

Lecture 3: Similarity

[Disclaimer: These informal lecture notes are not intended to be comprehensive - there are some additional ideas in the lectures and lecture slides, textbook, tutorial materials etc. As always, the lectures themselves are the best guide for what is and is not examinable content. However, I hope they are useful in picking out the core content in each lecture.]

Part 1: Introduction

As always, we open with a William James quote. This time, we have James arguing referring to the idea that different thoughts, objects, entities and events have a certain "sameness" or similarity to them, and that

this sense of Sameness is the very keel and backbone of our thinking

The idea is kind of intuitively plausible. In everyday life we so often say things like "oh, X is like Y" with a very particular goal in mind, namely to use what we already know about Y as a method for making good guesses about (or to learn about) X. Sometimes this kind of "analogical reasoning" is done explicitly, but most of the time we're just relying on the similarities that just automatically come to mind out at us (e.g., this morning is like yesterday morning, so I should probably go through the same routines as I went through yesterday)

Similarity is flexible

In in the opening slides, I try to highlight some of the ways in which our "sense of similarity" is actually a pretty flexible and slippery thing, as it seems to cover some quite disparate things:

- These leaves are all green (low level perceptual match)
- This human looks like that cat (high level perceptual match)
- This book about Star Trek is like that picture of Star Trek (conceptual match across different formats)
- This book about The Culture is like that book about Star Trek (conceptual match within the same format)
- DNA is like a zipper (similarity based on physical property)
- DNA is like a blueprint (similarity based on functional role)

Each of these could quite legitimately be thought of as "similarity" but they're all rather different to each other. As with many things in cognition, similarity isn't really "one" thing, it's lots of different things that are ... um... well... similar to one another.

Similarity is useful

Okay, so if similarity refers to this variety of closely linked phenomena, what is it there for? The key idea that reappears throughout the literature on psychology, philosophy and other disciplines, is that similarity is there to help us cope with the "snowflake problem"... namely the fact that everything in the world is unique in some sense. As noted in the slides:

- No two people are the same
- No two events are the same
- No two objects are the same
- Nothing is **literally** the same as anything else

Because everything is unique, you will never encounter a situation in life where your previous experience tells you exactly what you need to know. We're always guessing. Similarity is there to help us make these guesses.

To highlight this, I referred to a couple of quotes by W.V.O Quine (an influential philosopher) and Greg Murphy (one of the top researchers in the psychology of concepts and reasoning), that are worth reproducing in full. From Quine (1969):

"Similarity, is fundamental for learning, knowledge and thought, for only our sense of similarity allows us to order things into kinds so that these can function as stimulus meanings. Reasonable expectation depends on the similarity of circumstances and on our tendency to expect that similar causes will have similar effects"

From Murphy (2002):

"Although I've never seen this particular tomato before, it is probably like other tomatoes I have eaten and so is edible."

How do we measure similarity?

The lecture lists several different methodologies:

- Confusability: probability of mistaking A for B
- Reaction time: time taken to distinguish A from B
- Forced choice: is X more like A or more like B?
- Likert scales: how similar is A to B?

Very often in the literature psychologists tend to treat each of these as measuring the "same" notion of similarity, but it is important in practice to recognise that different methods elicit slightly different notions of "similarity"

Example: A reaction time task can be heavily influenced by visual pop-out effects (see lecture 2, visual search), whereas a Likert scale rating task is generally less influenced by those factors because it doesn't require people to make judgments quickly.

Part 2: Simple theories of similarity

Geometric models

One of the earliest and most influential approaches to stimulus similarity is the "geometric approach". The idea was prominently advocated by researchers like Roger Shepard in the 1970s and later, but you can trace back the history of the idea to methodological tools for "multidimensional scaling" (don't worry, you don't need to know this term) going back at least to the 1930s.

The simplest way to think about it is in terms of objects that possess multiple characteristics. For instance, colours are often represented in terms of *hue*, *saturation* and *value* (HSV), though there are more sophisticated ways of describing human colour perception (e.g. [CIELAB colour space](#)). More generally though, you could imagine that I think about a ball using perceptual dimensions like *size*, *weight* etc. So when thinking about the similarity between a tennis ball and a cricket ball, I might assess the *distance* between them (in "ball space") by comparing the differences in their sizes, the differences in their weight, etc.

