Supervised classification

Computational Cognitive Science 2014
Dan Navarro

Overview

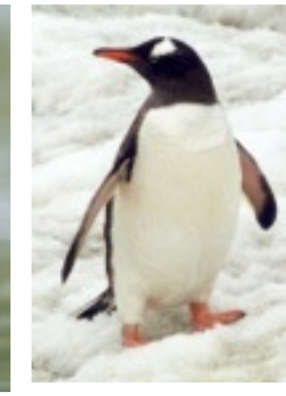
- Introduction
- Some psychological ideas
  - Abstracting a prototype
  - Storing exemplars
- Their machine learning counterparts
  - Simple probabilistic classifiers (prototype models)
  - Some non-parametric classifiers (exemplar models)

How do psychologists and cognitive scientists think about classification?

The celestial emporium of benevolent knowledge

It is written that the animals are divided into: **(a)** those belonging to the emperor, **(b)** embalmed ones, **(c)** tame ones, **(d)** sucking pigs, **(e)** sirens, **(f)** fabulous ones, **(g)** stray dogs, **(h)** those included in the present classification, **(i)** frenzied ones, **(j)** innumerable ones, **(k)** those drawn with a very fine camel hair brush, **(l)** others, **(m)** those having just broken the water pitcher, **(n)** those that from a long way off look like flies.”

-- Jorge Luis Borges



On what basis do we decide to refer to these different things as being examples of the same kind of entity?

Why do we give them the same **category label**?

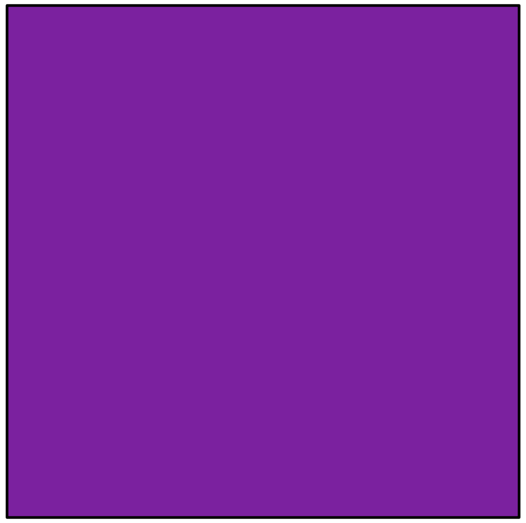
The classical view

- This is everyone's "intuitive" theory of categories
- The idea is simple: there must be a **definition** for the category
- More precisely, a category is defined by a features that are
  - **Individually necessary:** all category members <u>must</u> possess all of the features in the list
  - **Collectively sufficient:** if an item has all the features, it <u>must</u> be in the category.

The classical view

- Aristotle:
    - Definitions are “an account [or logos] that signifies the essence” (Topics I)
- Descartes, Locke, Hume:
    - Different approaches, but all seek “correct definitions” of abstract concepts like body, freedom, and miracles.
- Hull (1920):
    - “At length the time arrives when the child has a ‘meaning’ for the word dog… [T]his meaning is found to be actually a characteristic more or less common to all dogs and not common to cats, dolls, and ‘teddy-bears.’ (p6)

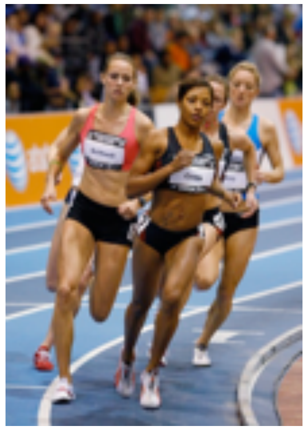
- is an integer
- is greater than 2
- is less than 4



- is a 2-D shape
- has four sides
- has all angles the same
- has all sides the same

Sport?

has a **ball** involved...
what about:



or

involves **running**...
what about:

or

involves **exertion**...
what about:

or

Games

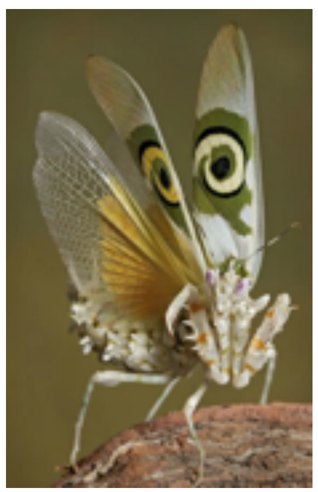
- Wikipedia offers wonderful examples of this sort of thing
- One attempt to define a “game” (Caillois):
  - **fun**: the activity is chosen for its light-hearted character
  - **separate**: it is circumscribed in time and place
  - **uncertain**: the outcome of the activity is unforeseeable
  - **non-productive**: participation does not accomplish anything useful
  - **governed by rules**: the activity has rules that are different from everyday life
  - **fictitious**: it is accompanied by the awareness of a different reality

Games

- It excludes a lot of things I'd call games
  - Political game playing is deadly serious, but it's a game
  - The outcome of a jigsaw puzzle is not uncertain, but those are games
- It includes a lot of of things I wouldn't call games
  - Teaching a class about Ptolemaic physics would be fun, rule-governed, unproductive, uncertain, fictitious etc.
  - Not really a "game". Certainly less than political gaming or playing a jigsaw

Classical theory fails :-(

In general, despite many decades of trying, nobody has succeeded in coming up with necessary and sufficient conditions for almost any (non-mathematical) everyday concepts. There are always awkward failures.



… And this problem is even more acute for abstract concepts like **freedom** or **duty** or **the**

Maybe we'll work backwards from the empirical data on how people classify things?

Graded membership… some category members feel “better” than others

Typical birds
(good examples)

Atypical birds
(not good examples)

Typical items are classified faster

Atypical items are classified slower

Typical items are easy to think of

Atypical items are harder to think of

People all agree that typical items belong in the category

furniture!

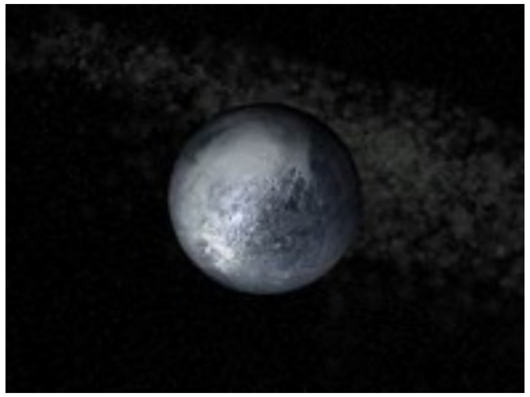
There is little consensus about atypical items

furniture?

People all agree that typical items belong in the category

Earth is a planet

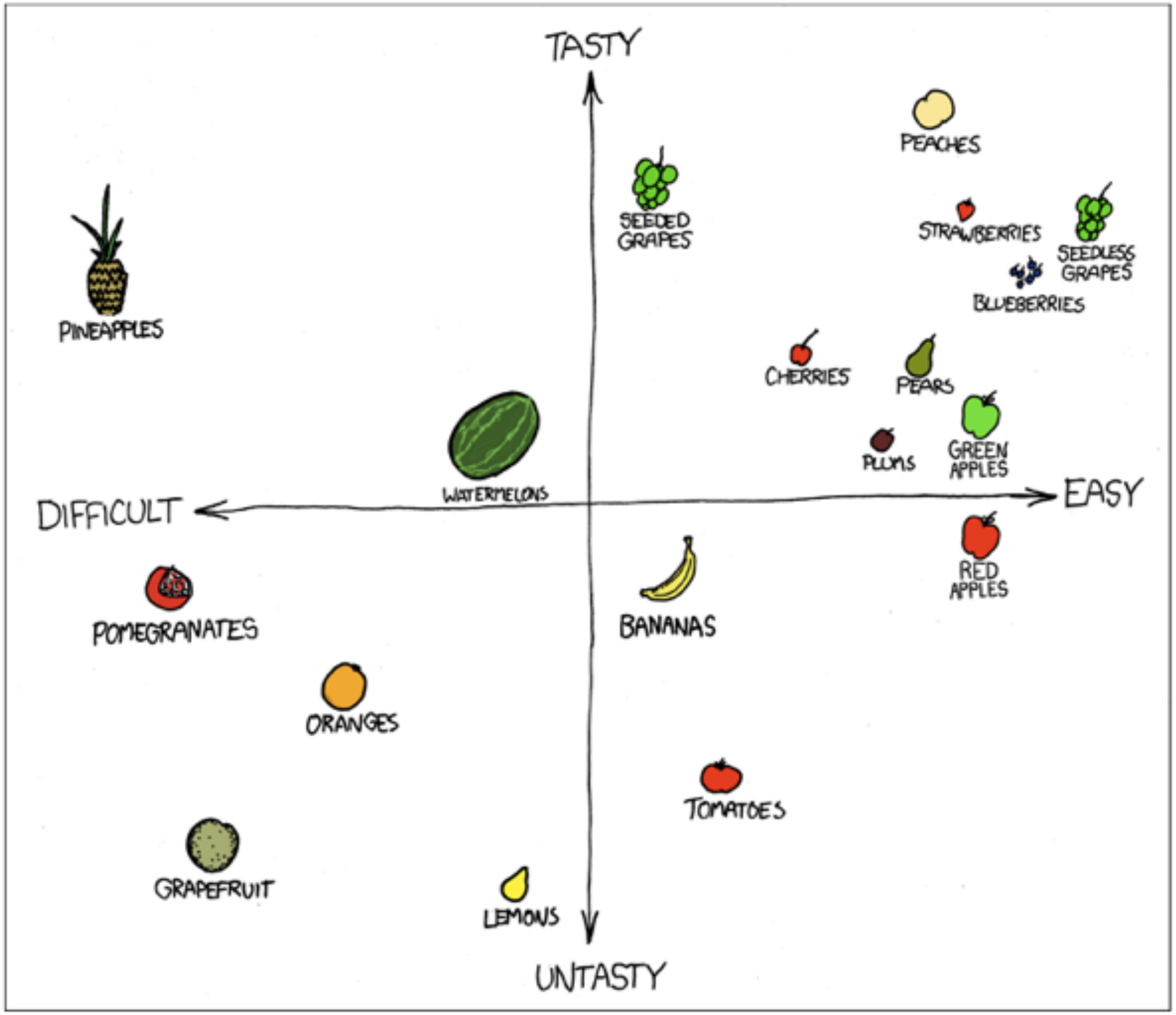
There is little consensus about atypical items



is Pluto?

