Reference Production Facilitates Listener Search Across Modalities ...but *Colour* is Special

Merrick Giles
Paula Rubio-Fernandez
Francis Mollica

Reference should be informative

“Look at the bird”

Grice, 1975; Zipf, 1949

The goal of reference is to point something out. For the past many decades linguists proposed that languages are very efficient: people say as much as they need, but no more.

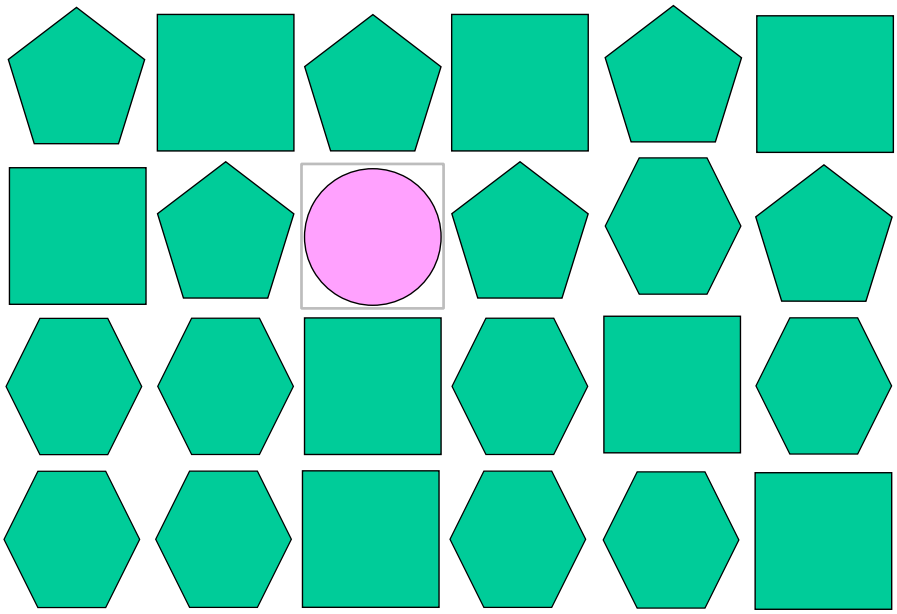
But it's *Overi*nformative

"Look at the ***blue*** bird"

Pechmann, 1985; Sedivy, 2003

But people frequently say more than what they need: they say look at the *blue* bird. Why are wasting each other's time? There's only one bird, I can obviously see it's blue.

Search Efficiency



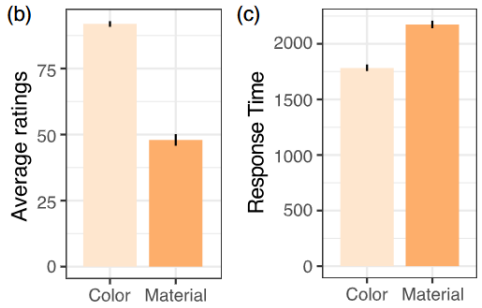
Jara-Ettinger & Rubio-Fernandez, 2022

The best explanation we have is search efficiency. Speakers don't just want you to find what they're talking about, they want you to find it quickly. And that means the game is not just to inform, but to search.

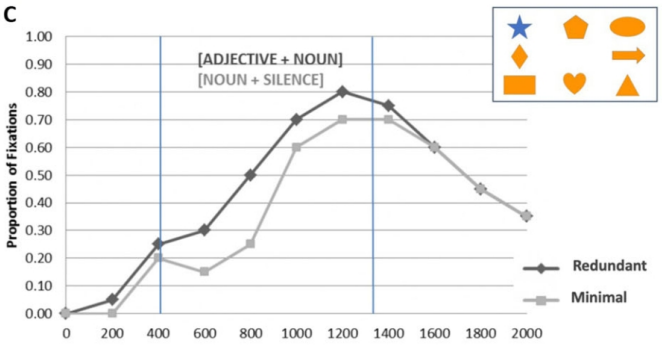
So even though there's only one circle here, you might help out the listener's search by saying 'pink circle'.

Production → Comprehension

Jara-Ettinger & Rubio-Fernandez, 2022



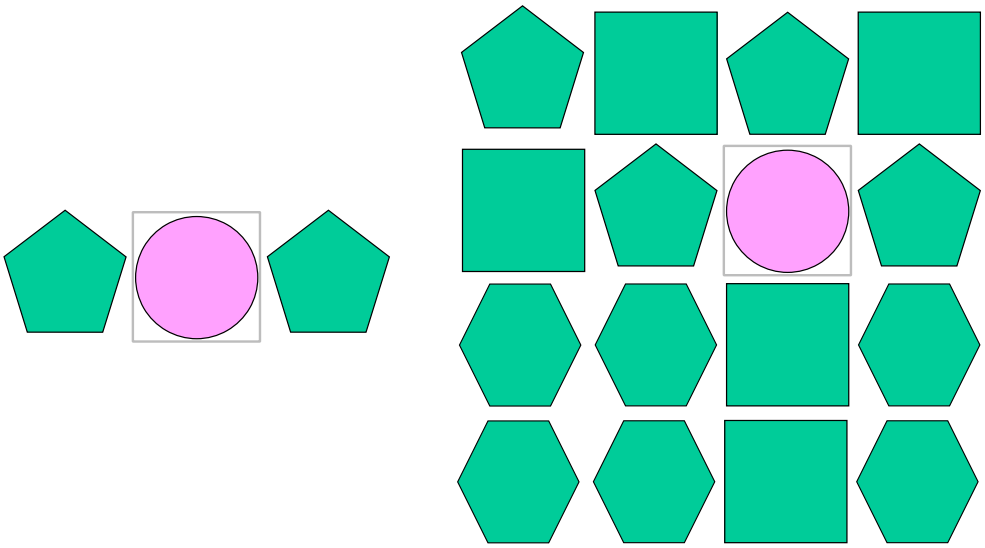
Rubio-Fernandez, 2021



Also: Kursat & Degen, 2021; Tourtouri et al., 2019; Rehrig et al., 2021, and more.

And the search efficiency drivers in production translate really well to comprehension: so people tend to overinform when the redundancy is helpful for search – as measured by response times, eye-tracking, eeg.