The critical idea in this approach is the claim that these "psychological spaces" can be thought of as if they were geometric entities, and if so then ideas like "distance" should be (a) meaningful, and (b) closely related to "similarity". Entities that are near each other within psychological space (e.g., *yellow* is near *orange*) are treated as similar, whereas entities that are far apart (e.g., *orange* is distant from *purple*) are treated as dissimilar.

The idea is intuitively plausible, and there's some evidence for it. In the lecture, we talked about Shepard's "universal law of generalisation" which proposes that there is a consistent relationship between "distance in psychological space" and the "probability of generalising from one thing to another" (e.g., if this yellow fruit is poisonous, what is the probability that this orange fruit is also poisonous?) Across a pretty wide range of data sets, covering different species, sensory modalities and stimulus sets, there seems to be a consistent pattern, in which the shape of the relationship between the two looks like an exponential decay (see slides 45-46)

Featural models

The geometric approach does have some problems, however. For instance, one problem is that it makes a strong claim that similarity should be *symmetric*: yellow should be as similar to orange as orange is to yellow. When we're talking about colours that seems to make sense, but as we talked about in lecture, there are exceptions. Which is more "natural" to say:

- "A horse is like an okapi" OR "An okapi is like a horse"
- "Earth is like Venus" OR "Venus is like Earth"
- "A bird is like a bat" OR "A bat is like a bird"

As these examples illustrate (see also the pictures in the slides), we have a definite preference to say

[unusual thing] is like [usual thing] or to say [unfamiliar thing] is like [familiar thing]. A geometric approach can't easily explain this, because the distance from A to B is always the same as the distance from B to A.

To address this, Amos Tversky proposed a "featural" approach to similarity. Tversky's idea was that we mentally represent objects in terms of discrete "features" that they either possess or do not possess. For instance, a horse "has hooves", "is quadrupedal", etc. When we start thinking about mental representation in this way, we notice that for some entities, we have very detailed representations: we can easily list of many many features of horses, but comparatively few for okapis. Similarly, we know a lot about the planet Earth, and much less about Venus. We know more about birds than about bats. And so on. According to the featural approach, it is this asymmetry that produces asymmetric similarity.

Example: Here are some things I know about Tyrannosaurus Rex...

- a dinosaur
- a predator
- big
- was in Jurassic Park
- Sue the T. Rex is a famous skeleton
- there was a band called T. Rex
- had tiny arms

In comparison, here's what I know about Allosaurus

- a dinosaur
- a predator
- big
- is not a T. Rex

When you compare these two lists, notice that there are 3 *common features* shared by both T. Rex and Allosaurus. However, while I can think of 4 *distinctive features* that a T. Rex possesses, I can only think of a 1 for Allosaurus.

Tversky's idea, then, is that when I ask "is an Allosaurus like a T. Rex", I focus on the second list, and I notice that 75% (3 out of 4) of the features possessed by an Allosaurus are also possessed by T. Rex... so I judge them to be very similar. However, when I ask "is a T. Rex like an Allosaurus" I focus on the first list, and only 43% (3 out of 7) T. Rex features are also possessed by Allosaurs, and so I judge them to be less similar.

Part 3: Structured models of similarity

The geometric approach and the featural approach are somewhat different to each other, but they're both "simple" theories of similarity in the sense that they describe similarity in terms of fairly simple mental representations (dimensions & features). More recent theories have tended to emphasise the fact that our

mental models of the world tend to be a lot more *structured* than this, and this structure shapes the way we perceive similarity. In the lecture we talked about two different ideas about how this could matter.