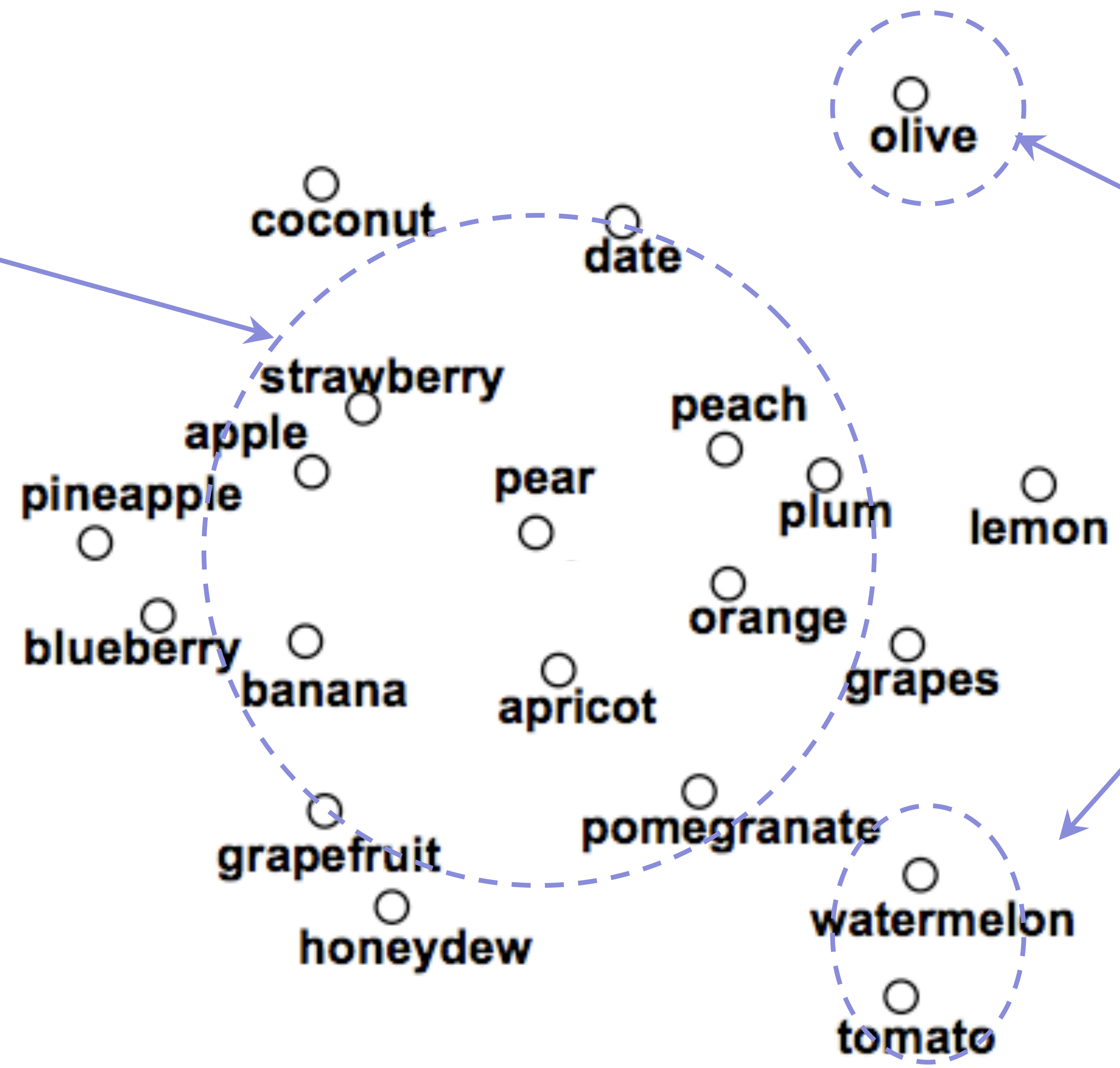
The family resemblance view

Natural categories have some scatter, with some members more central and others not



Here is a multidimensional scaling solution showing people's representations of fruit…

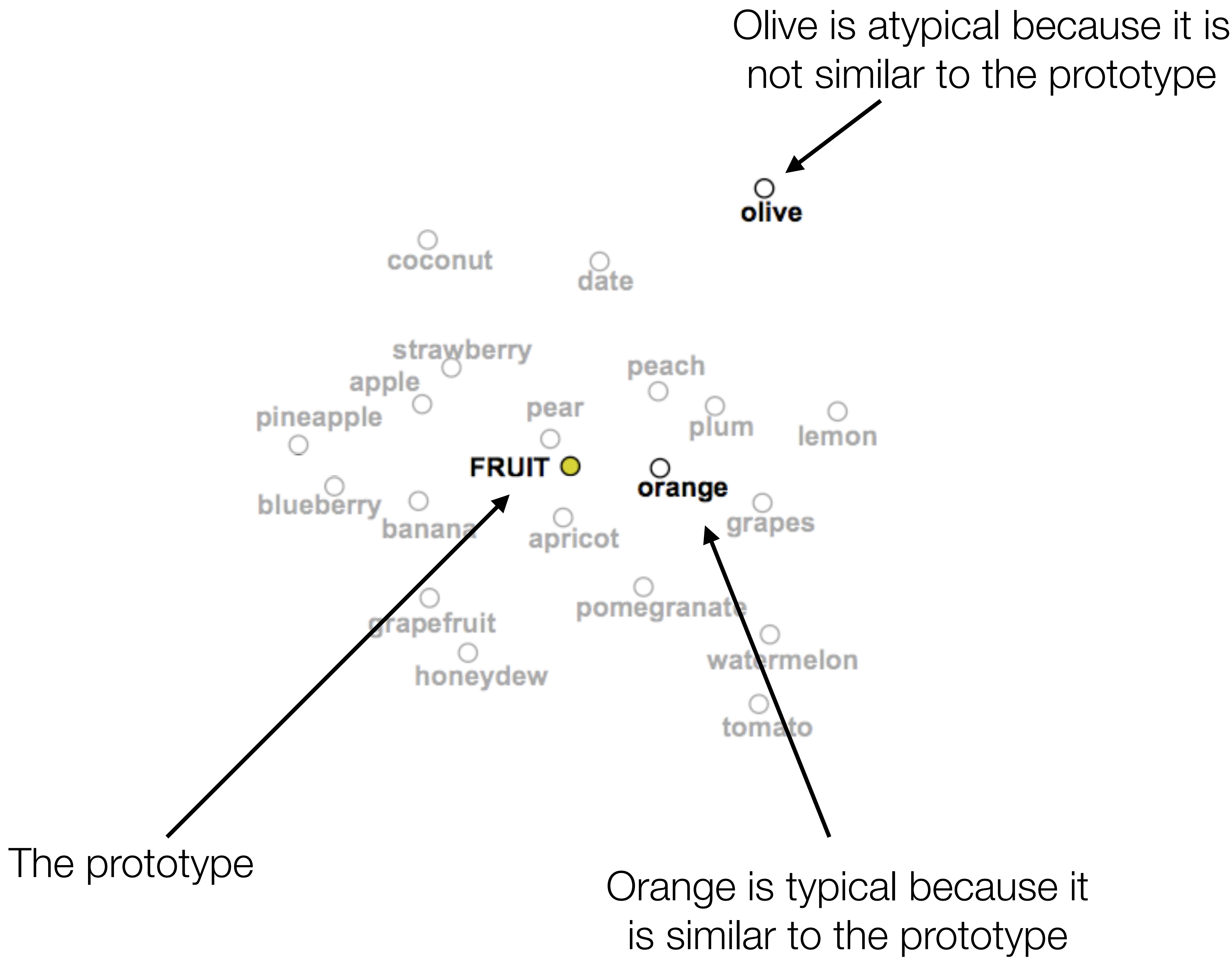
Central "core" of typical category members



Atypical members

The prototype view

- The prototype view is one way of thinking about family resemblance
- It proposes that people represent a category in terms of a single “ideal” member
    - This “ideal” category member is called the **prototype**
    - It may or may not correspond to a real object
    - The prototype is the “best possible” category member
    - Sometimes thought of as being in the middle of all the others

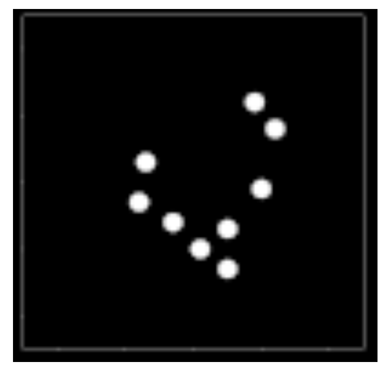

Olive is atypical because it is
not similar to the prototype
olive
coconut
date
strawberry
peach
apple
pear
pineapple
plum
lemon
FRUIT
orange
blueberry
banana
apricot
grapes
pomegranate
grapefruit
watermelon
honeydew
tomato
The prototype
Orange is typical because it
is similar to the prototype

A classic study…

Prototype

Showed

Low distortion



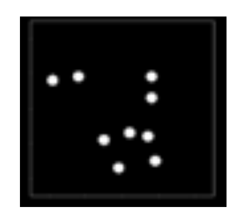
High distortion

Tested: all novel items

Prototype

Low distortion

High distortion



Random dots

Posner & Keele (1969)

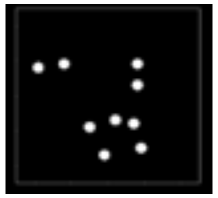
A classic study…

People were faster at responding to the prototype, and (after a week) more likely to think they'd seen it in the original testing

Prototype

Low distortion

High distortion

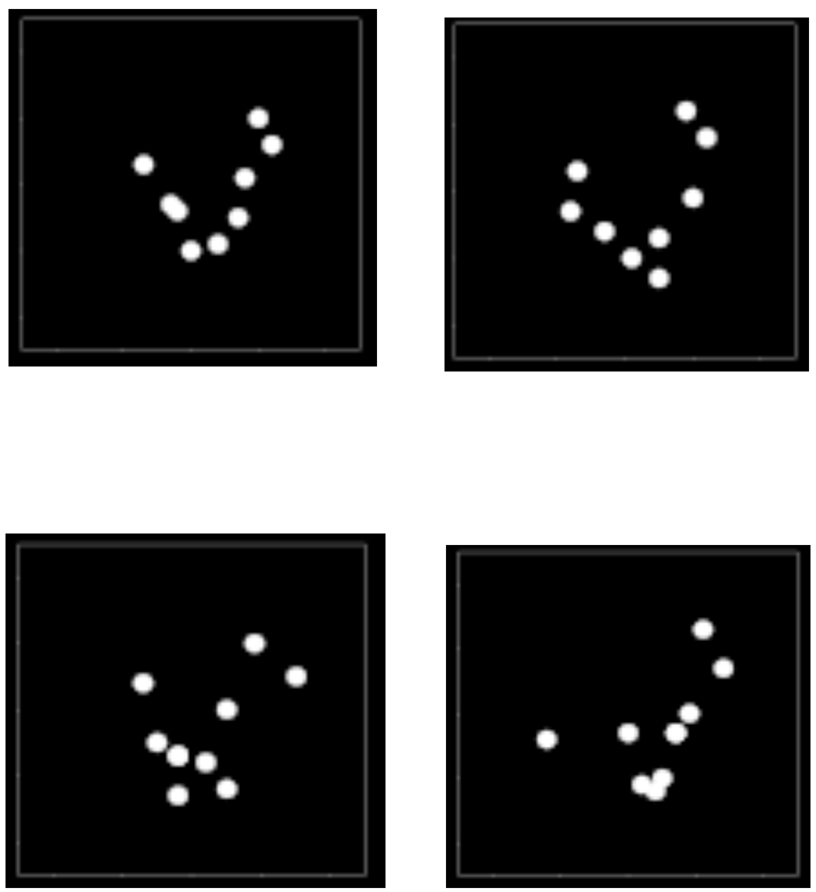


Random dots

Posner & Keele (1969)