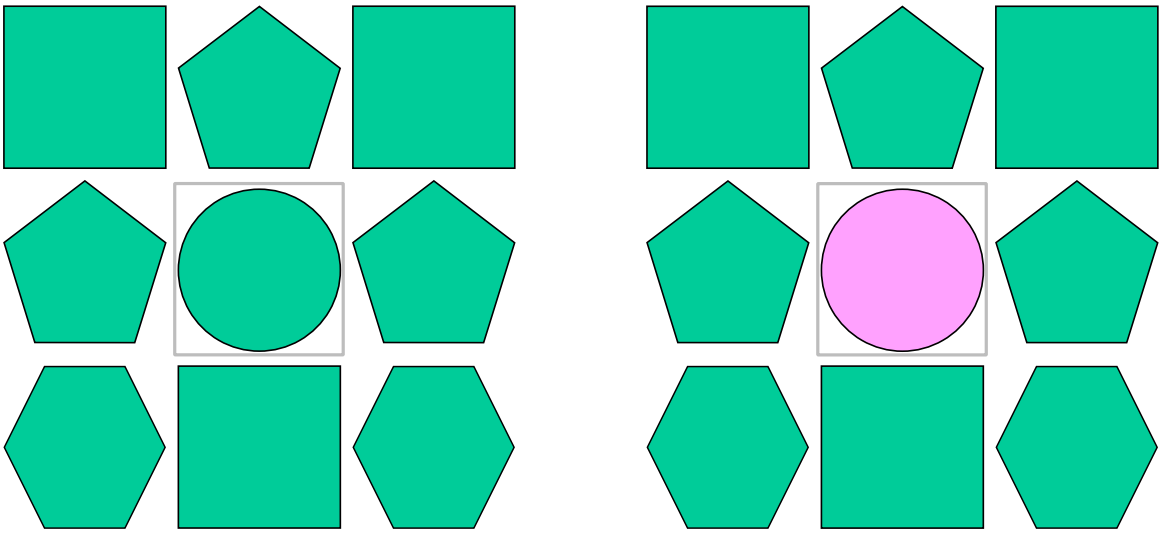
In dense displays, people overinform a lot.



Rubio-Fernandez, 2019; Degen et al. 2020

So I'll show two examples where speakers seem real well calibrated to listener search. The first is this display density thing: In a sparse display like the left, barely anyone will overinform. But in a dense display like the one on the right, speakers overinform all the time – colour here helps cut through the challenge of a dense display

Overinforming scales with contrast

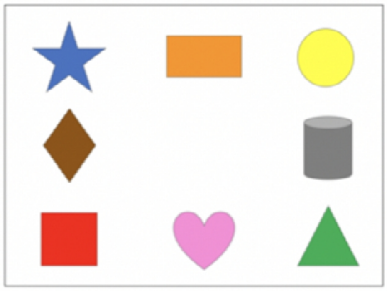


Jara-Ettinger & Rubio-Fernandez, 2022; Kursat & Degen, 2021; Viethen et al., 2017

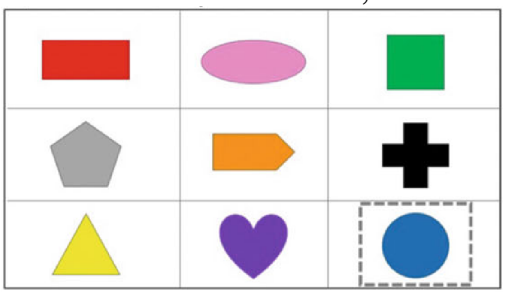
The second case is that overinforming scales with contrast. When redundancy isn't contrastive like on the left, you'll just say the circle. But if it's contrastive like on the right, you'll say the pink circle. And contrastive attributes capture visual attention, so mentioning them speeds up listener search.

A general theory of reference

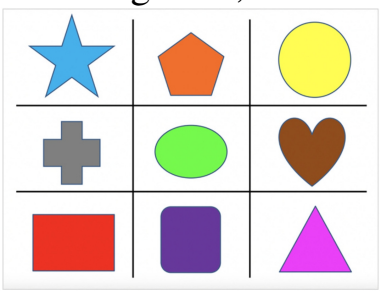
Rubio-Fernandez, 2019



Rubio-Fernandez, 2021



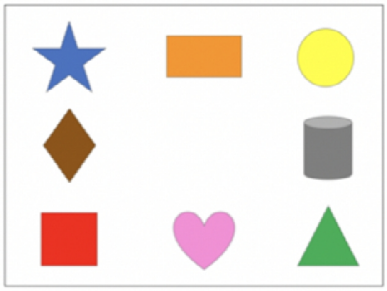
Long et al., 2021



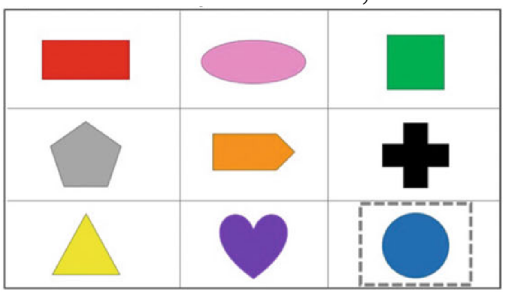
By now the predictions of Search Efficiency have been confirmed in a bunch of experiments – making it a strong candidate for a general theory of reference.

A general theory of reference?

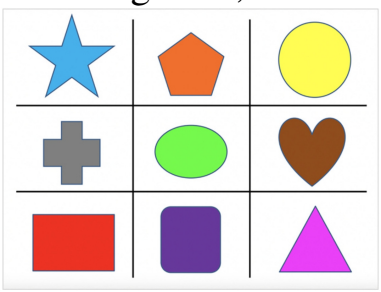
Rubio-Fernandez, 2019



Rubio-Fernandez, 2021



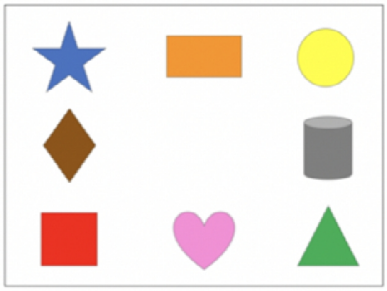
Long et al., 2021



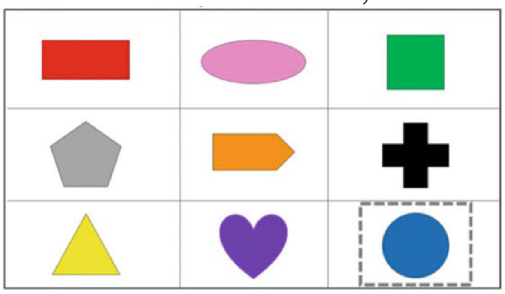
But if you've been paying close attention to this talk – which I know you all have -- you might have notice something about all the stimuli – search efficiency is really only being manipulated using super distinctive colour stimuli.