Structure alignment

The "structure alignment" view emphasises the fact that our mental models of entities don't merely list "features". Rather, we also acknowledge that these features attach to a particular "role" or "slot". For instance, that list of Allosaurus features I gave above should probably have been written more like this:

- biological category: dinosaur
- biological role: predator
- size: big

In and of itself, this doesn't seem all that important, but as we noted in the lecture, it does lead to some neat predictions. Consider the "cartoon bugs" on slides 62-70, which can be described using the following structured lists:

```
Bug A {
  head: yellow
  body: brown
  wings: blue
}

Bug B {
  head: yellow
  body: green
  wings: blue
}

Bug C {
  head: green
  body: yellow
  wings: blue
}
```

As we showed in the lecture, most people intuitively think Bugs A and B are more similar to each other than Bugs A and C. All three bugs are composed of the same three body parts (*head*, *body*, *wings*) and match on two of the three colours (*yellow* and *blue*), yet we intuitively feel that the fact that Bugs A and B both have *yellow heads* is important, and so we judge these to be more similar.

This leads to the terminology:

- A *match in place* (MIP): occurs when two objects have the same feature (*yellow*) in the same slot (*head*).
- A *match out of place* (MOP): occurs when two objects share a feature (*yellow*), but it occurs in different

places (head vs body)

In lecture, we described an experiment by Rob Goldstone (1994) in which he systematically manipulated the number of MIPs and MOPs shared by different objects, and showed that both kinds of match will tend to increase the perceived similarity between entities; but that the effect of MIPs is much larger than the effect of MOPs.

The conclusion from this kind of study is that the "structure" (i.e., the manner in which features are bound together into object representations - compare to the "feature binding" comments in the attention lecture) of an object plays a very important role in shaping how we perceive its similarities to other objects.

Transformation

The final idea we talked about is the notion of stimulus *transformation*. According to this theory, the main idea is that "similarity mirrors processes", in the following sense - when we assess the similarity between two entities, we are sensitive to the number of "mental operations" that are required to transform one object into another. These kinds of mental operations don't necessarily have to be physically possible (e.g., in lecture I gave a silly example where "delete human" and "add cat" were listed as mental transformations), just easily imagined. For visual stimuli, the idea is not dissimilar to a kind of mental "photoshop". In real life, transforming a short person into a tall person would be terribly difficult if not impossible - but as a mental transformation it's not too difficult to do at all.

The transformational approach leads to some unique predictions of its own. For instance, consider these two bug stimuli:

```
Bug A {  
  head: yellow  
  body: yellow  
}  
  
Bug B {  
  head: yellow  
  body: blue  
}
```

If I wanted to transform Bug A into Bug B using a typical image manipulation program like photoshop, it would take two steps -- I'd first have to "create" blue (by selecting it from the palette), and then I'd have to "fill" the body with this colour. Going the other way (turning B into A), however, only takes one step: because the image of Bug B already has yellow in it, I can go straight to "filling" the body with yellow.

Admittedly, in real photo editing it's a bit messier than this story, and of course real world similarity judgment is also messier than this, but hopefully you can get a sense of what the theory implies... Among other things it offers an explanation of asymmetric similarity. If it's easier to transform A into B than vice versa, it means we will perceive the similarity from A to B as higher than the similarity from B to A.

In the lecture, I described an experiment that aimed to test this idea (Hodgett & Hahn 2012 - see slide 92 in particular) in which they showed exactly this effect on reaction times. In their task, they presented one stimulus (A) on screen briefly, then replaced it with another one (B) shortly afterwards and asked people to judge if it was the same object or a different one. The experiment included many different pairs of stimuli, and (of course) on half the trials the second stimulus was the same as the first one (so you can't guess!). Critically, the design included cases where the sequence went from A to B, and other trials with the same two items in reverse order (from B to A). The items A and B were chosen so that one direction was predicted to be "shorter" (i.e., fewer transformations, higher similarity) than the other; and what they found was that people were slightly slower (about 12ms) to perceive the difference in those cases, suggesting that the transformational theory does successfully predict asymmetric similarities.

Summary

- Intro
 - What similarity is
 - How it is measured
- Geometric similarity
 - "Universal" law of generalisation
 - Symmetry prediction
- Featural similarity
 - Asymmetry due to different knowledge
- Structure alignment
 - MIPs and MOPs
 - Goldstone experiment
- Stimulus transformation
 - Asymmetry due to different structure
 - Hodgetts & Hahn experiment