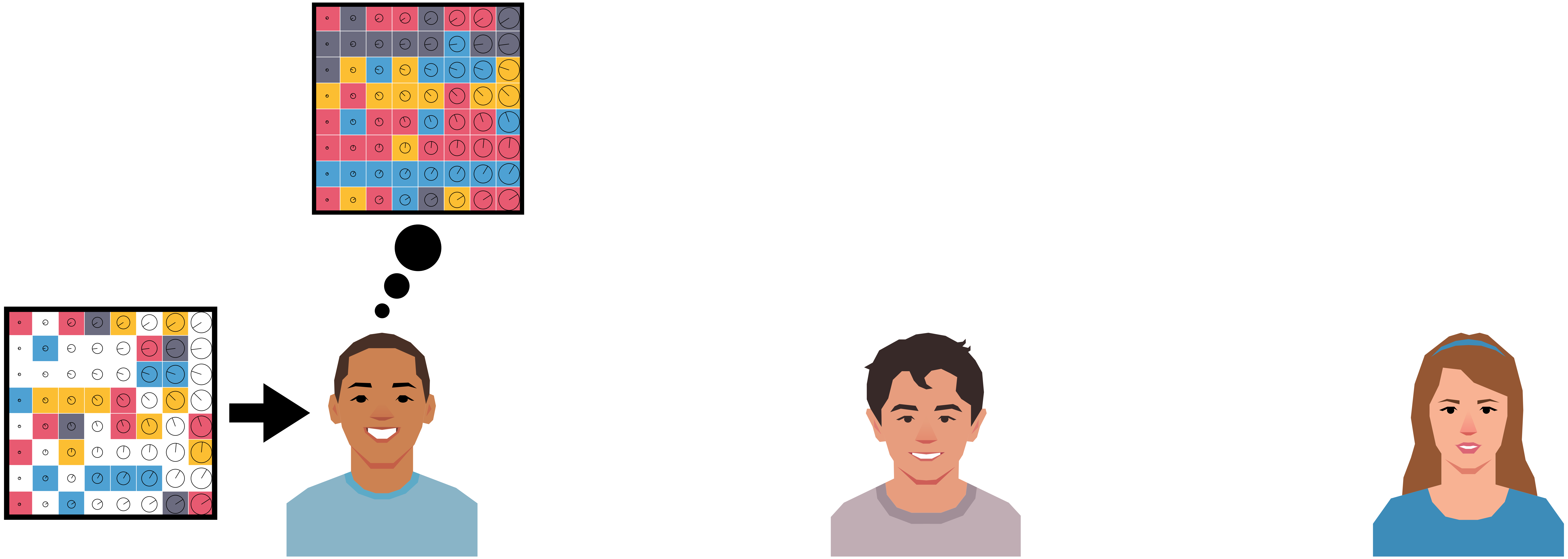


Set	Labels			
1	*pov*	*reb*	*wud*	*zix*
2	*gex*	*juf*	*vib*	*wop*
3	*buv*	*jef*	*pid*	*zox*
4	*fod*	*jes*	*wix*	*zuv*