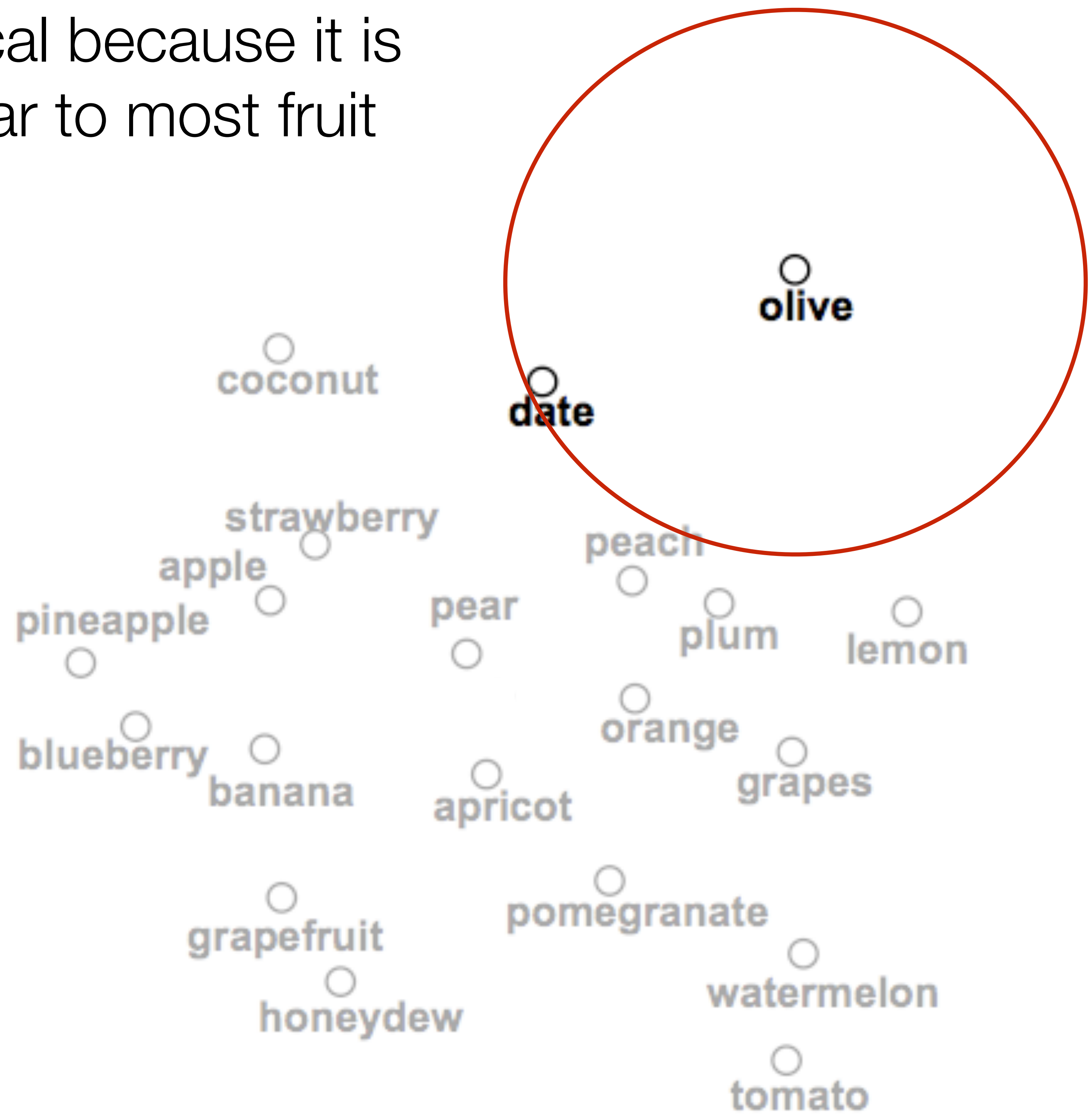
Exemplar theory

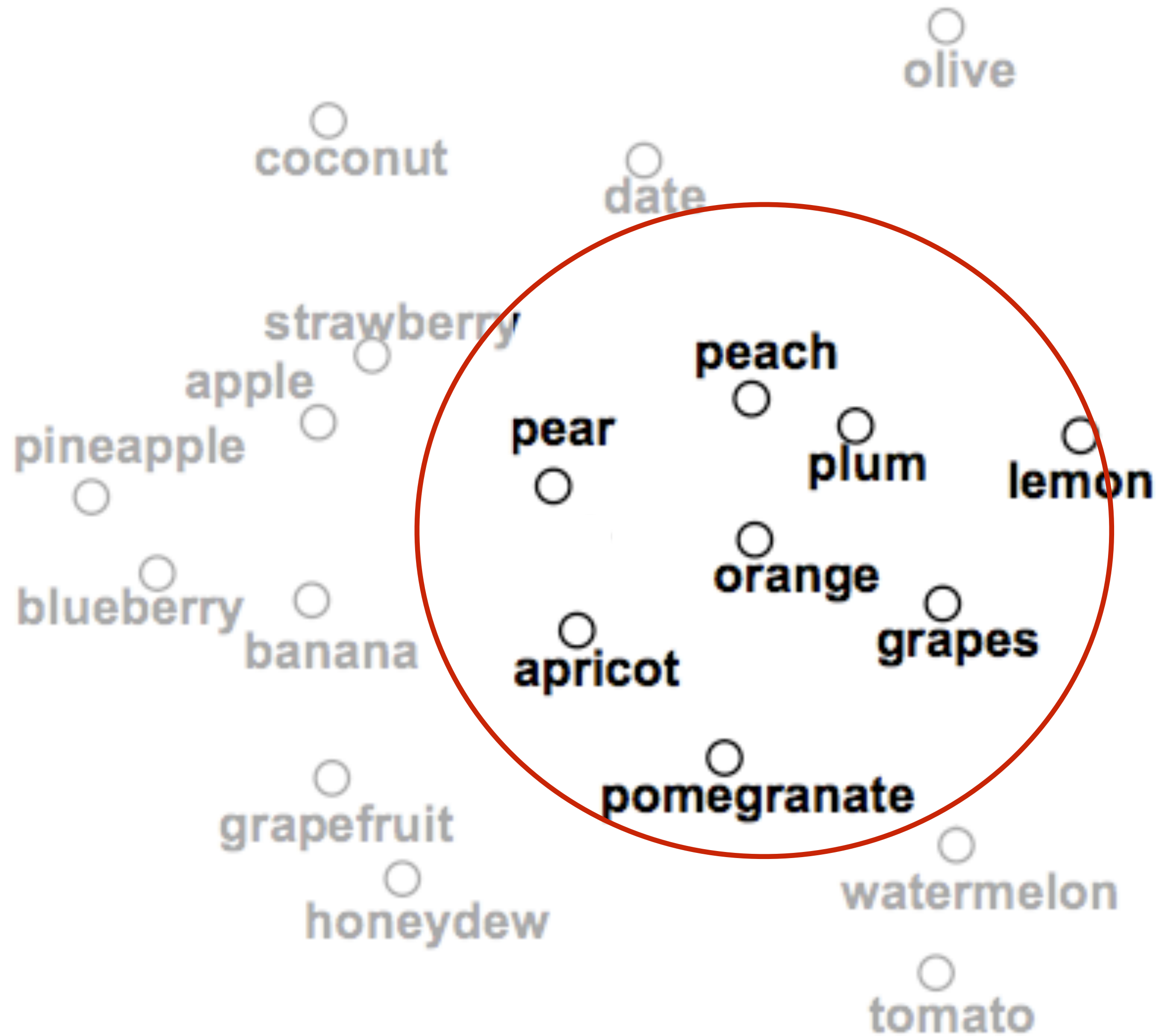
An alternative to prototype theory. It suggests that rather than forming an abstract representation, we remember every specific instance (or **exemplar**) individually.



Then when we do see a new category member, its typicality is determined by its overall similarity to all of those instances

Olive is atypical because it is very dissimilar to most fruit





Orange is typical because it is similar to lots of other fruit

Some thoughts before moving on

- Both seem to have problems…
  - **Prototype theory**: seems to imply that we ignore or don't remember specific instances. Yet there it feels like we do (and there's empirical evidence for "specific exemplar effects"
  - **Exemplar theory:** seems to imply we don't form any abstract representations of a category. And it implies very high memory demands. Could we really remember everything we've seen?
- We're ignoring a lot of richer ideas about categories…
  - More on this later in the class

Supervised classification from a statistical learning perspective

Different kinds of learning problems

- Supervised learning:
  - You get shown a set of objects
  - You know what category labels they have
  - You need to guess labels for new objects

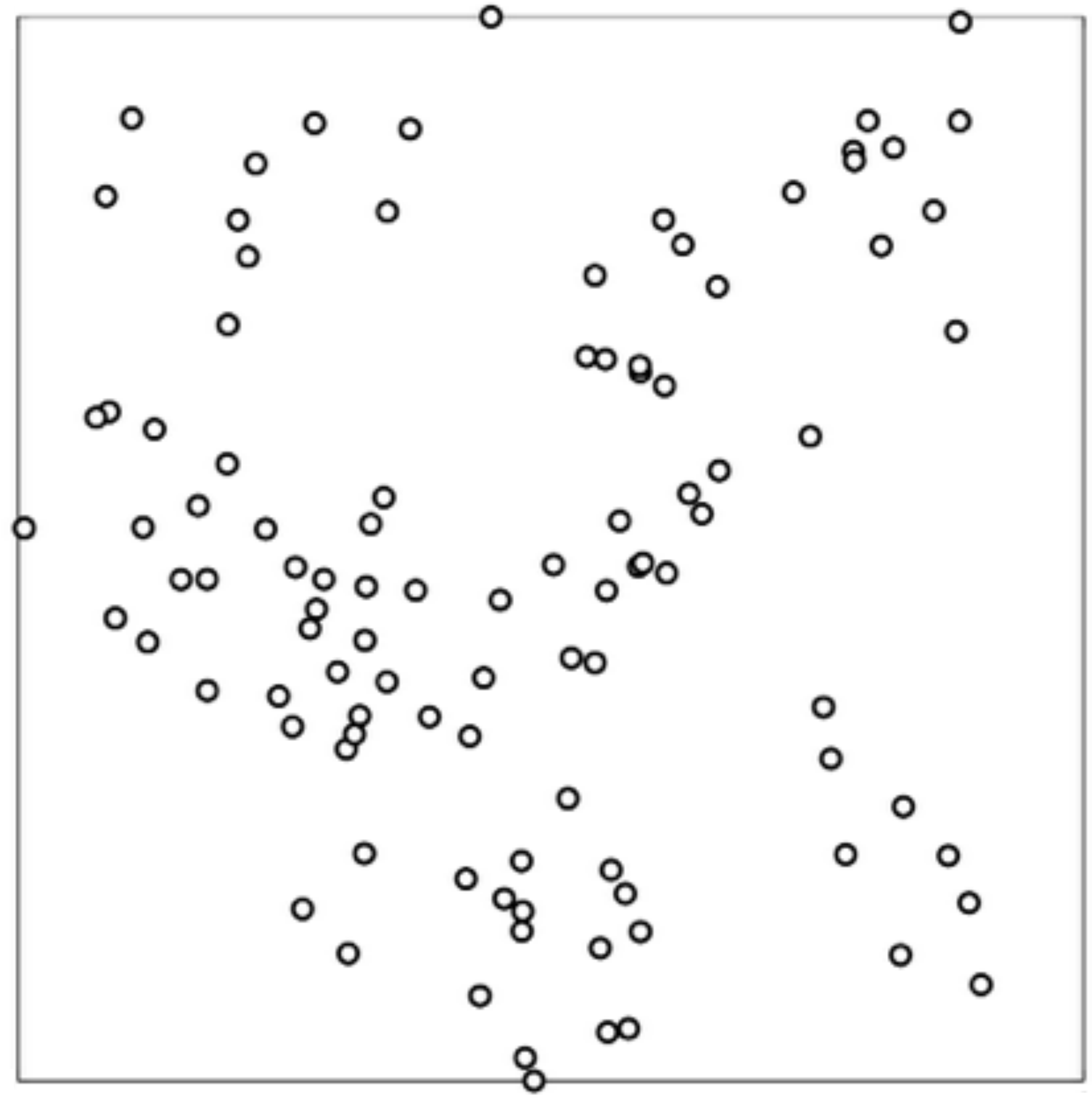
This lecture and the next

Different kinds of learning problems

- Supervised learning:
  - You get shown a set of objects
  - You know what category labels they have
  - You need to guess labels for new objects
- Unsupervised learning
  - You get shown unlabelled objects
  - You need to group them into sensible categories
- Semi-supervised learning
  - You get shown some labelled and some unlabelled objects
  - Figure out the labels for everything