A general theory of reference?

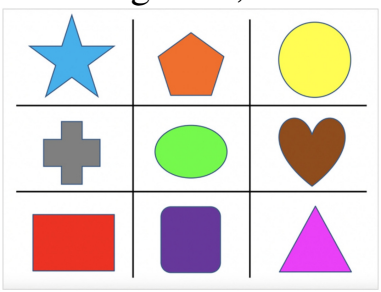
Rubio-Fernandez, 2019



Rubio-Fernandez, 2021



Long et al., 2021



Kursat & Degen, 2021

Jara-Ettinger & Rubio-Fernandez, 2022

Though to be fair, some have tried to move beyond colour to other *absolute* adjectives like material. But search efficiency manipulations haven't ever panned out for material in-itself. Or any other attribute as far as we know.

Today we will fix this. We will look at whether search efficiency drives reference for both colour AND material – is it strong as a general principle.

A general theory of reference?

How to validate Search Efficiency for reference of all kinds?

So why hasn't the whole search thing panned out for material? Well first, material is highly correlated with colour. So using stimuli like these is confounded with colour.

A general theory of reference?

How to validate Search Efficiency for reference of all kinds?

As soon as we remove colour, visual material is really hard to see. Like:

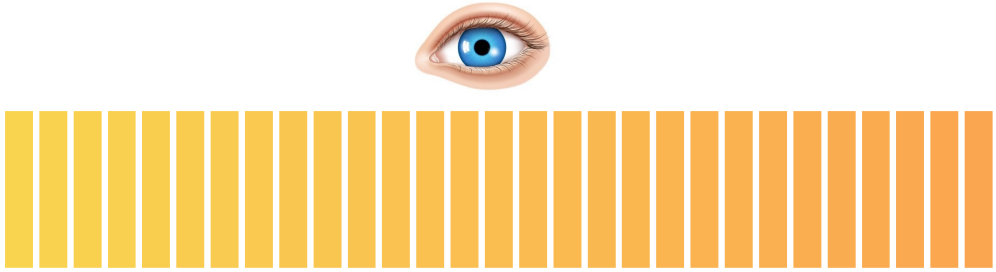
A general theory of reference?

How to validate Search Efficiency for reference of all kinds?

Entire game shows have been made about how hard material is to see.

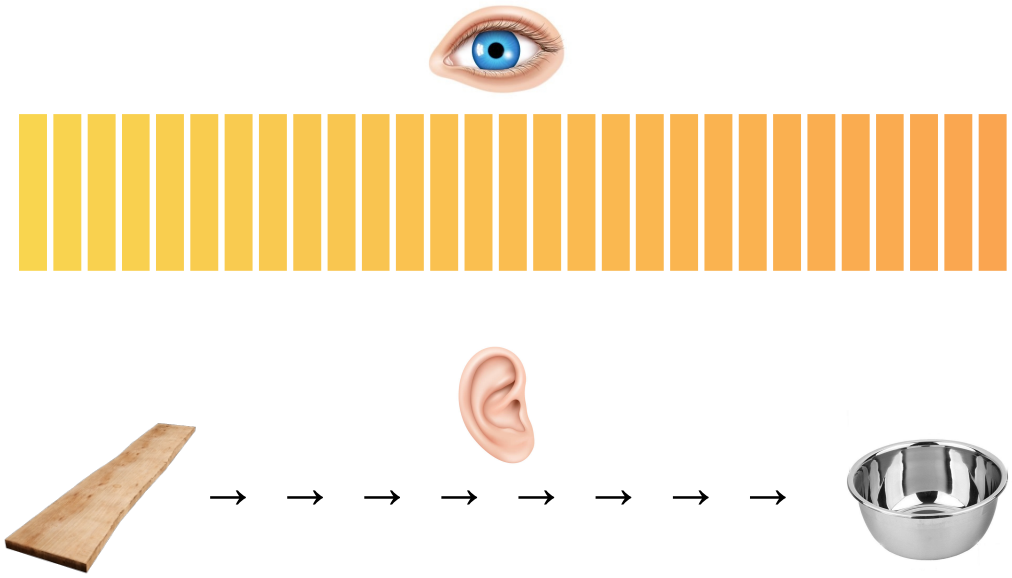
You can’t manipulate search efficiency here – material in-itself won’t be search efficient.

Does search efficiency go beyond colour?



And man, colour is just so good for search manipulations – it has this great perceptual range that we can use: it's a continuum.

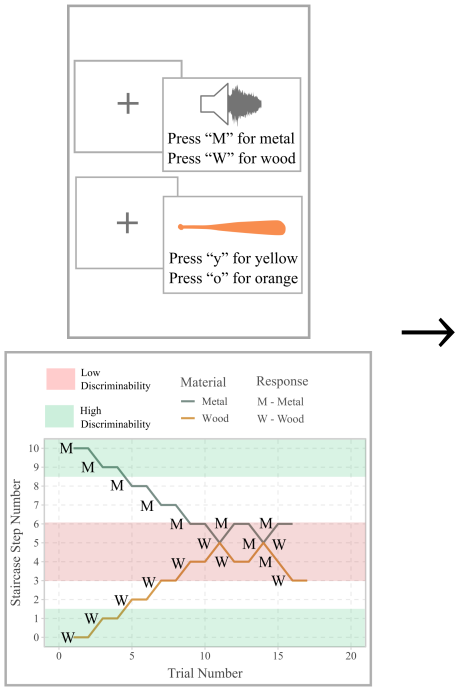
Does search efficiency go beyond colour?



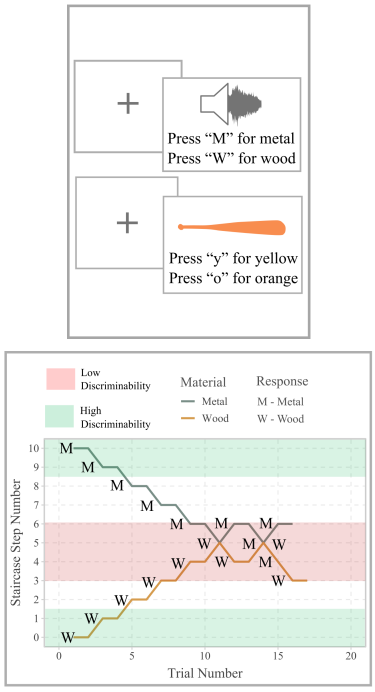
But we can make material a continuum by using sound! We can take the sound of wood hitting a surface and the sound of metal hitting a surface, and make a continuum between them.