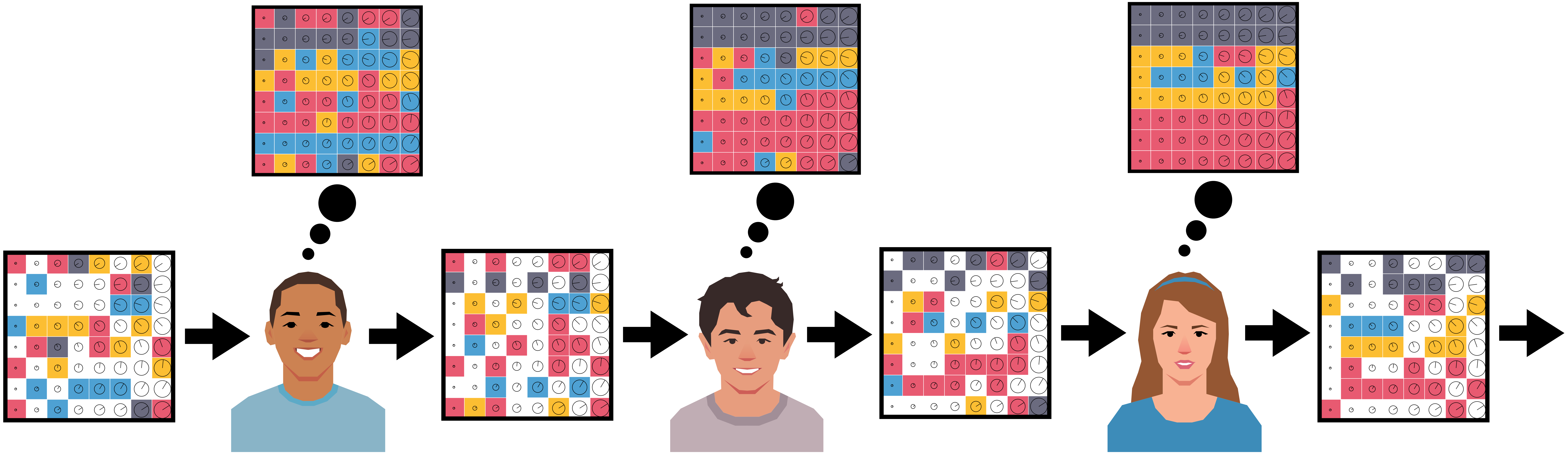
Iterated learning with humans

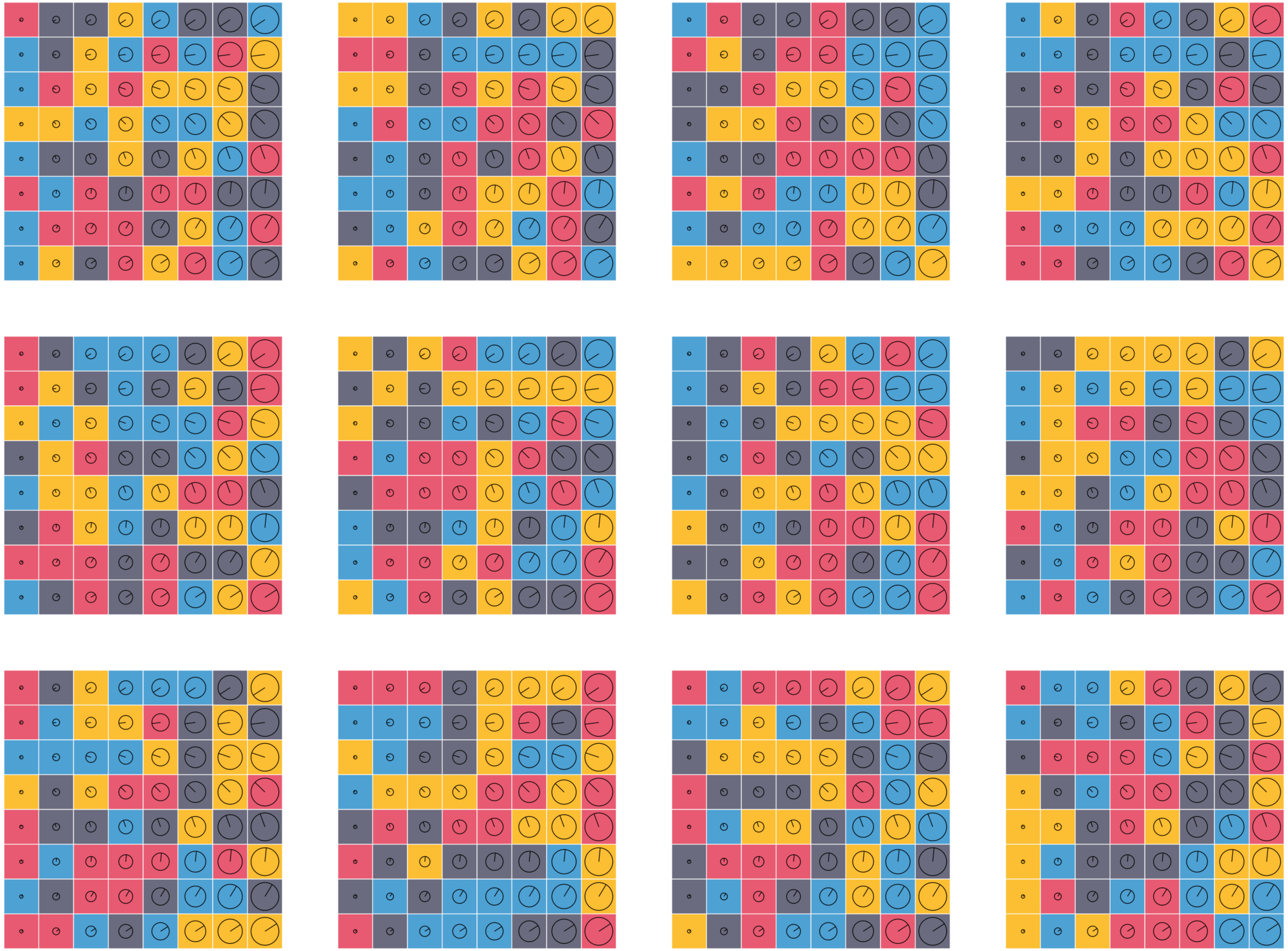


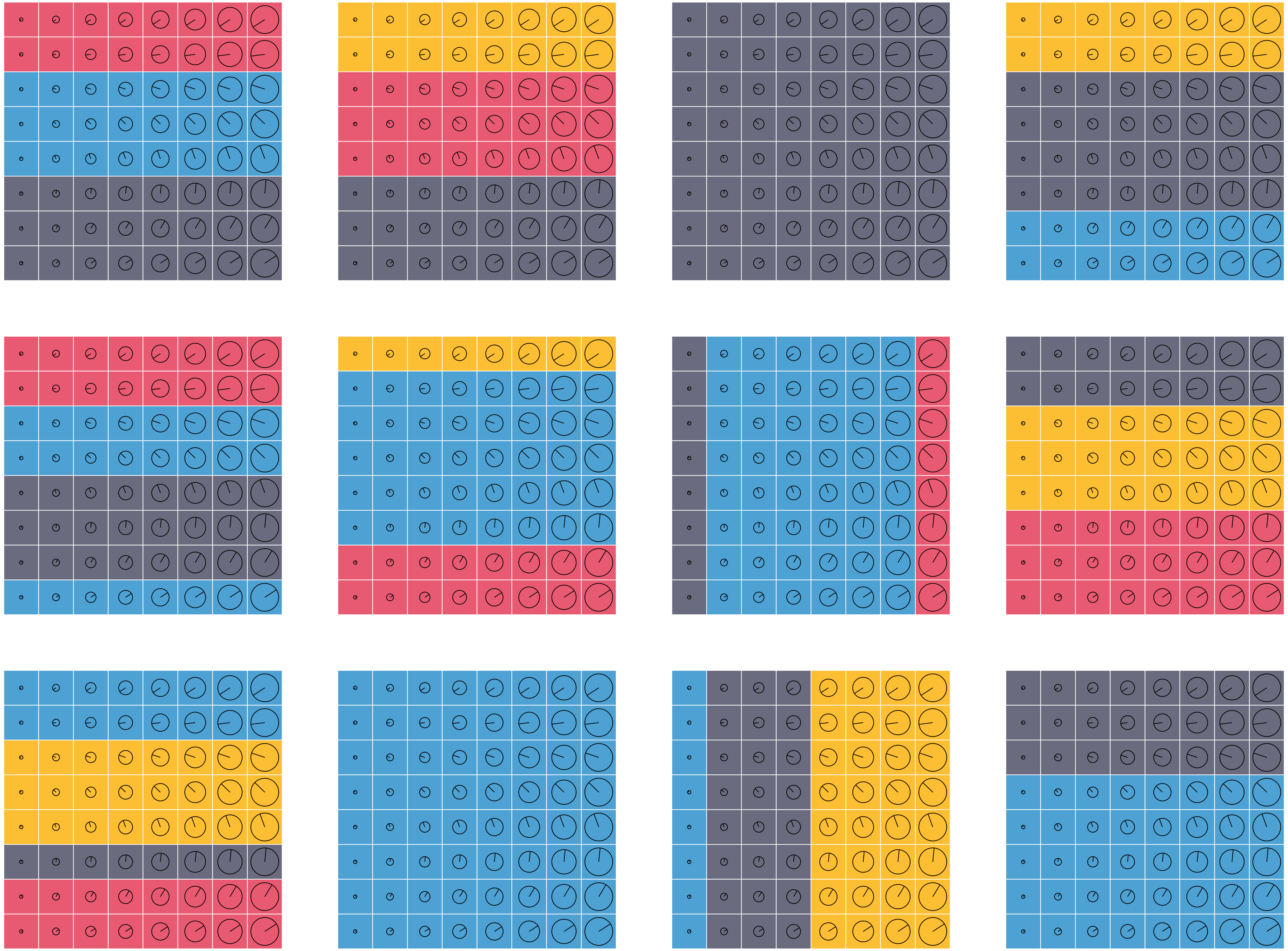