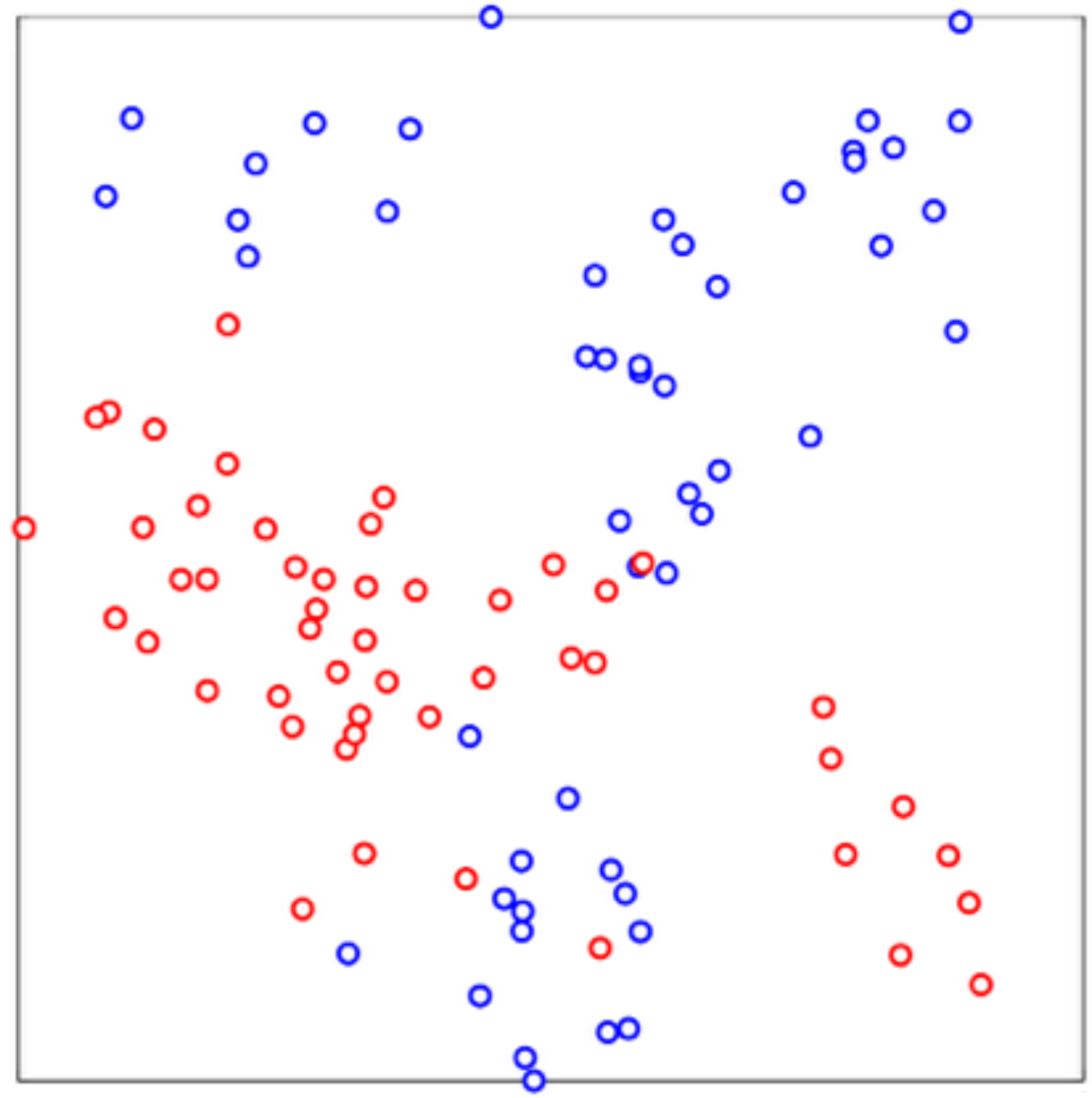
Later lectures

An observation x is a point in some space



An observation x is a point in some space

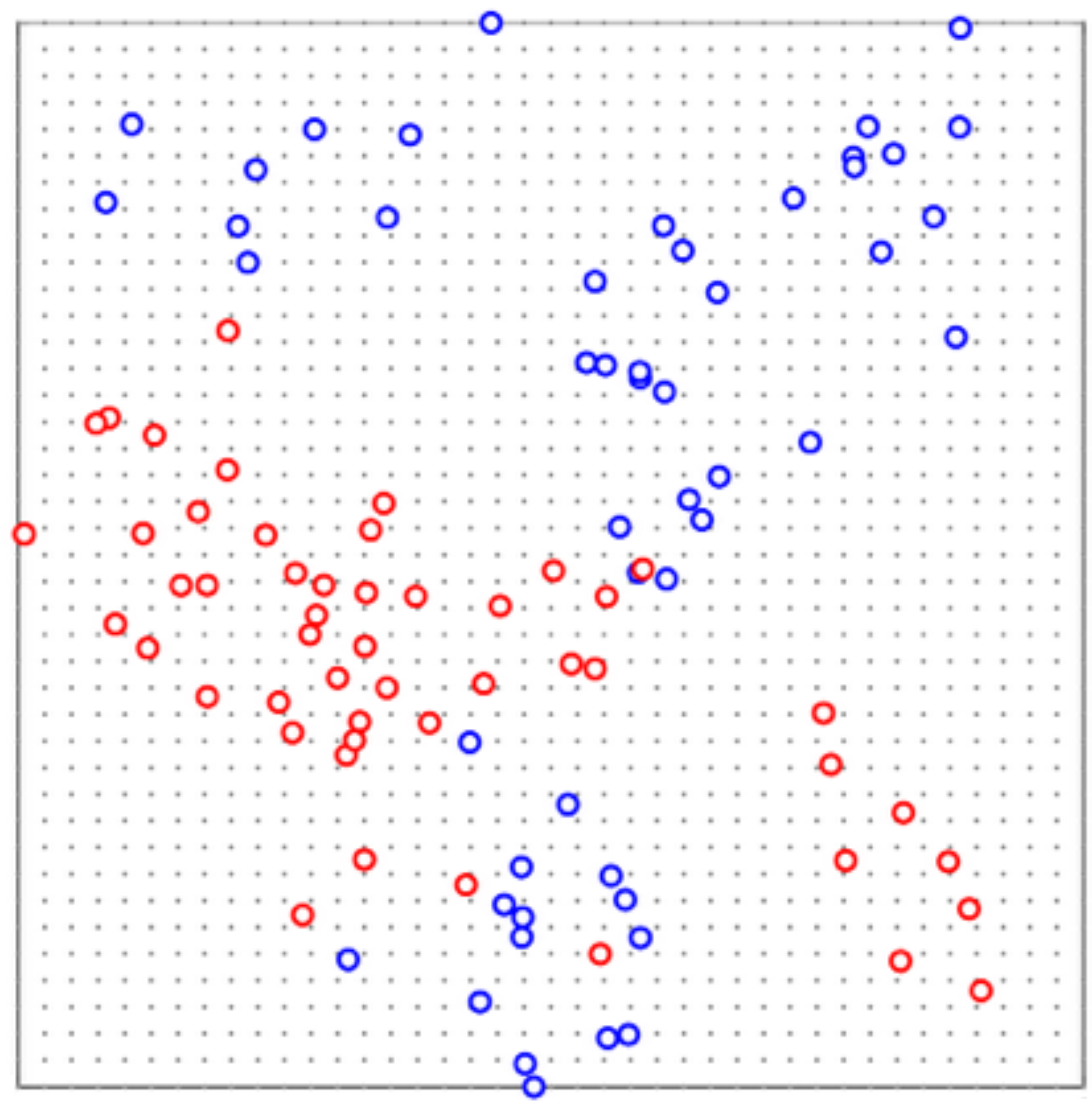
It has a category label $l(x)$ taken from some set of possible labels