So now we got material, we got perceptual range, and we got no colour association. So we can answer the question of whether this whole search efficiency thing generalise beyond colour?

We’re going to run an experiment and find out -- manipulating search efficiency using visual colour and auditory material.



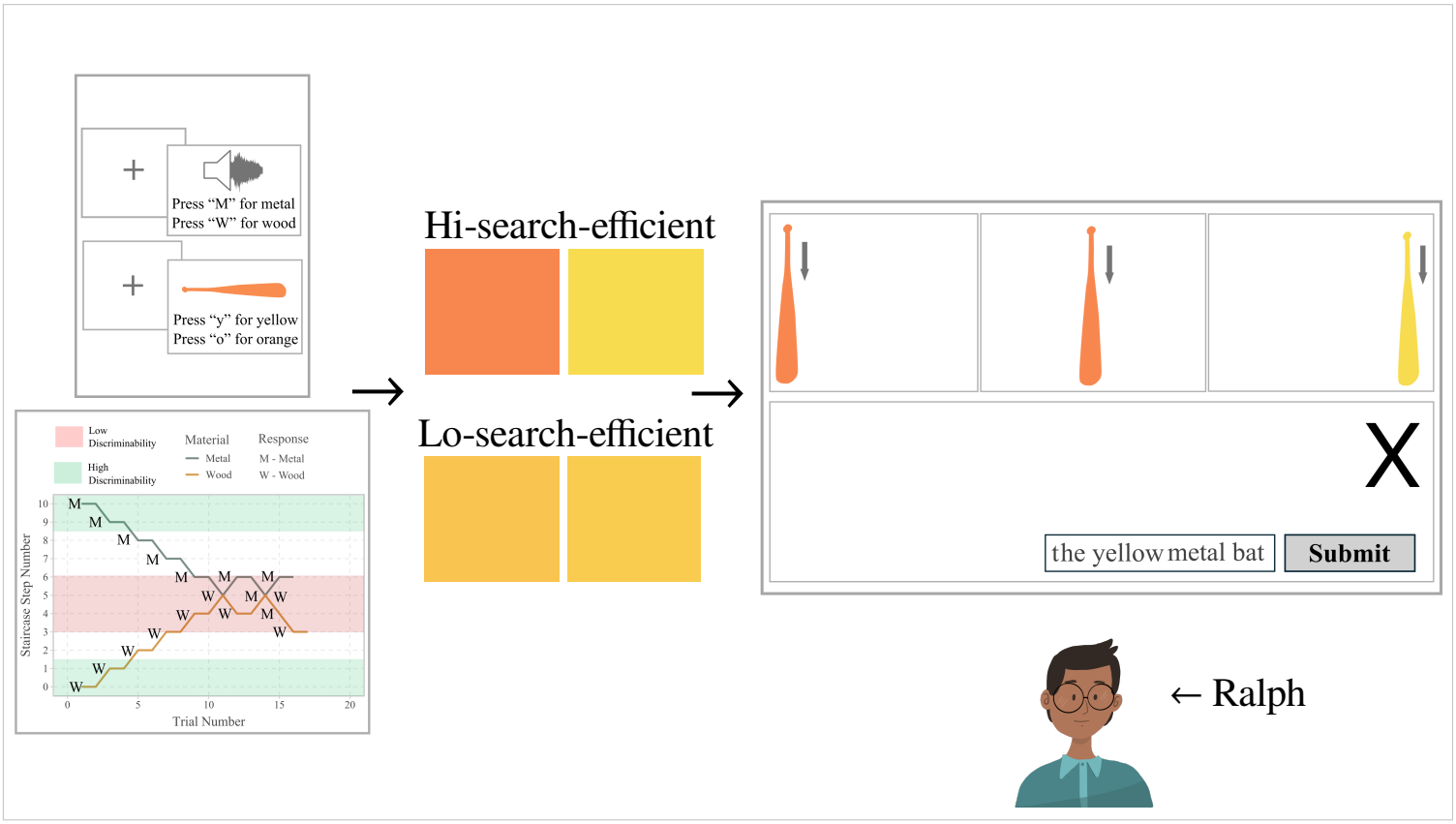
So let me give you a roadmap of the experiment. First we need a valid way to determine what peoples perceptual limits are – so what we’re going to do is classic psychophysical staircase.



Hi-search-efficient

Lo-search-efficient

The staircase lets us pull out high and low search efficient stimuli for colour and material – I’ve shown colour here but we of course have both.



Then. Taking the participants individualised stimuli, we run a variation of the classic director task. We needed a cover story to let us capture both visual and auditory events, so participants’ job was to be quality control at a baseball bat factory, describing the bats to their coworker ralph.