Iterated learning with humans



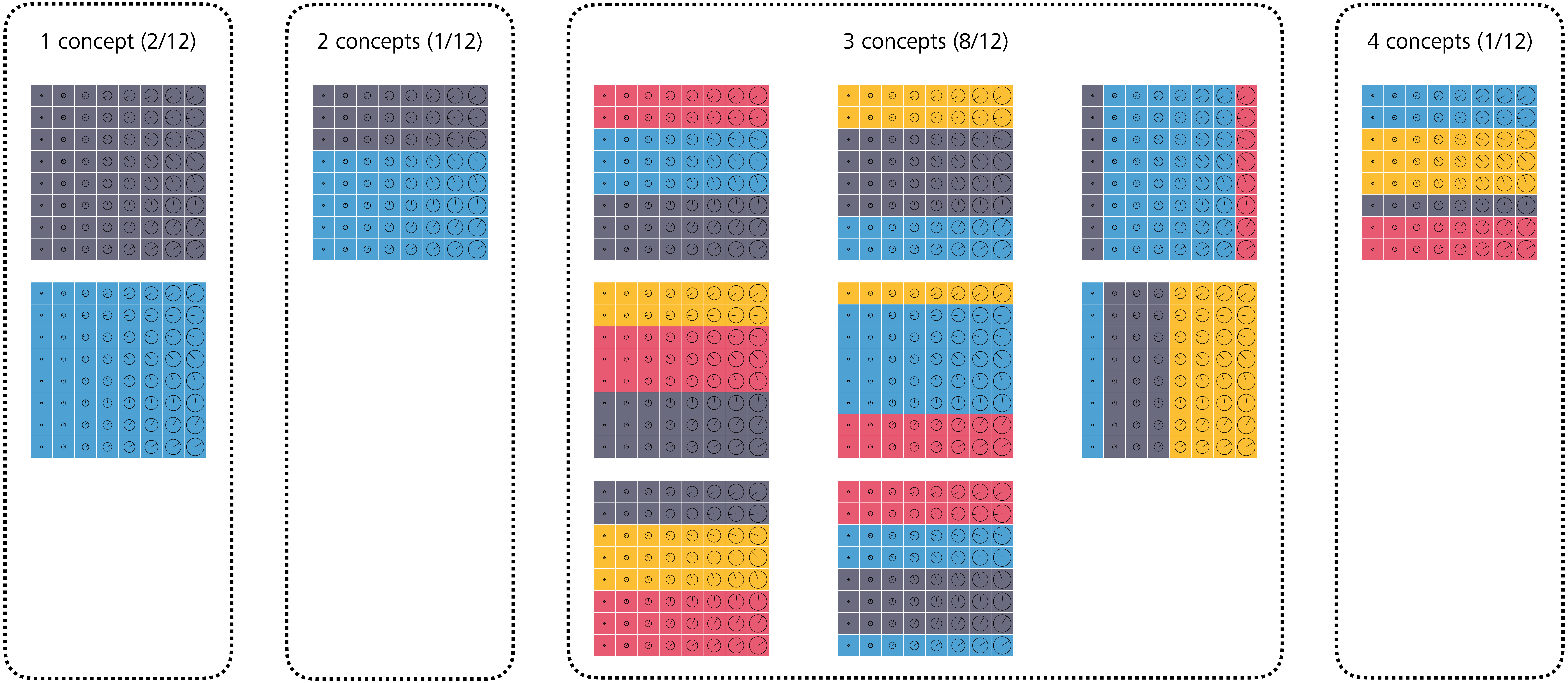
Iterated learning with humans



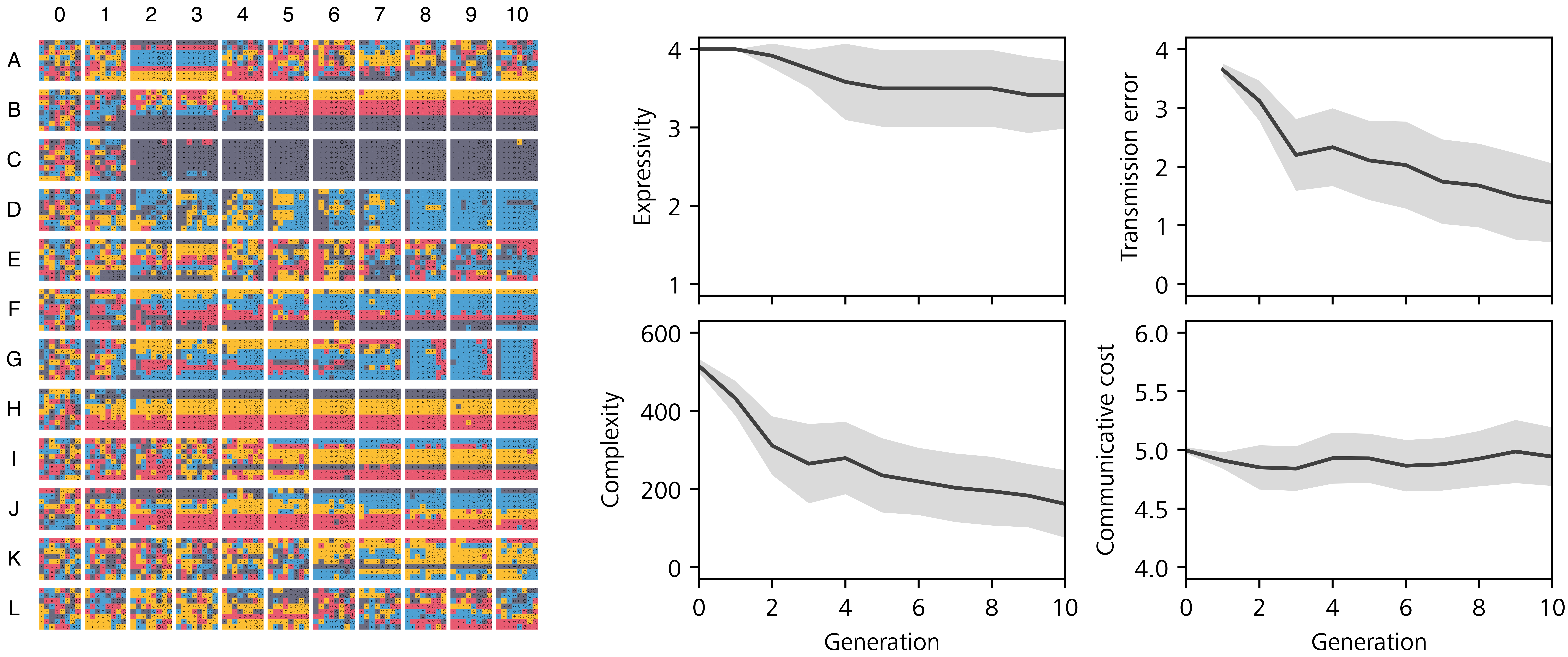