An observation x is a point in some space

It has a category label $l(x)$ taken from some set of possible labels

There are other possible observations, y

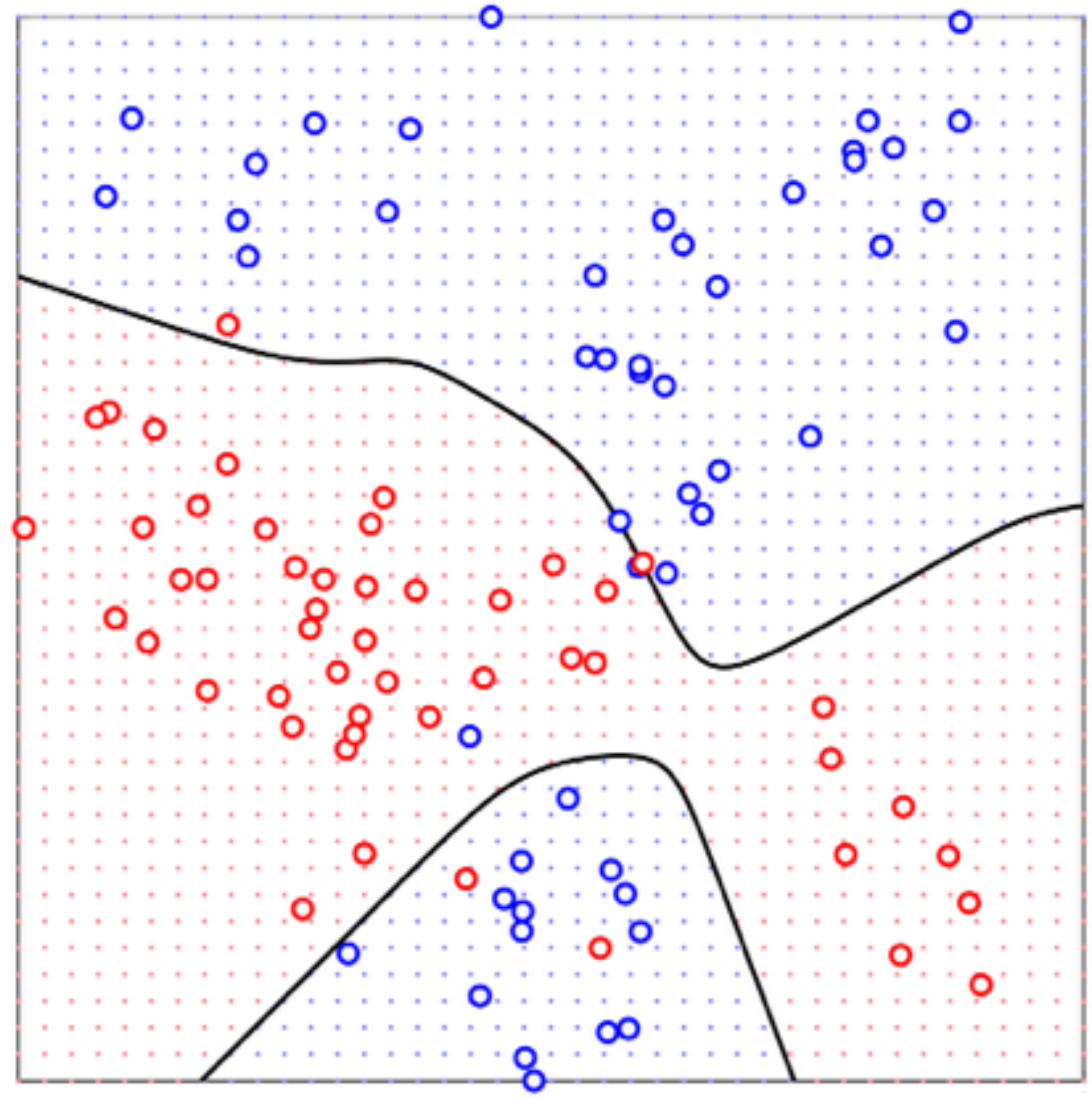


An observation x is a point in some space

It has a category label $l(x)$ taken from some set of possible labels

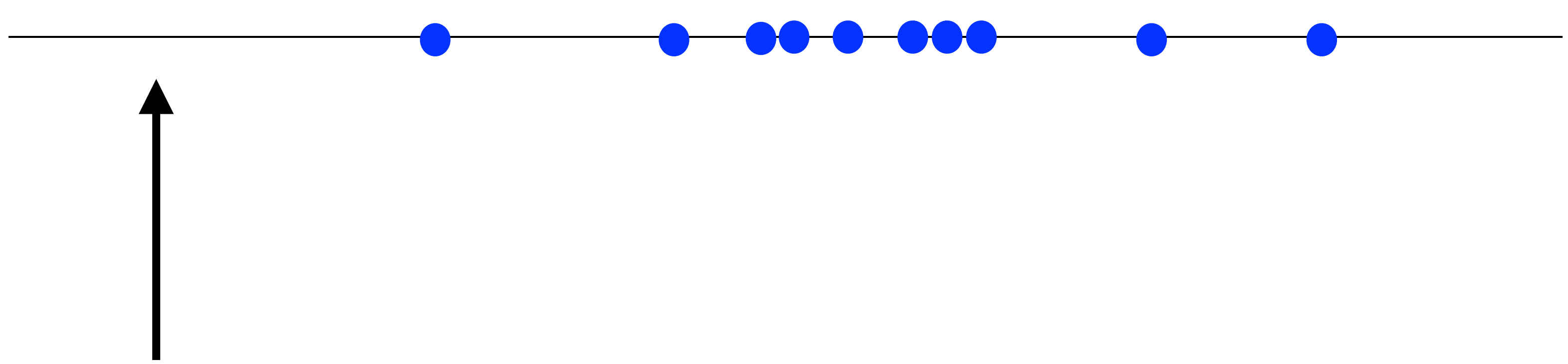
There are other possible observations, y

Our goal is to predict the label $l(y)$ using some model



Very simple Bayesian classifiers

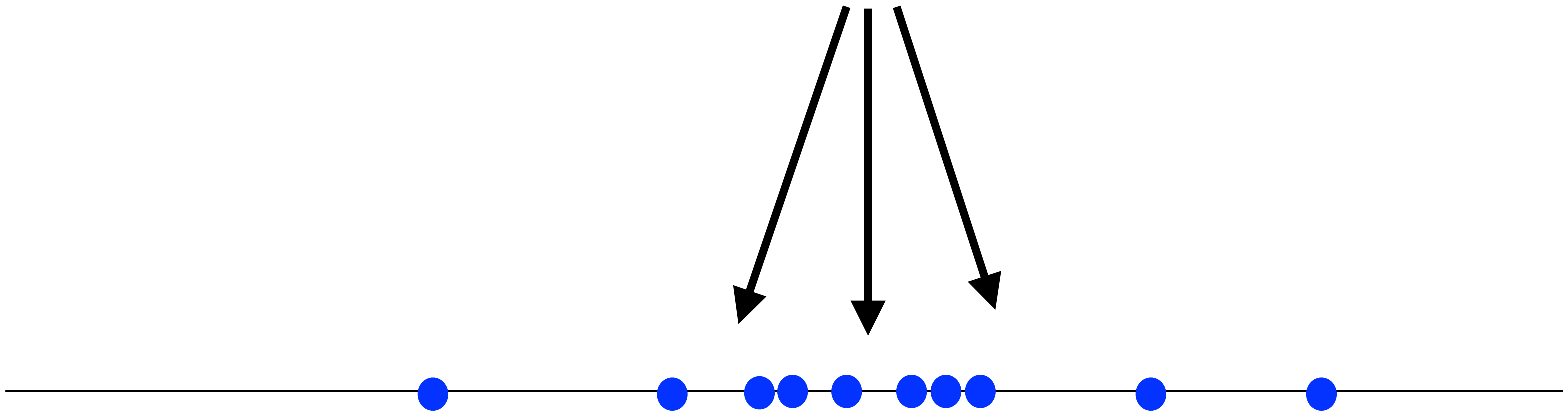
The simplest kind of observations!



The “space” of possible observations is one dimensional, to make it a little clearer on the screen

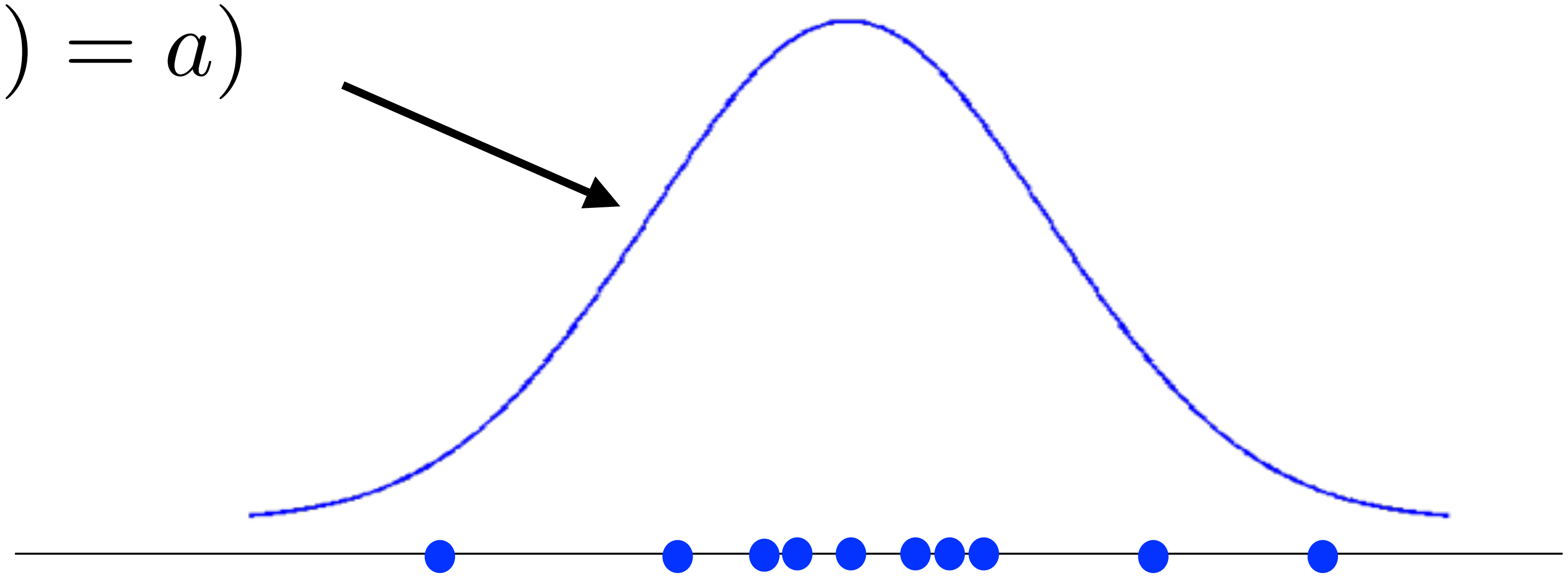
Some observations from the category

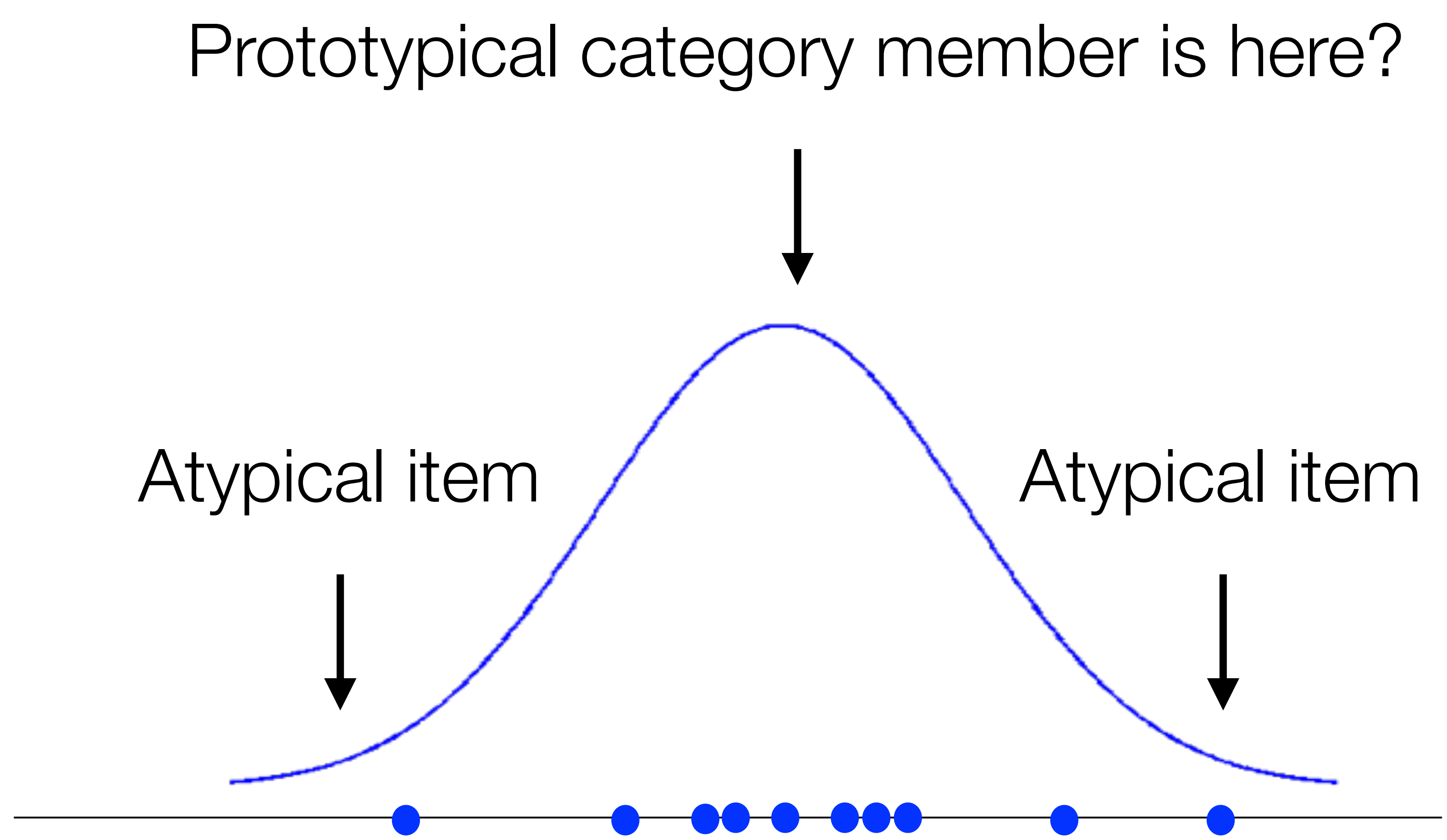
These are the observations x that belong to category a. They have label $l(x)=a$