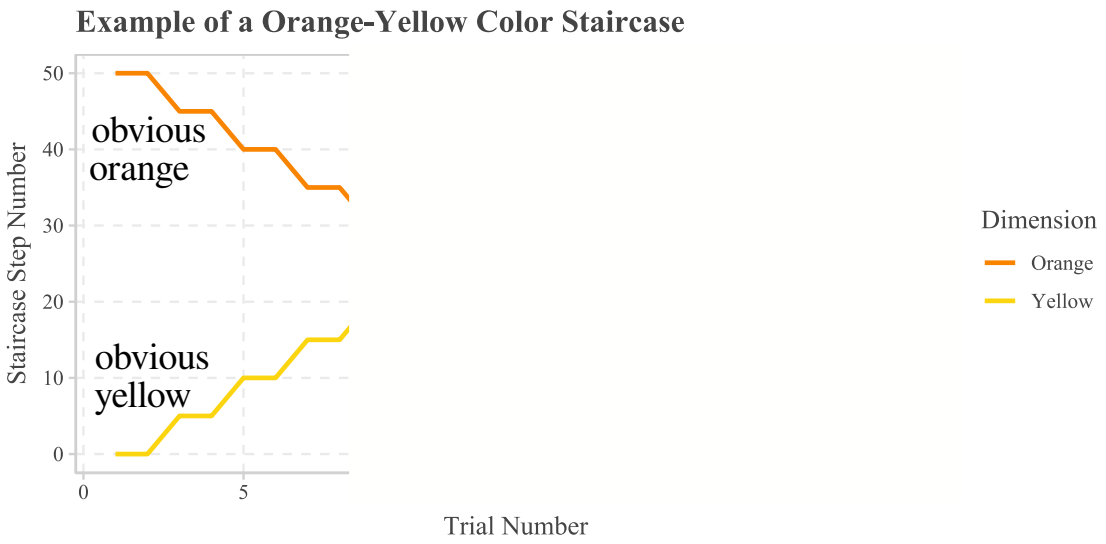
Wood


Metal

Wood

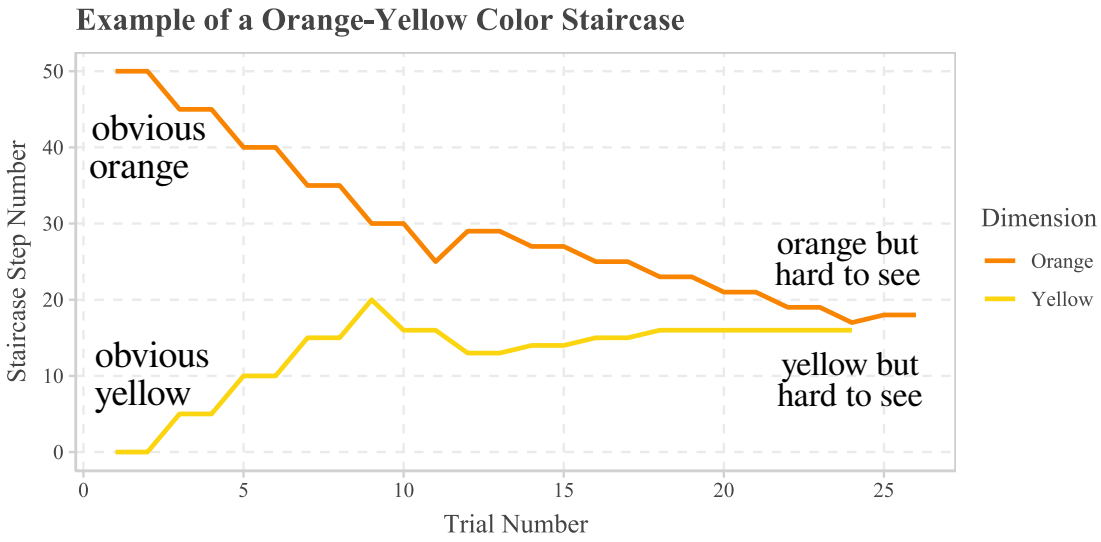
Type your description here:

Phase One: Stairs

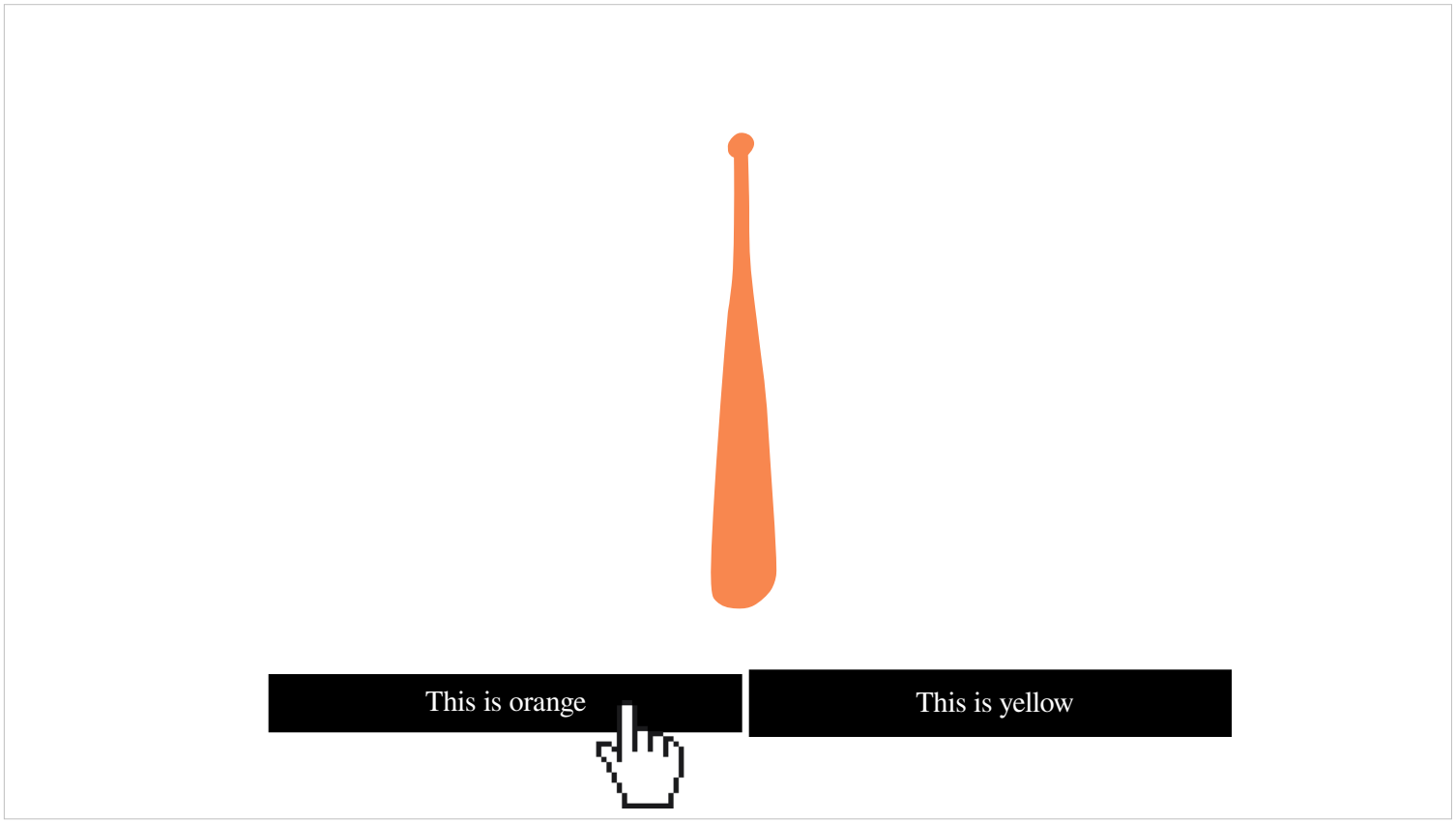


To make the stimuli, we need individual high and low points of discriminability. So we run this psychophysical staircase, where you classify stimuli as, say blue or green. And you start the task at really far ends of colour or material space so it’s really obvious.