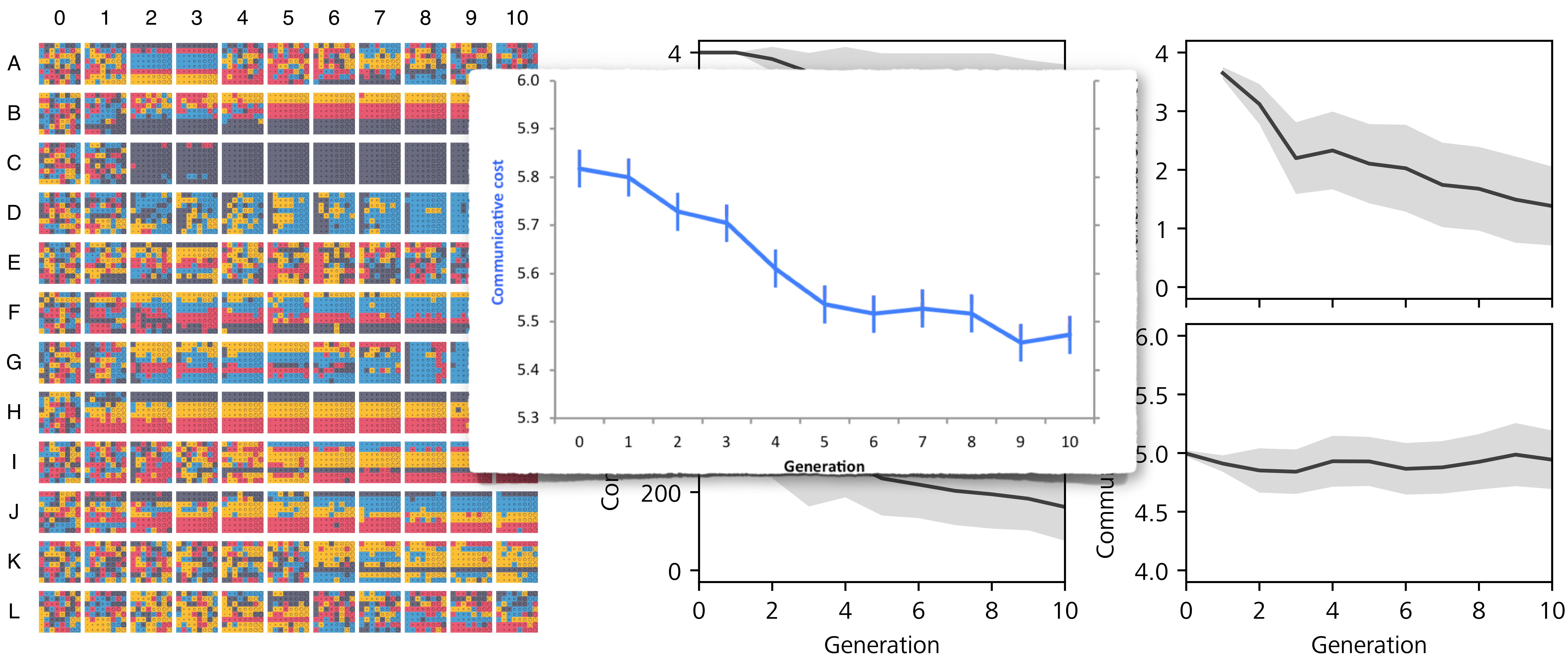
Systems converged on



Experimental results

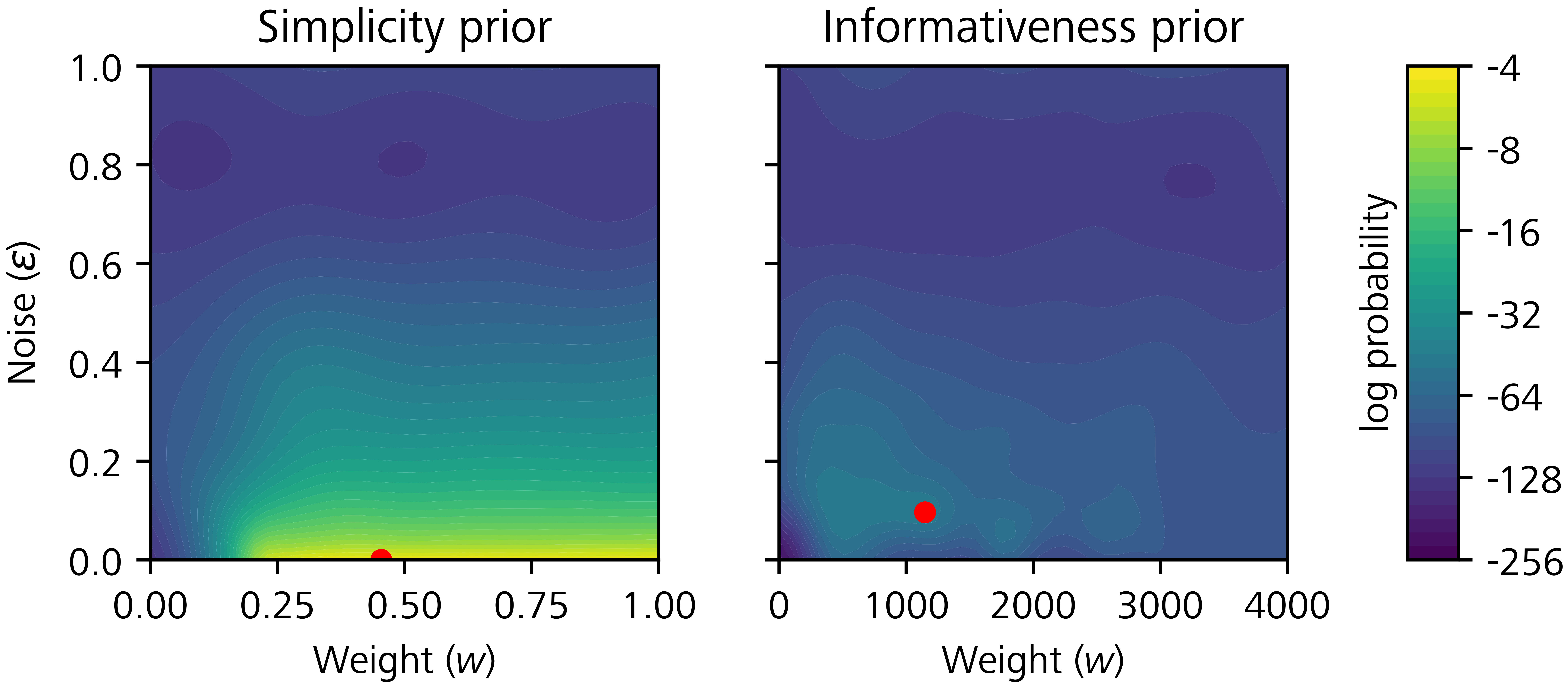


Experimental results