We represent this as a probability distribution over possible observations

$P(x|l(x) = a)$



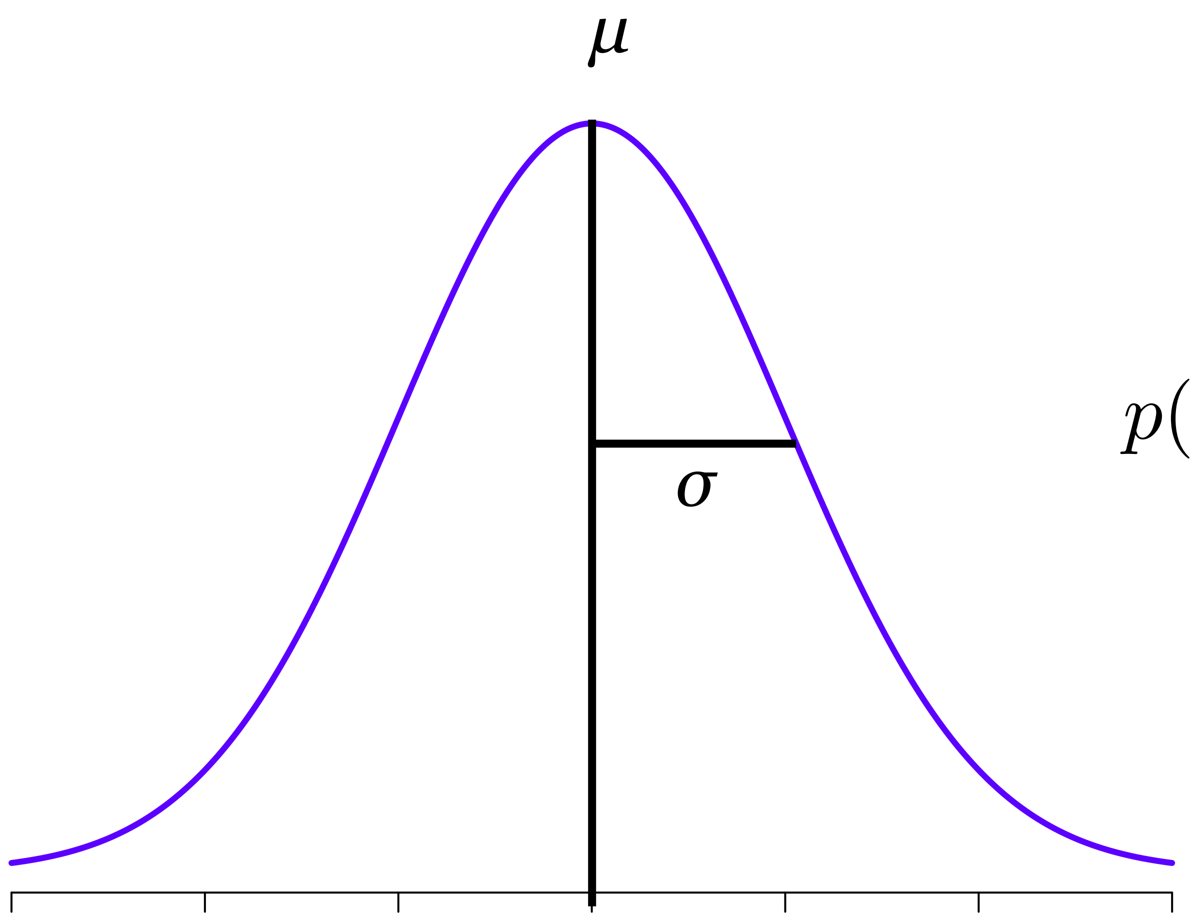


This kind of classifier is very closely linked with **prototype** theory in psychology

Brief digression: The Gaussian distribution and the concept of probability density

Normal (Gaussian) distribution

- Described by two parameters
- The mean, μ, and the standard deviation σ

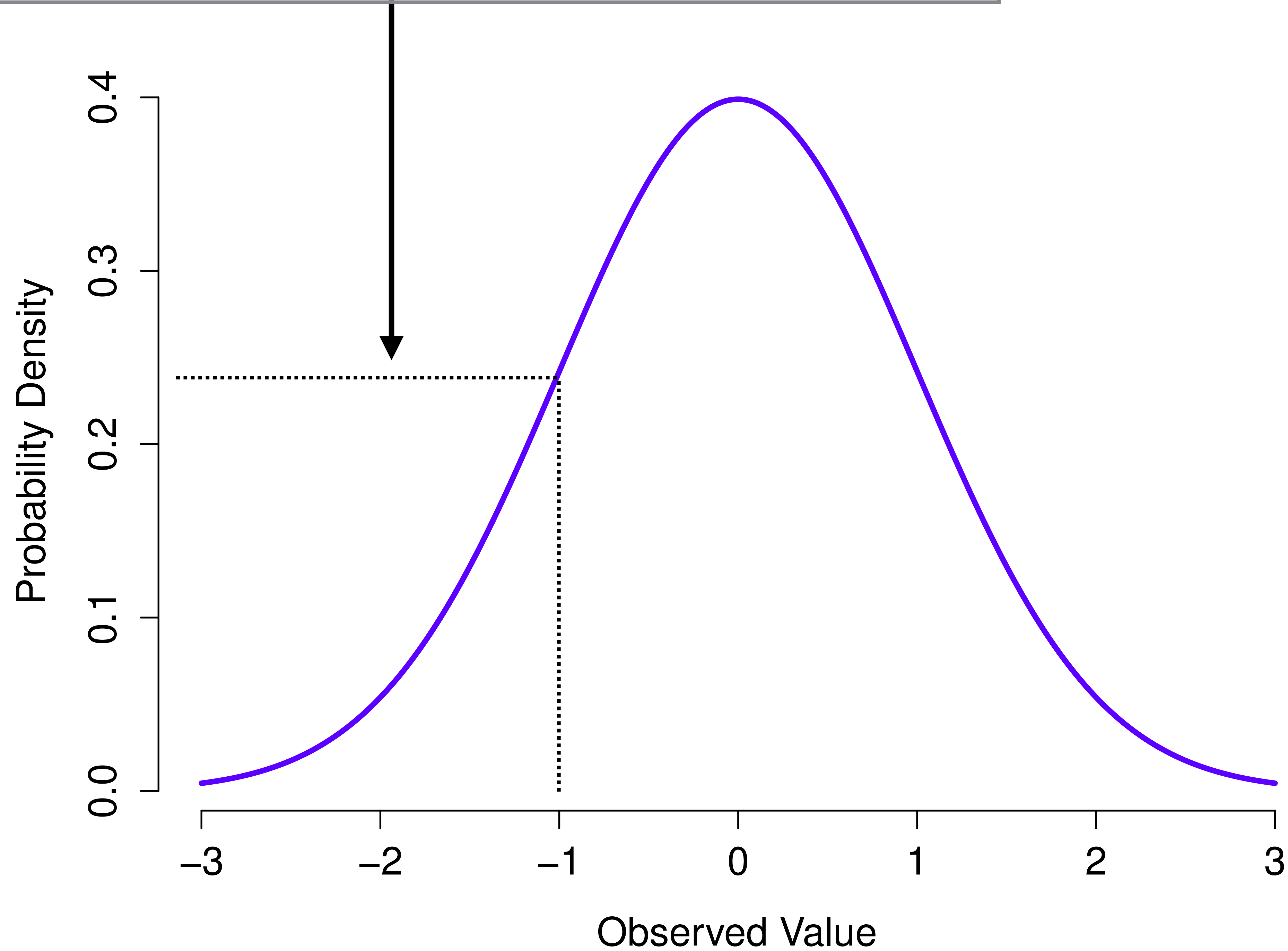


$$p(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$$

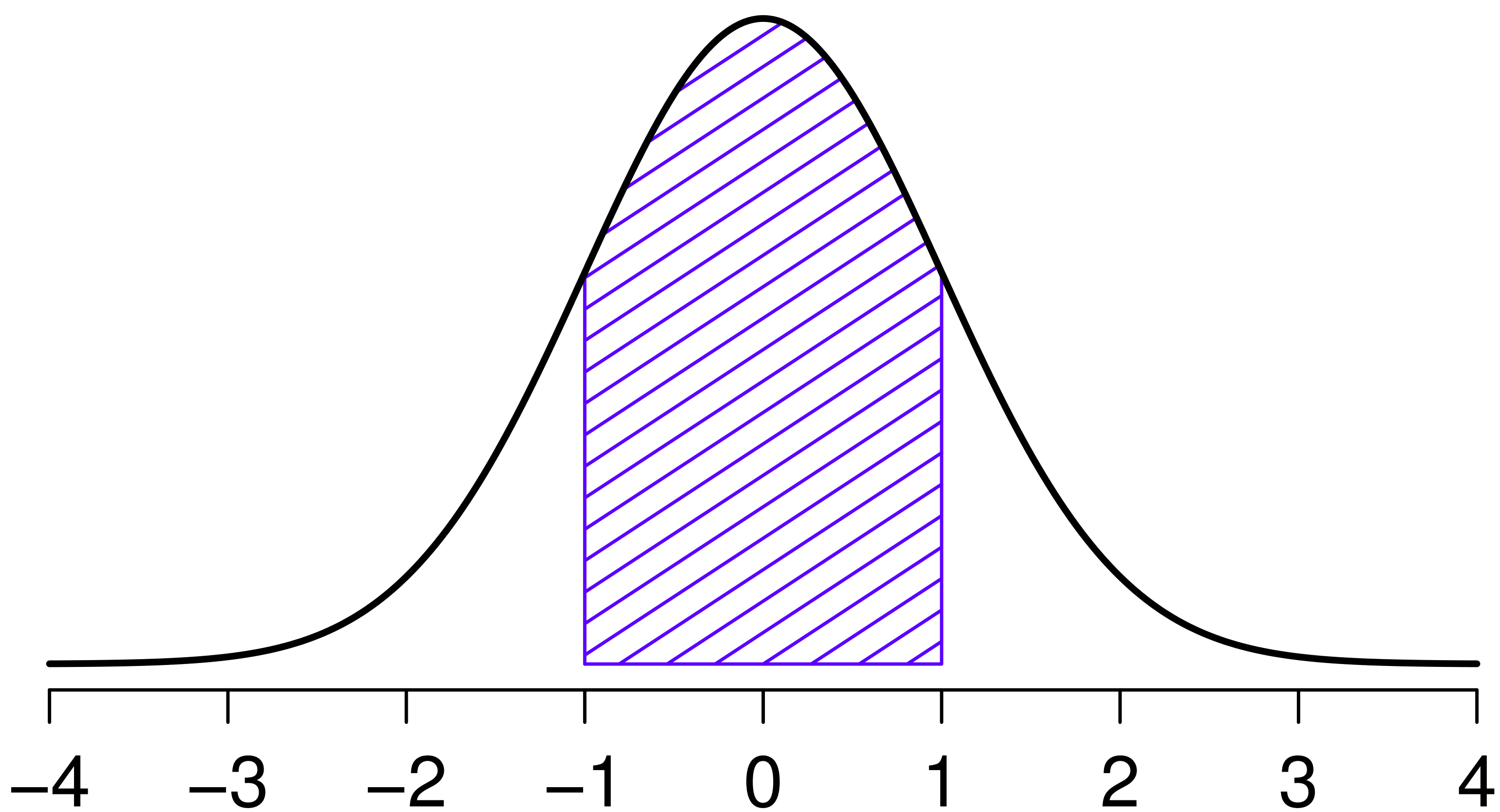
Continuous variables are nasty

- What's the probability of a reaction time of **exactly** 451ms?
  - Not 451.001.
  - Not 450.99288
  - Exactly 451.00000
- The answer has to be **zero**, right?
- Or, more precisely,
  - For continuous variables, the height of the curve isn't a **probability**
  - It's a "probability density"… it describes the tendency for observations to fall in a particular location
  - To get a real probability….

The “probability density” of a score of -1 is 0.24. That **doesn’t** mean that 24% of the data will fall here!