Phase One: Stairs

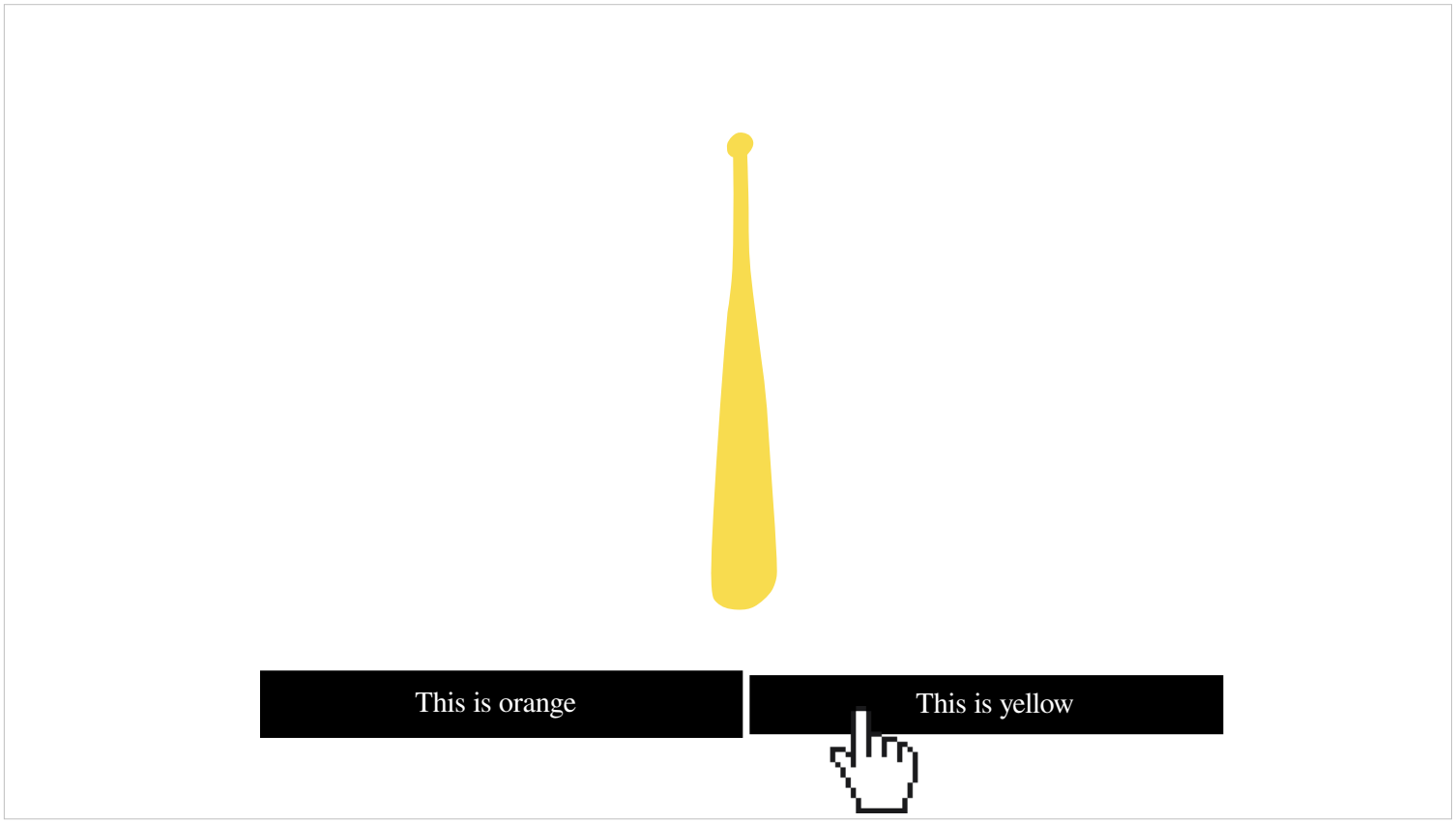


But as the task goes on you get closer in colour space and it becomes ambiguous.