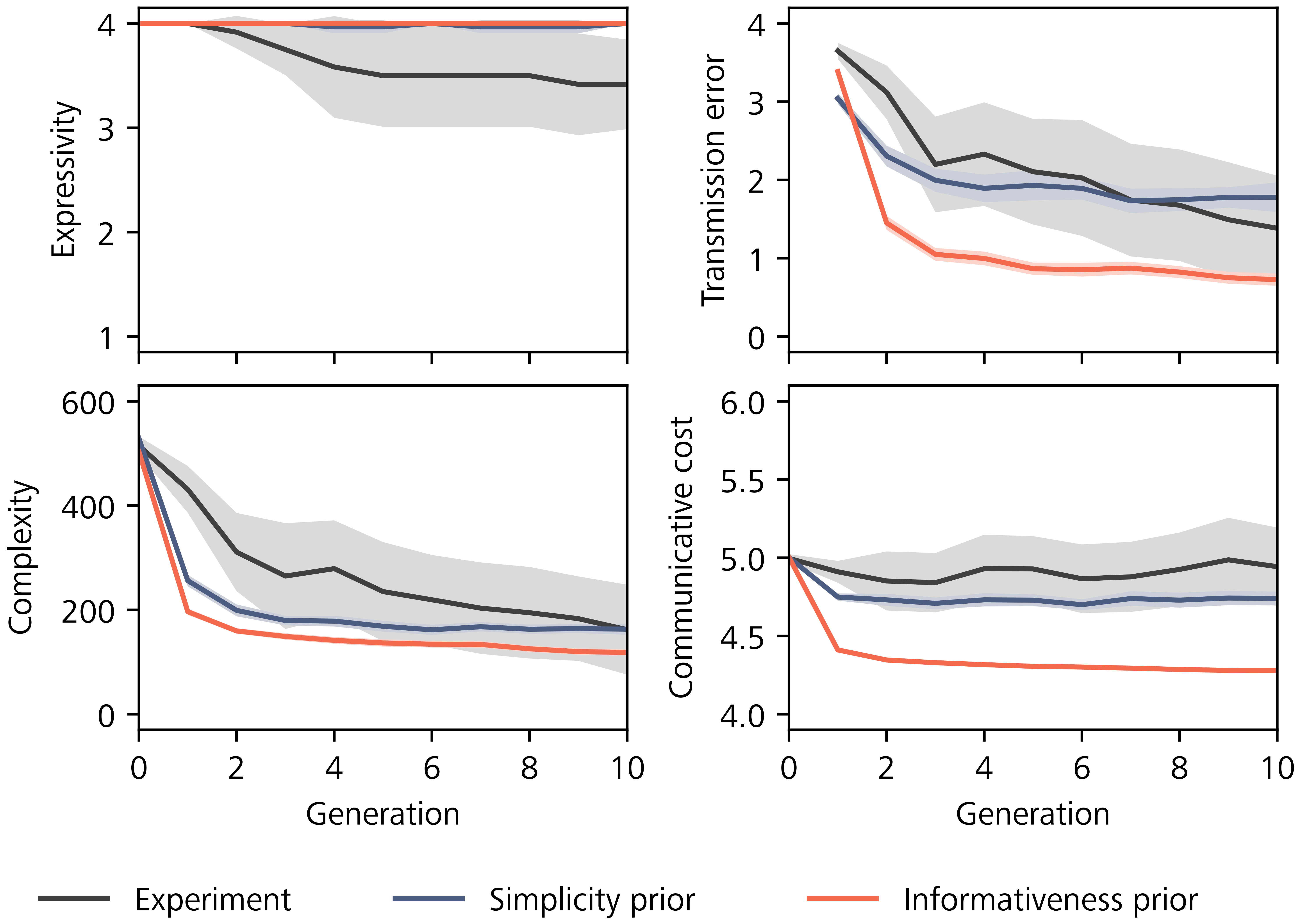


Model fit

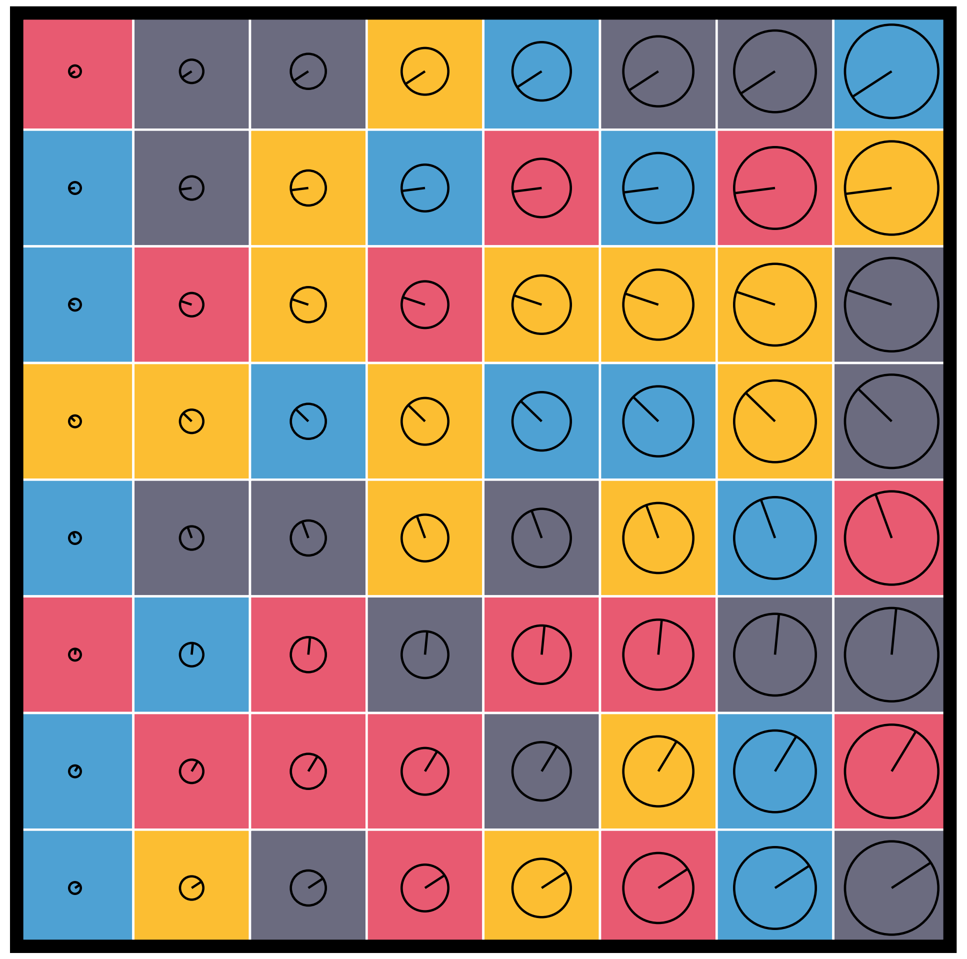
Estimating unknown parameters of the model