The area under the curve between -1 and 1 is 0.683…
There really **is** a 68% chance that an observation falls within this region



$$P(x \in [a,b] | \mu, \sigma) = \int_a^b \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) dx$$

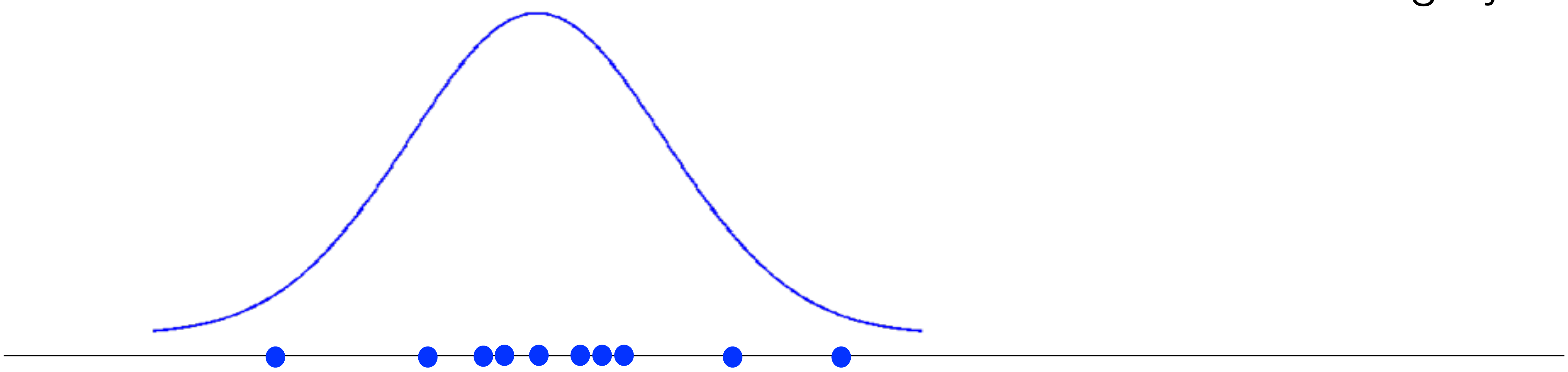
Working with probability distributions

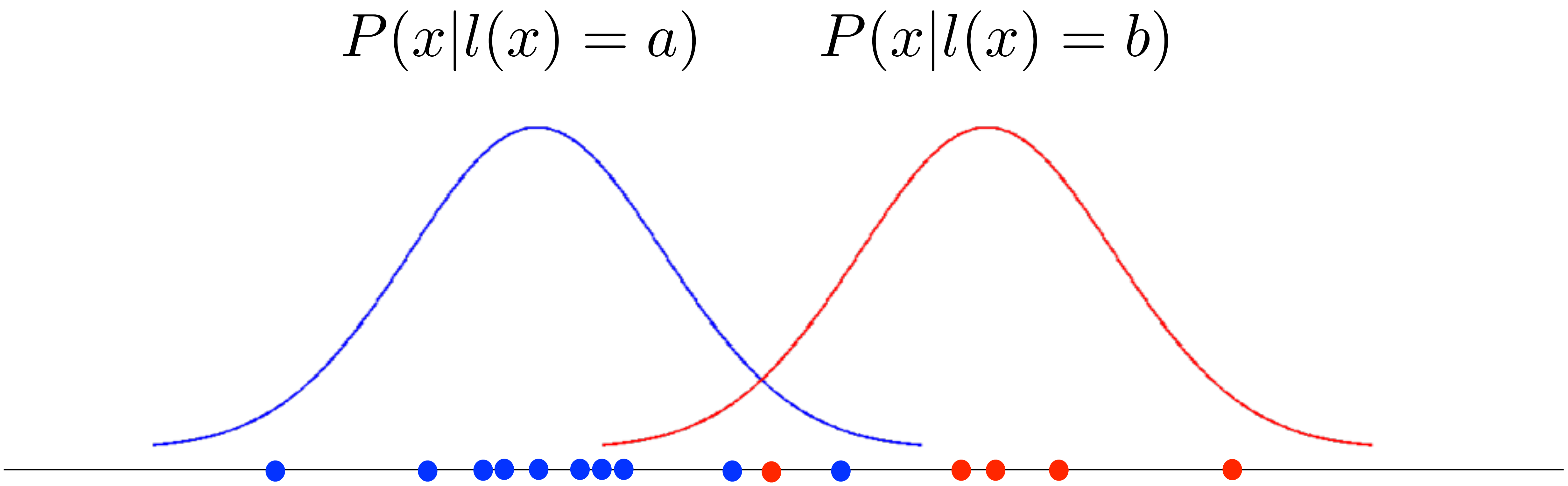
- R has a lot of functionality to let you play with distributions
- It's usually the same structure:
  - dnorm(), dbinom() - Probability (density) of a specific outcome
  - pnorm(), pbinom() - Chance that the outcome does not exceed some threshold
  - qnorm(), qbinom() - Compute some quantile of the distribution
  - rnorm(), rbinom() - Sample a random number from a distribution
- See also: gaussian() function in the classifiers.R code

Back to our classifier

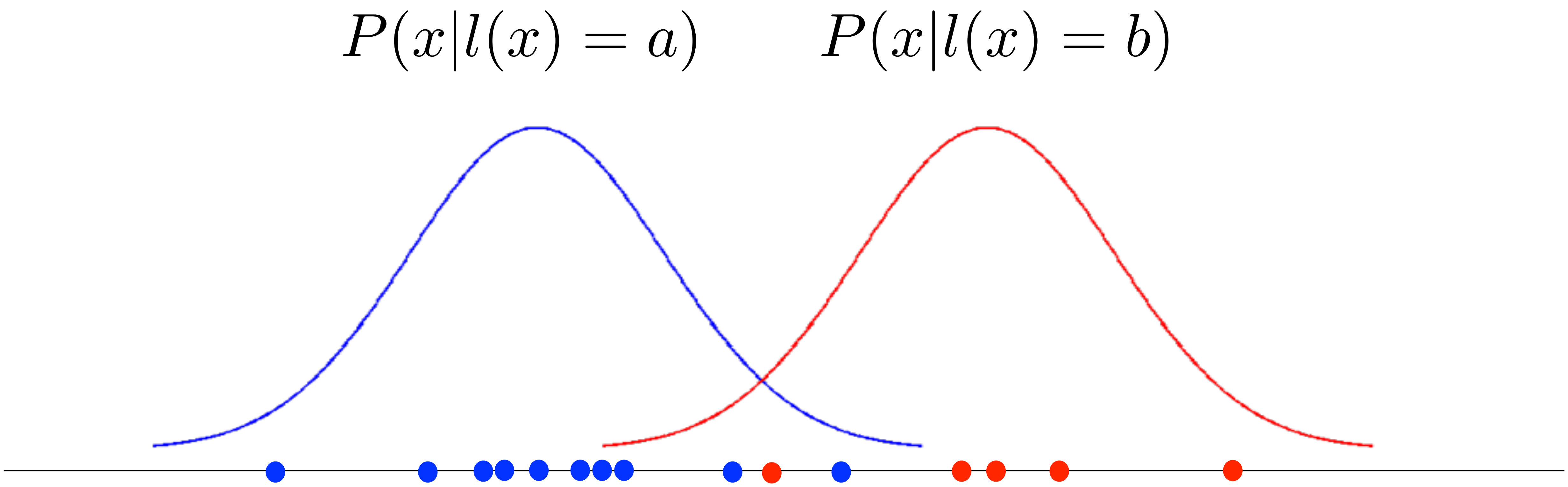
$$P(x|l(x) = a)$$

The probability distribution associated with category *a*

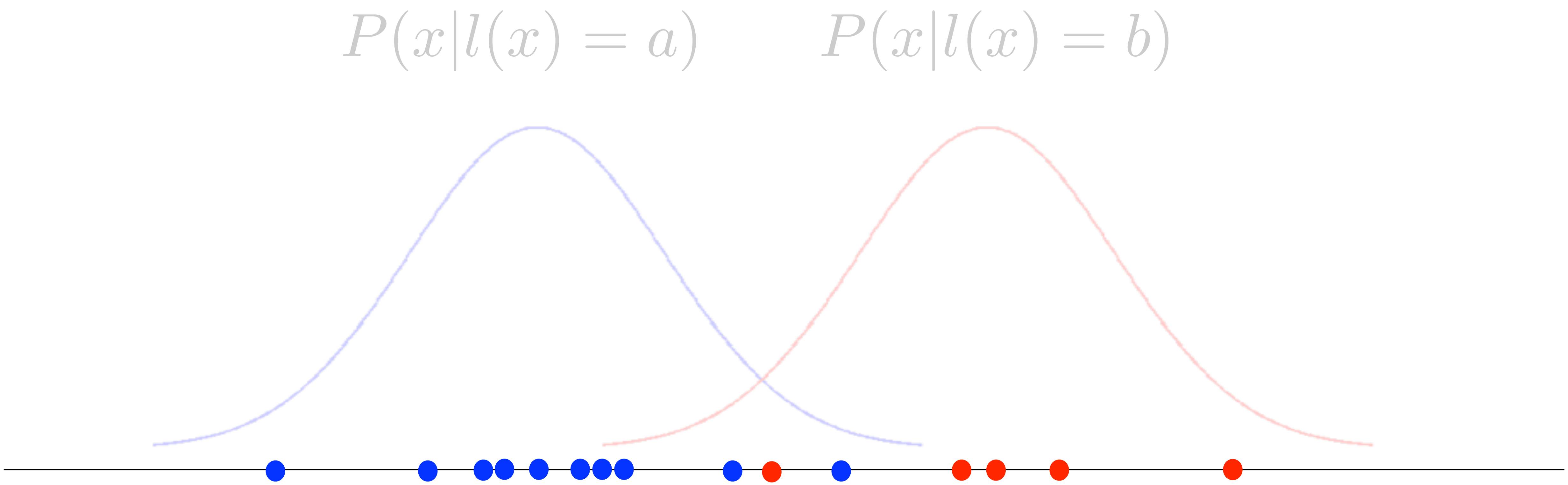




The probability distribution
associated with category b



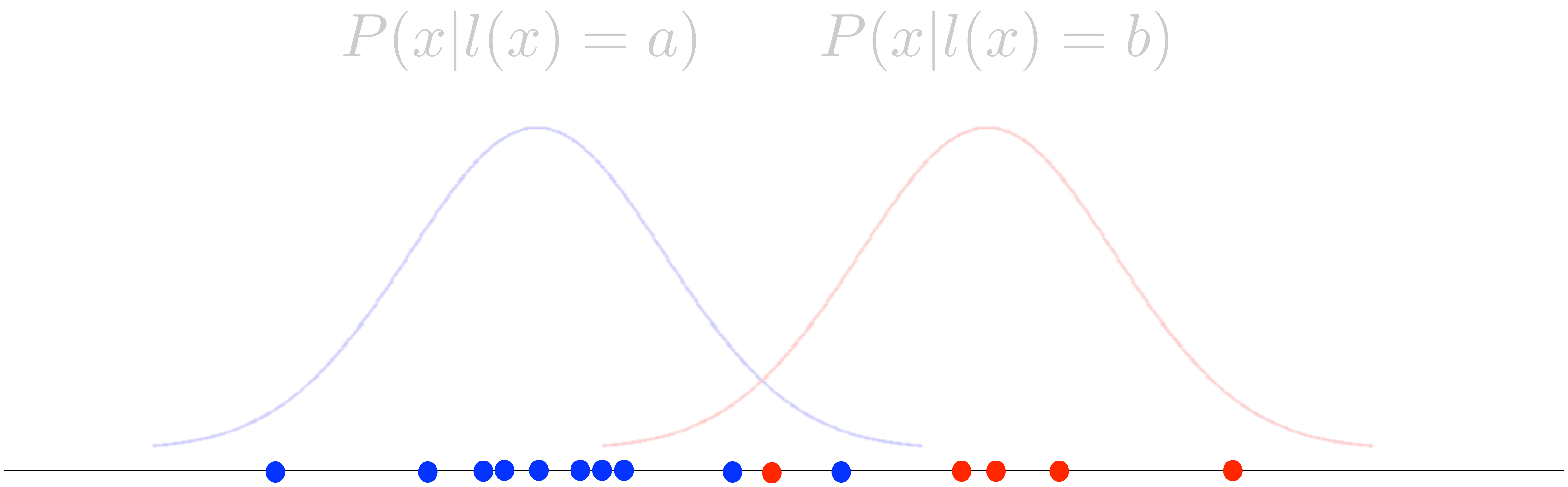
These define the **likelihood** functions associated with our two hypotheses (category a and category b)