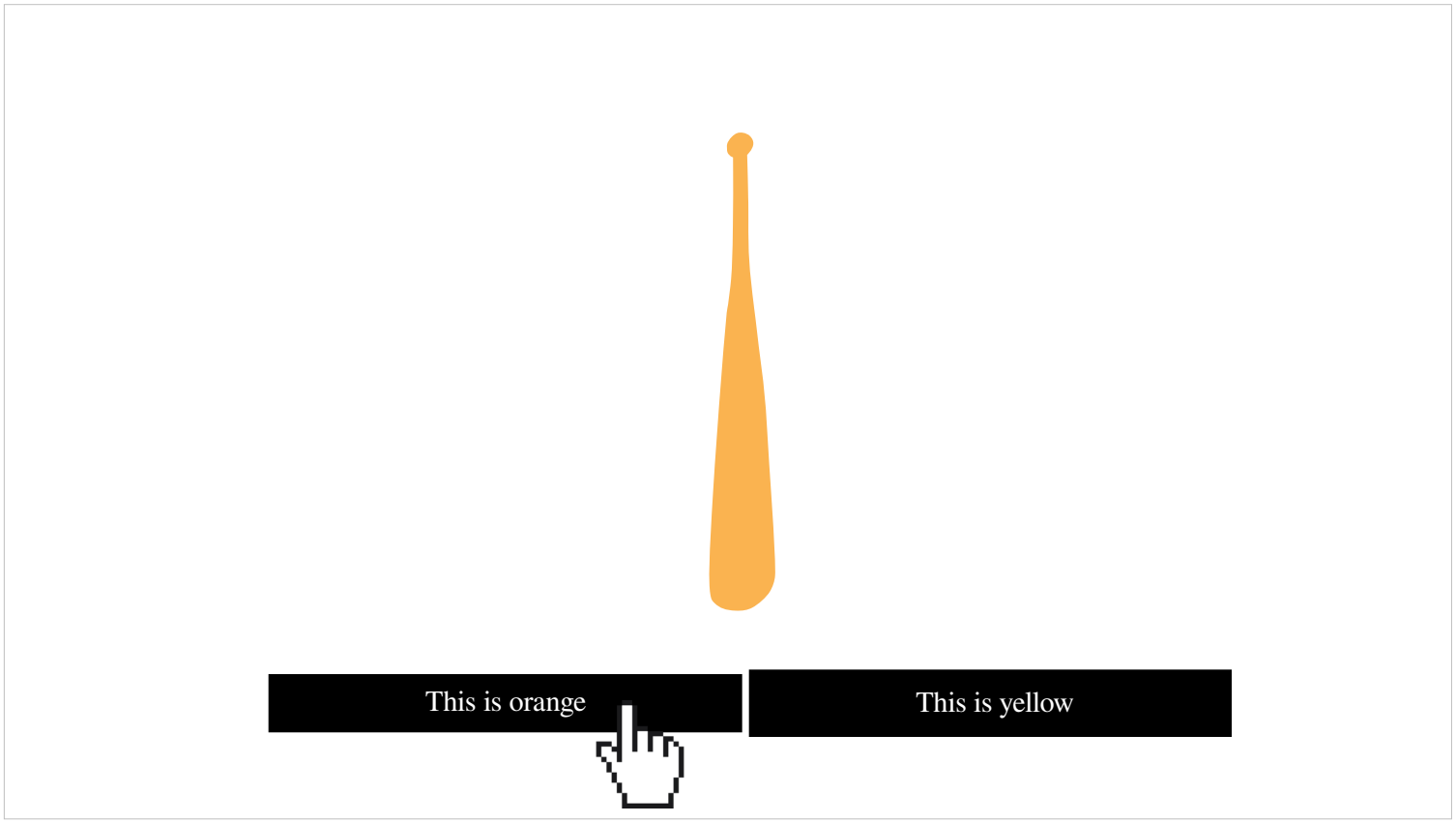


The actual task looks something like this.

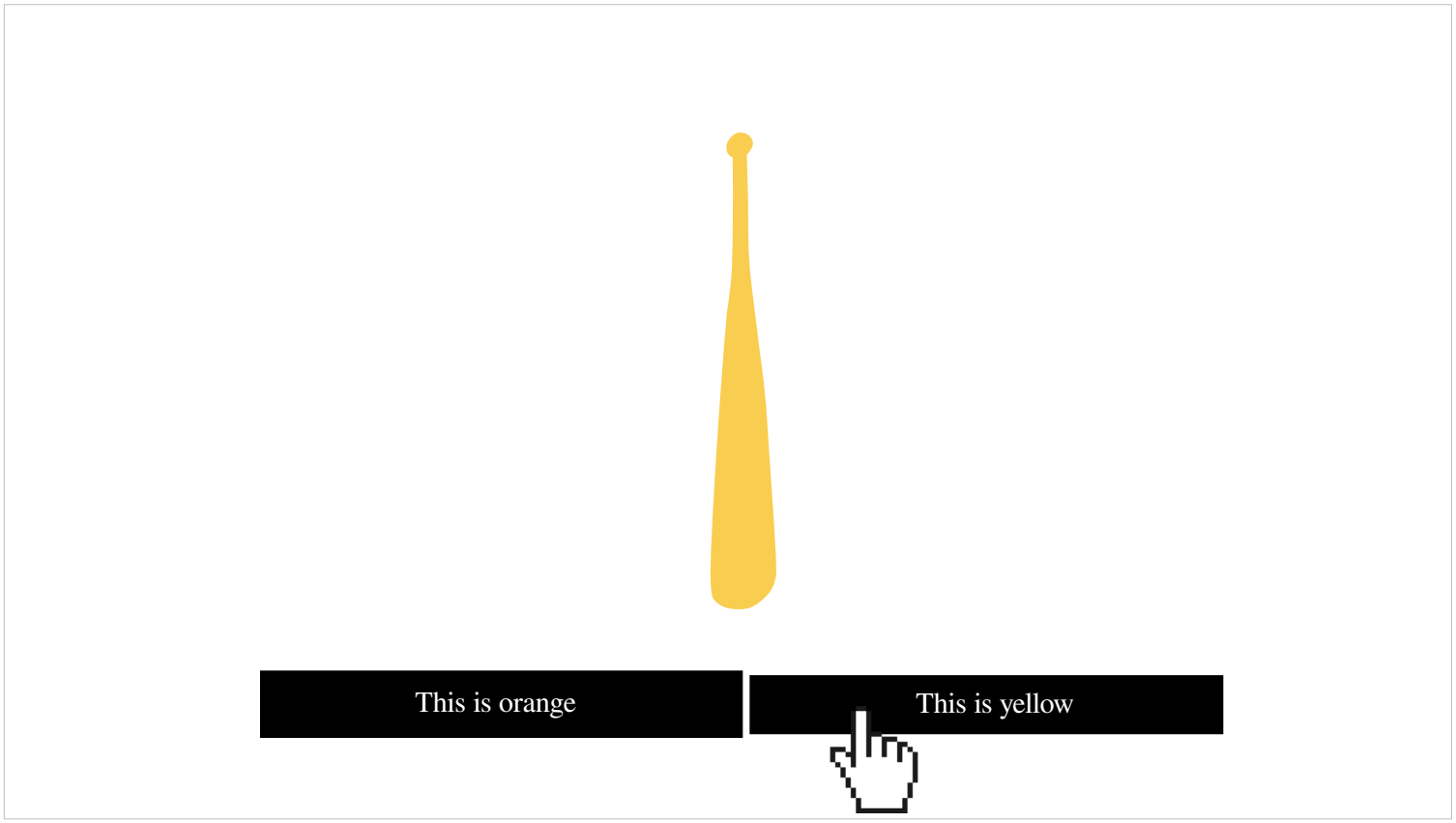
So we start out at the high-discriminability end of the continuum – we've started out at orange rather than yellow by chance.