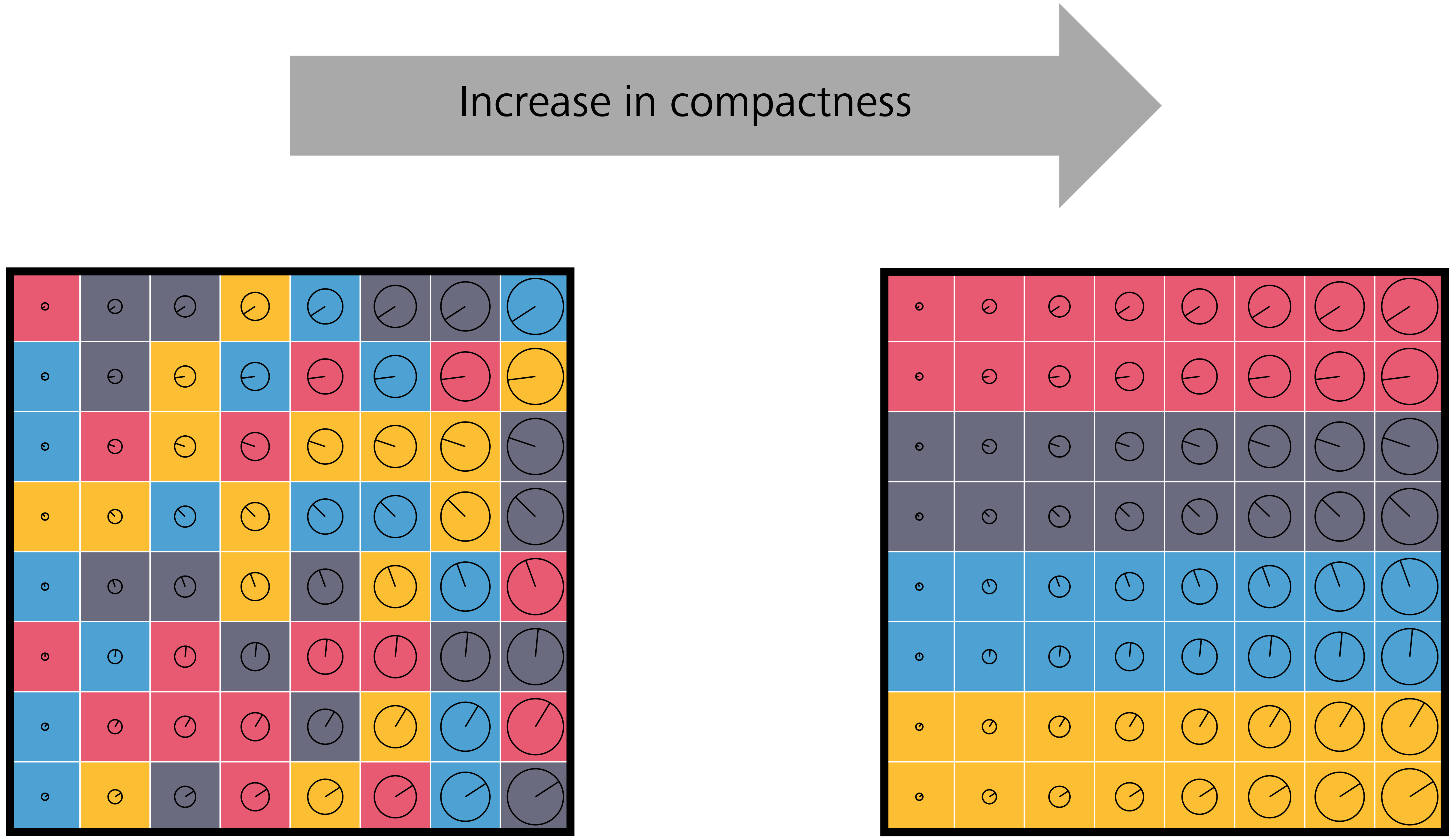
Rerun the model with parameters estimated from the experiment