We've observed 10 examples of category *a*

We've observed 5 examples of category *b*

$$P(l(x) = a) = 2/3$$



The base rate of the two types of observation defines our **prior** distribution

$$P(l(x) = a) = 2/3$$

A simple Bayesian classifier

Posterior probability that
item y belongs to category a

$$P(l(y) = a|y) \propto P(y|l(y) = a)P(l(y) = a)$$

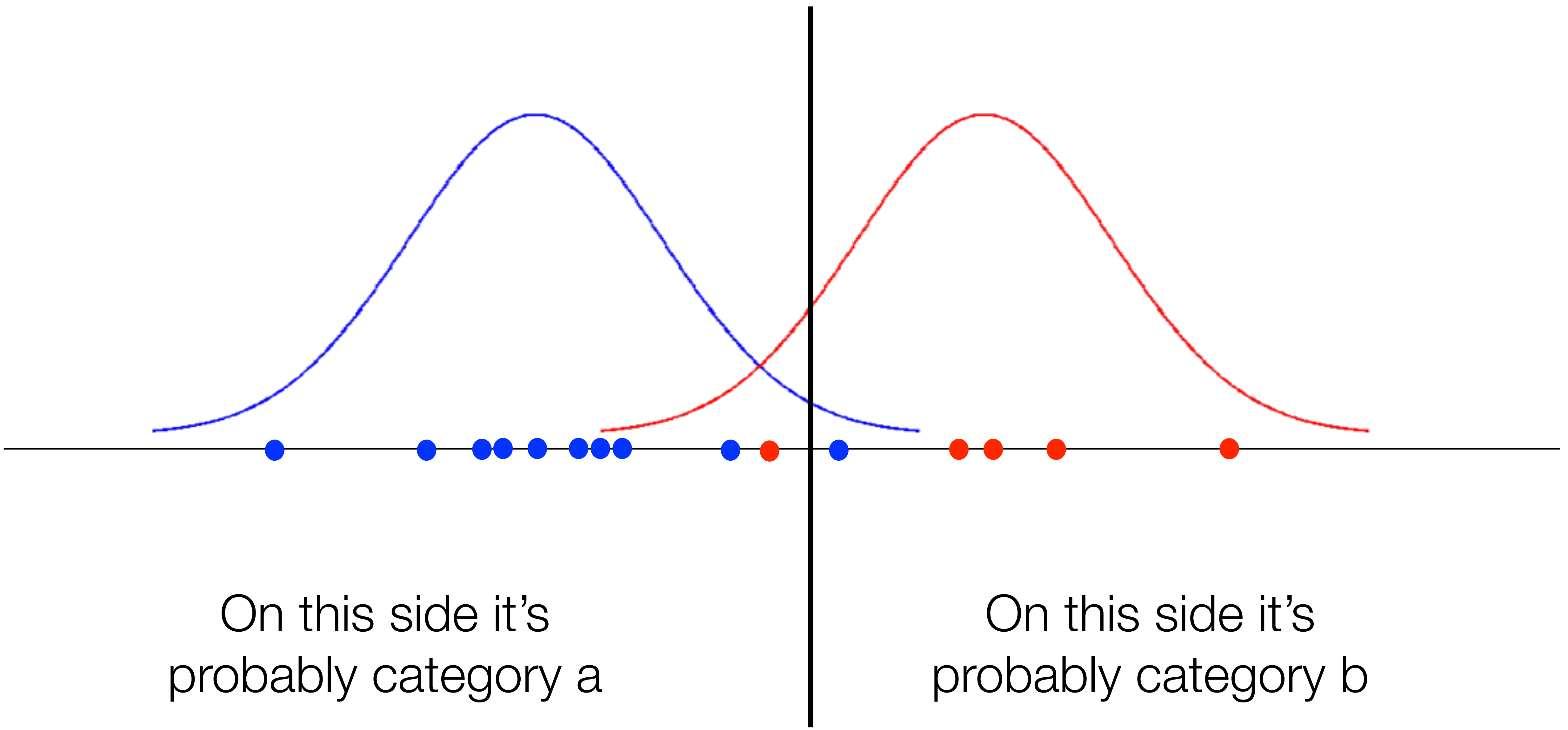
Likelihood of observing a
member of category a at the
location of item y

Prior probability that a new
observation would belong to
category a

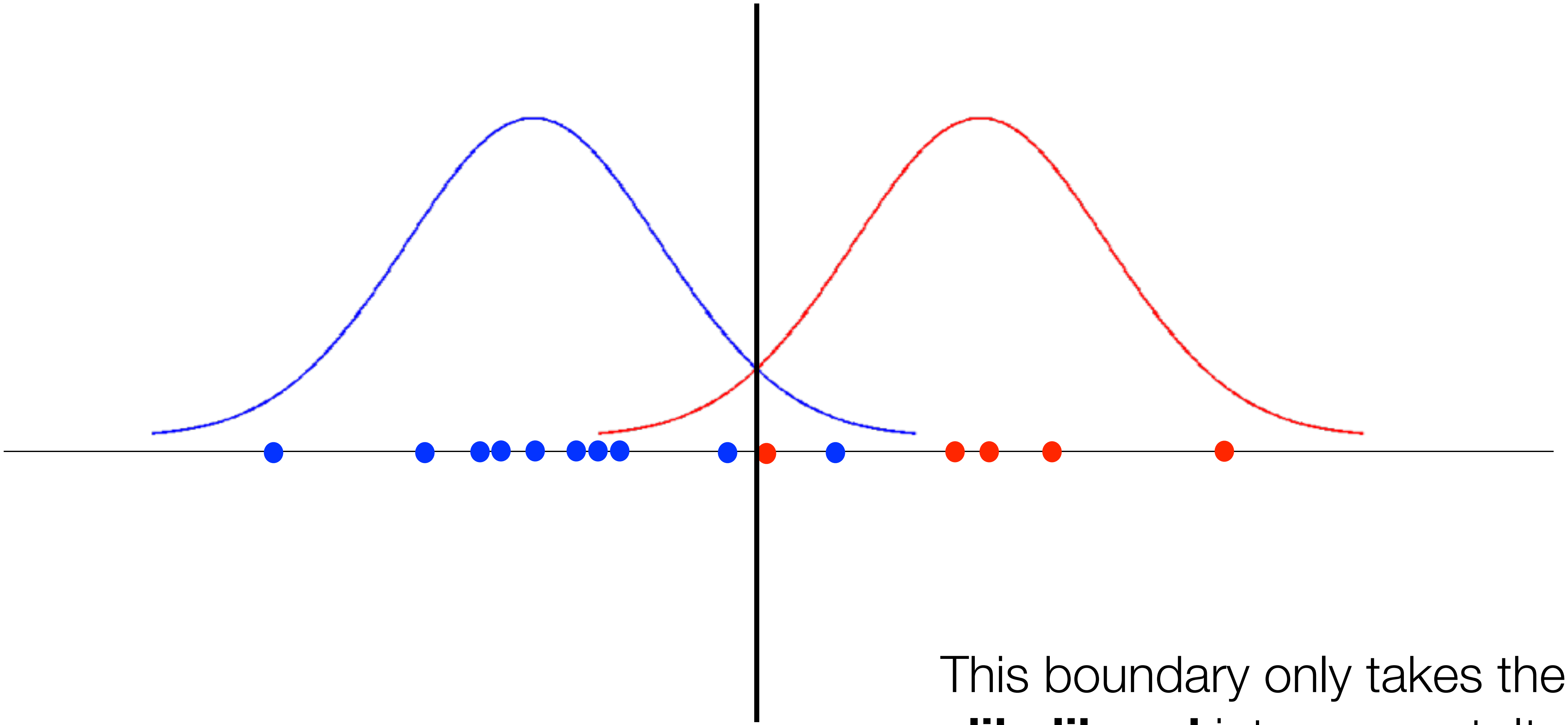
It's a bit uglier when we include the denominator term, but it's not conceptually difficult

$$P(l(y) = a|y) = \frac{P(y|l(y) = a)P(l(y) = a)}{P(y|l(y) = a)P(l(y) = a) + P(y|l(y) = b)P(l(y) = b)}$$

The Bayesian classifier…

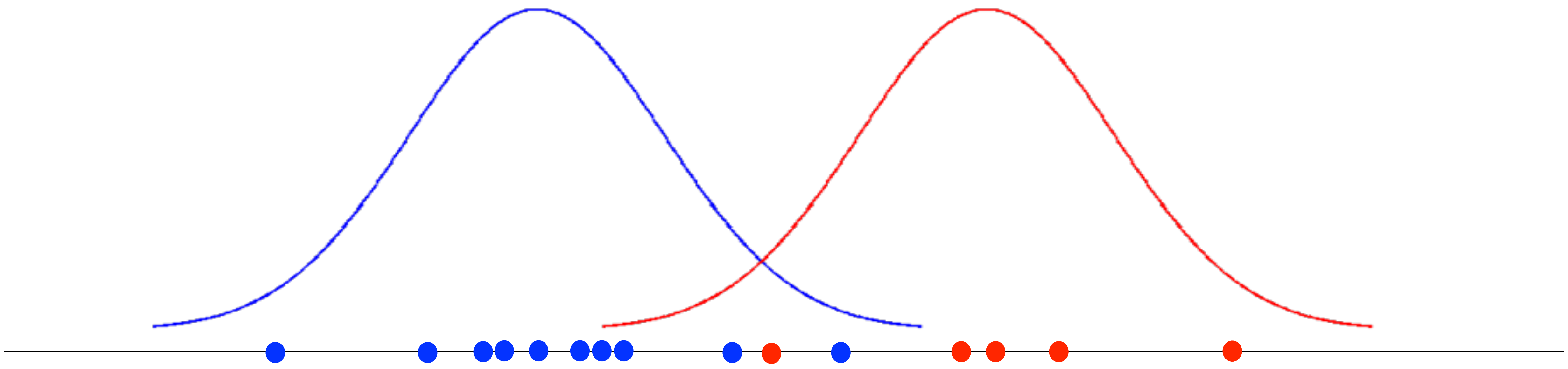


Why isn’t the “boundary” here?



This boundary only takes the **likelihood** into account. It ignores the fact that the **prior** favours category a

Pedantic comment



Our classifier has to estimate five parameters:

μ_a : Mean of category a

μ_b : Mean of category b

σ_a : Standard deviation of category a

σ_b : Standard deviation of category b

θ : Base rate for category a

Pedantic comment

You should write down your priors over all these quantities, and do Bayesian inference over those too!

Meh.

I’m too lazy to do that for this class, especially since it doesn’t make a big difference except when you have very little data

Demonstration code

(classifiers.R, simpleGaussianClassifier function)

Summary

- Ideas from cognitive science
  - The classical view and why it fails
  - Two family resemblance views: prototypes and exemplars
  - Hints about richer structure?

- Ideas from statistical machine learning
  - Supervised, unsupervised and semi-supervised learning
  - A simple Gaussian classifier (linked to prototype models)

- Next time:
  - An extension of the Gaussian classifier
  - More classifiers (linked to exemplar models)