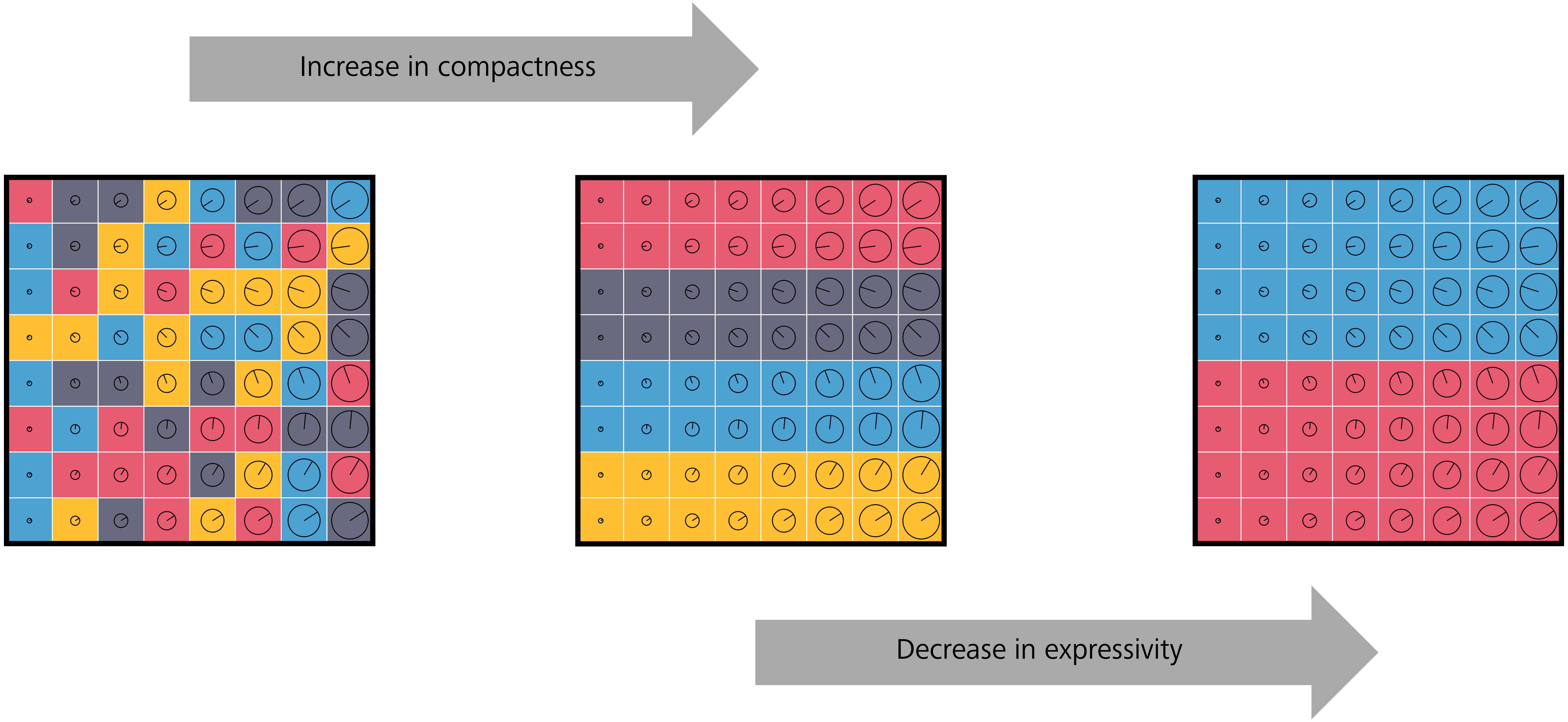
Two ways of achieving simplicity



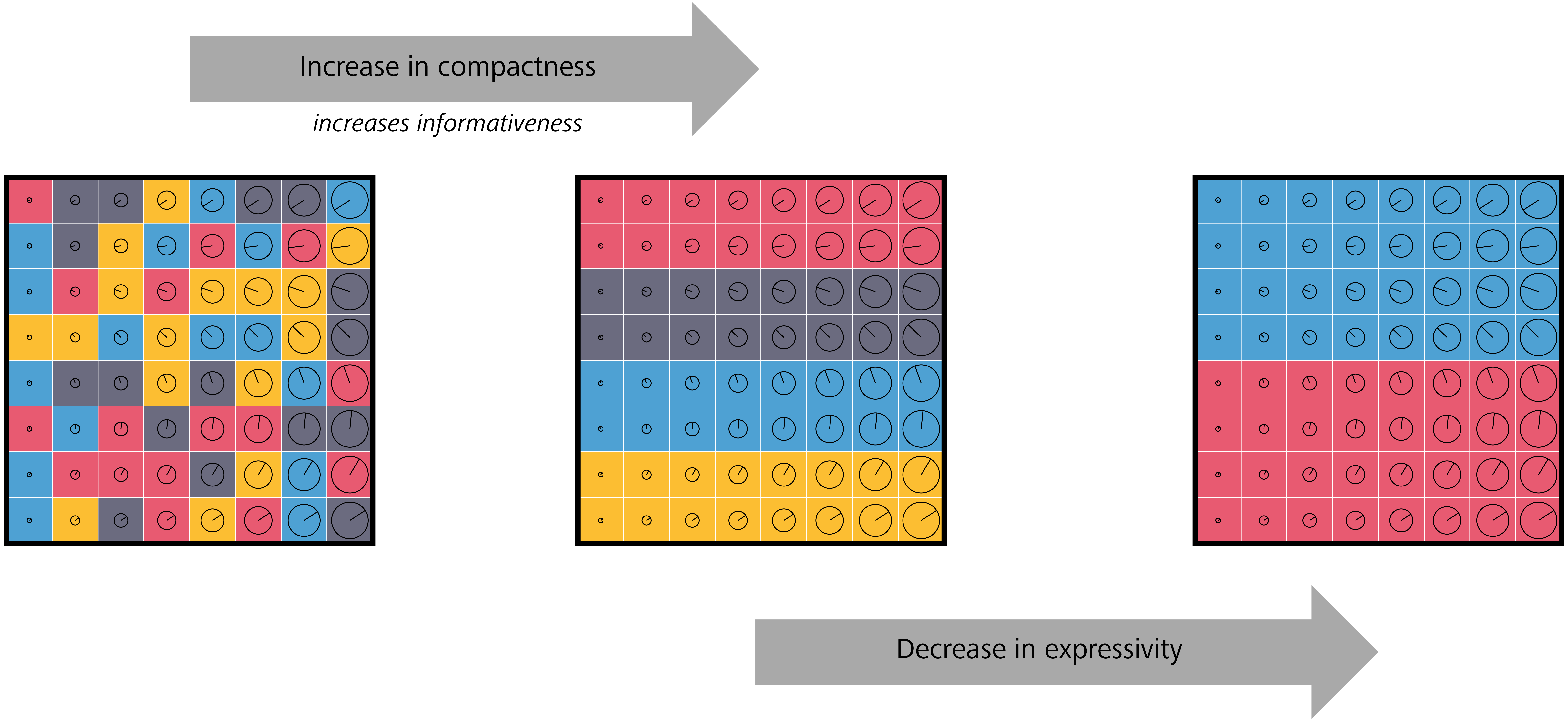
Two ways of achieving simplicity



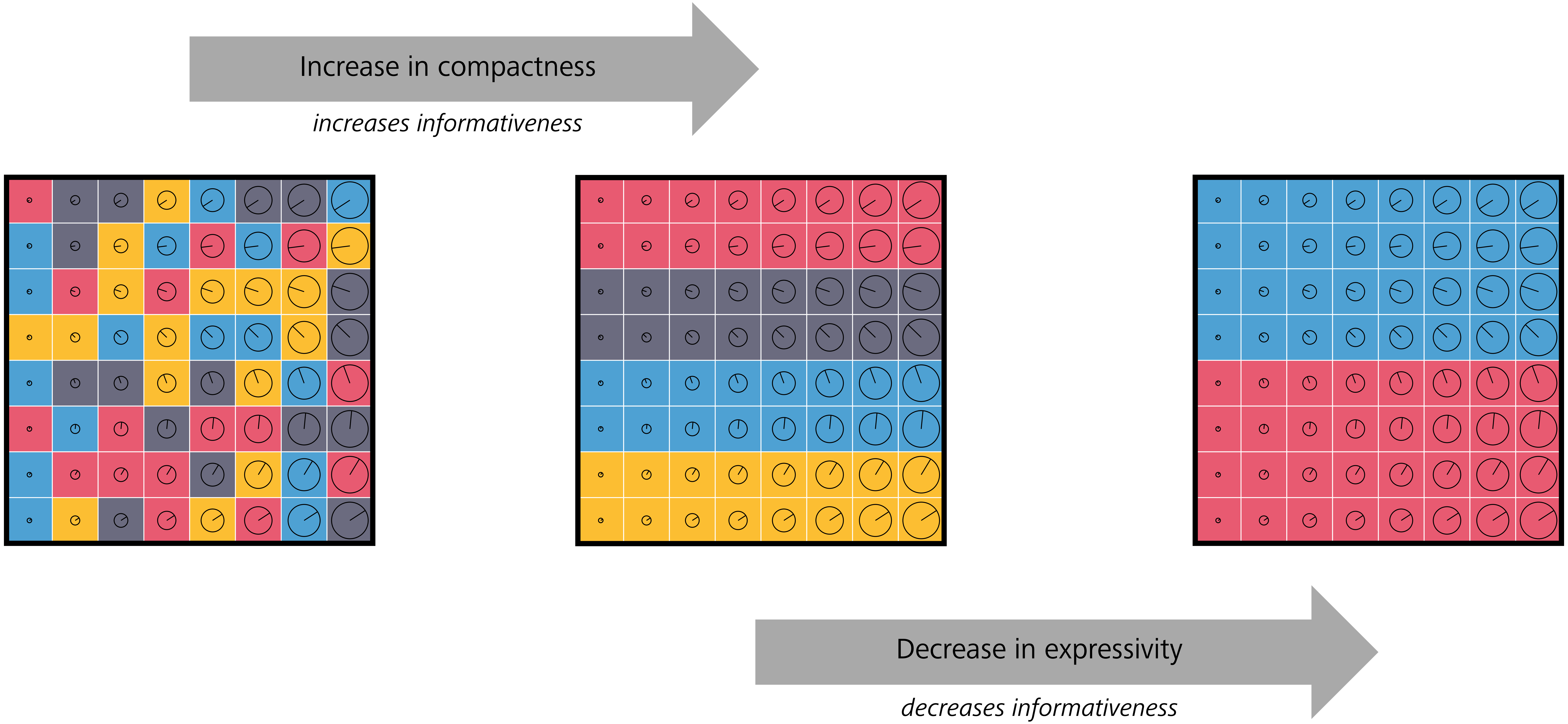
Two ways of achieving simplicity



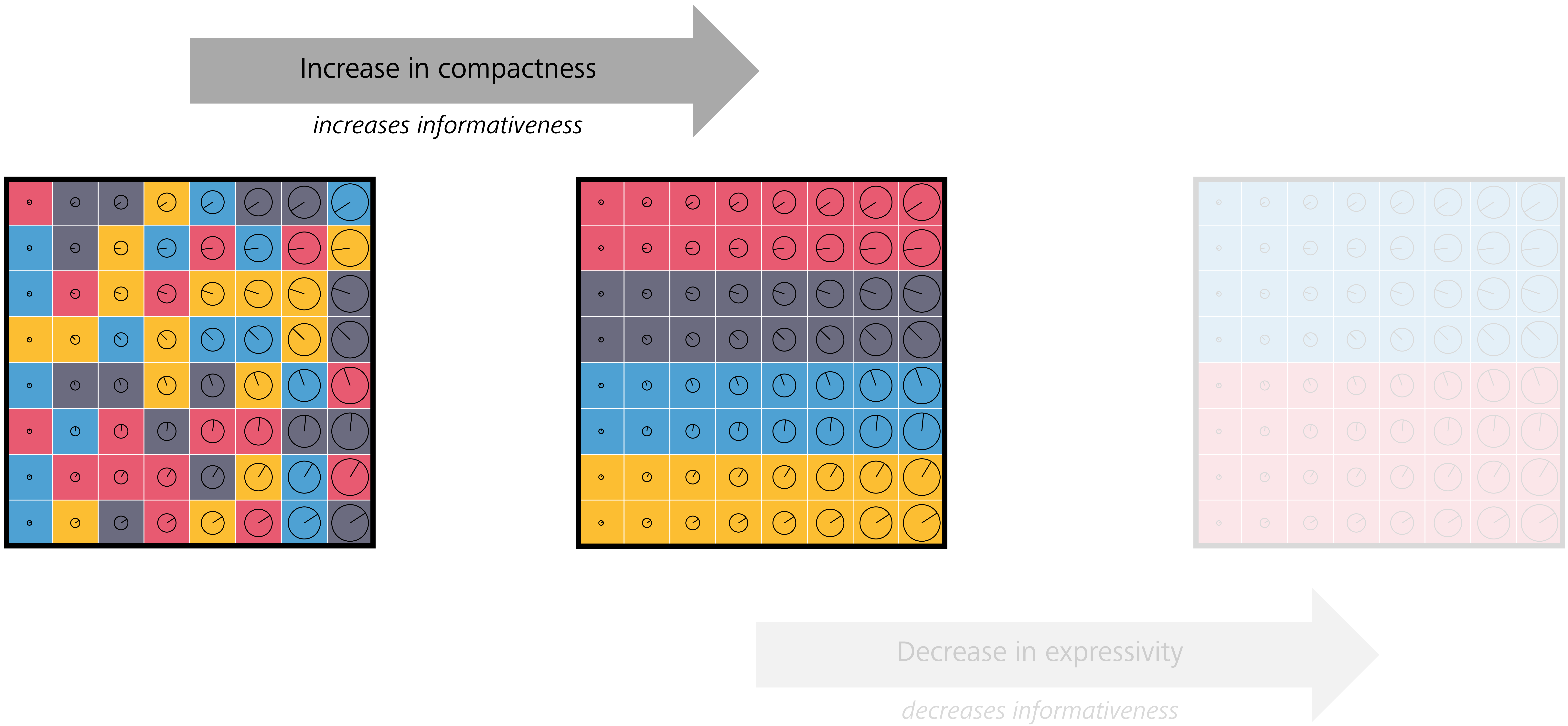
Two ways of achieving simplicity



Two ways of achieving simplicity



Two ways of achieving simplicity



Conclusions

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

For a rational learner, induction contains a simplicity bias to prevent overfitting noise, and to aid reasoning about unseen meanings

Iterated learning (repeated induction) converges to the prior bias, favouring languages that are as simple as possible:

> **Loss of expressivity:** Loss of words/concepts to aid learning
>
> **Compact categories:** Reorganization of the space to aid learning

In the process, some informativeness may come along for the ride, potentially obscuring the causal mechanism

Nevertheless, some kind of interactional dynamics (e.g. learning based on communicative success) must restrain languages from total degeneration

Thanks!