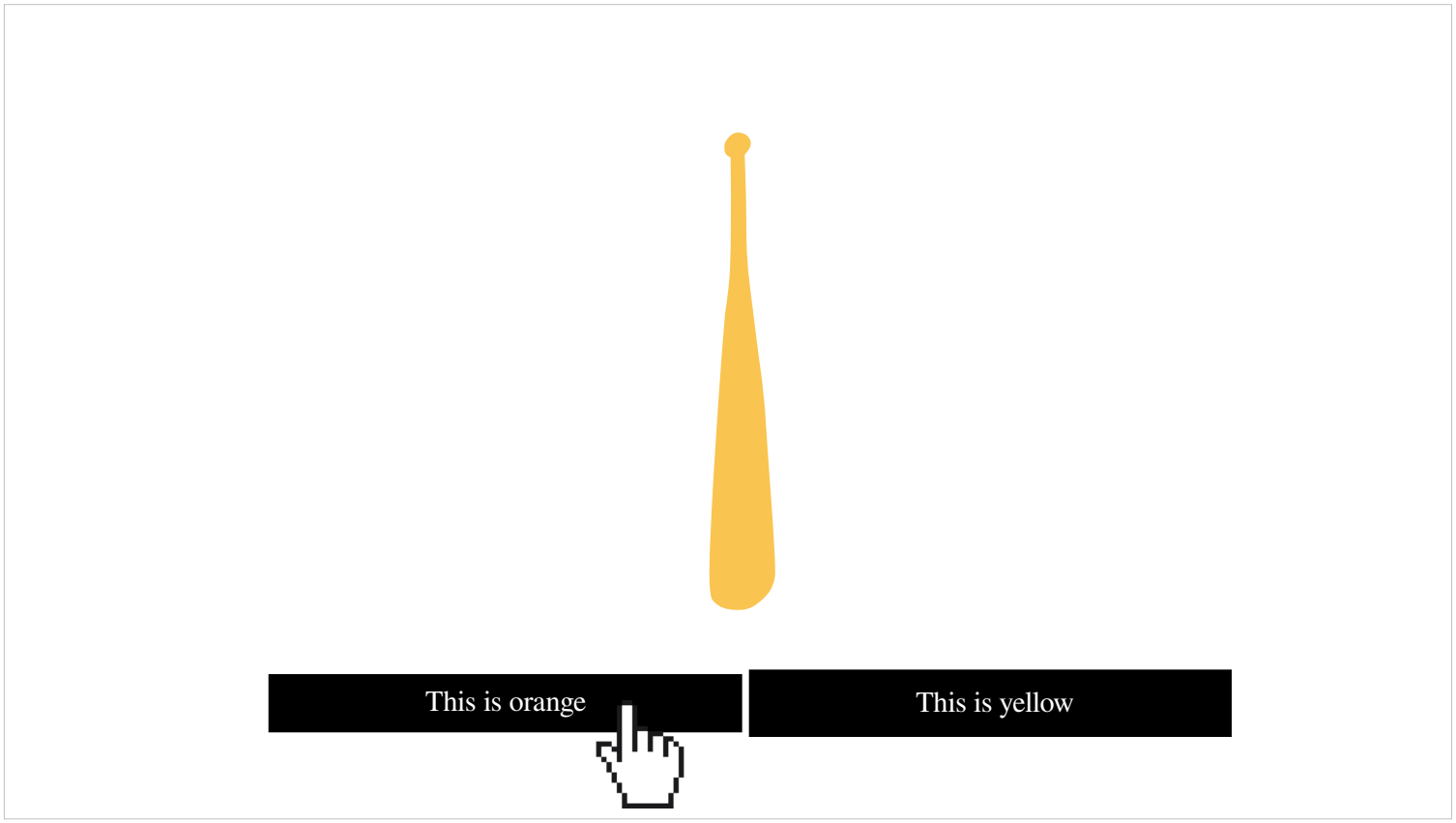
Then we move to the high discrim at the other end – so we got the high-discrim yellow.



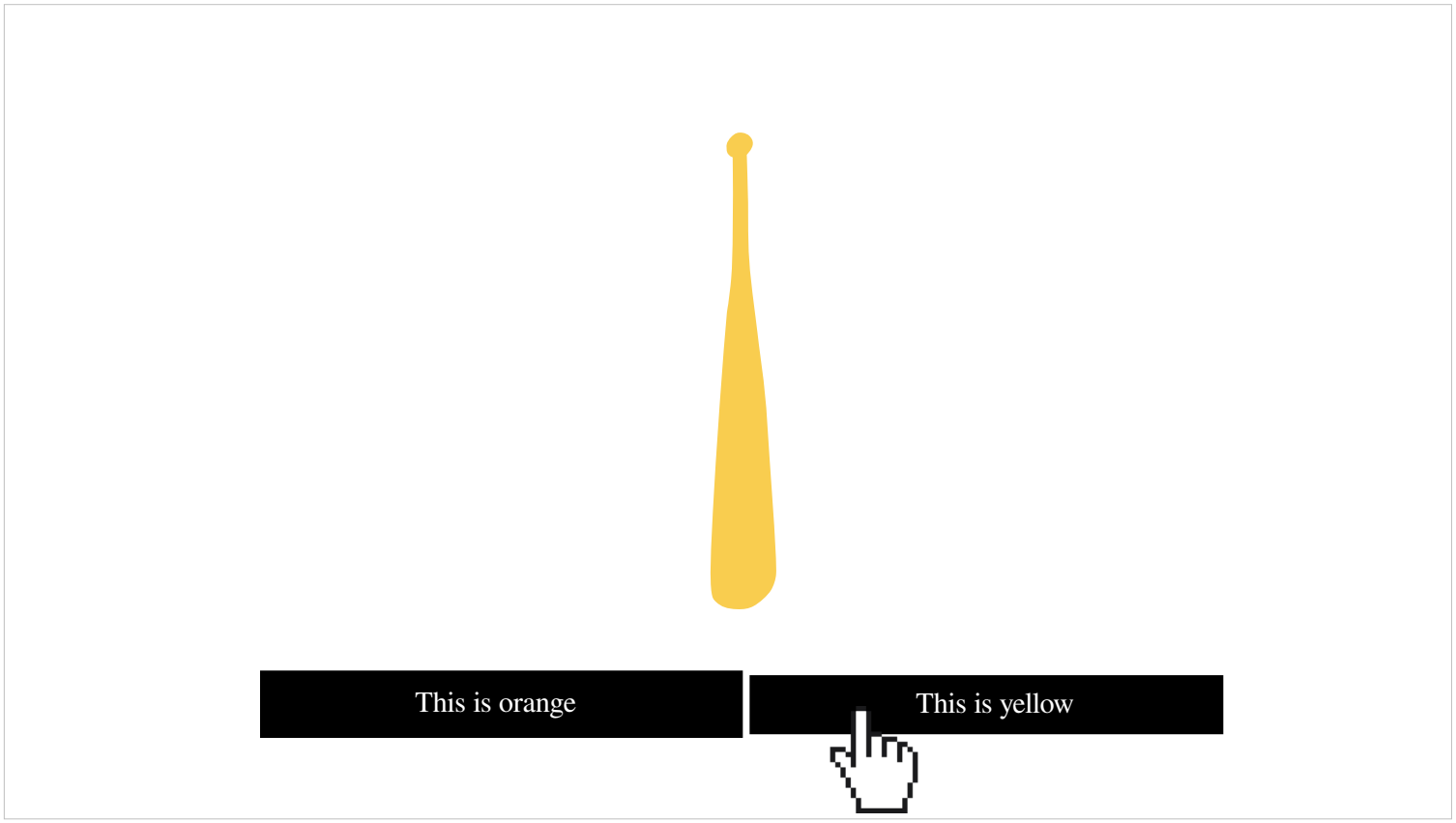
Orange – moving closer to yellow.



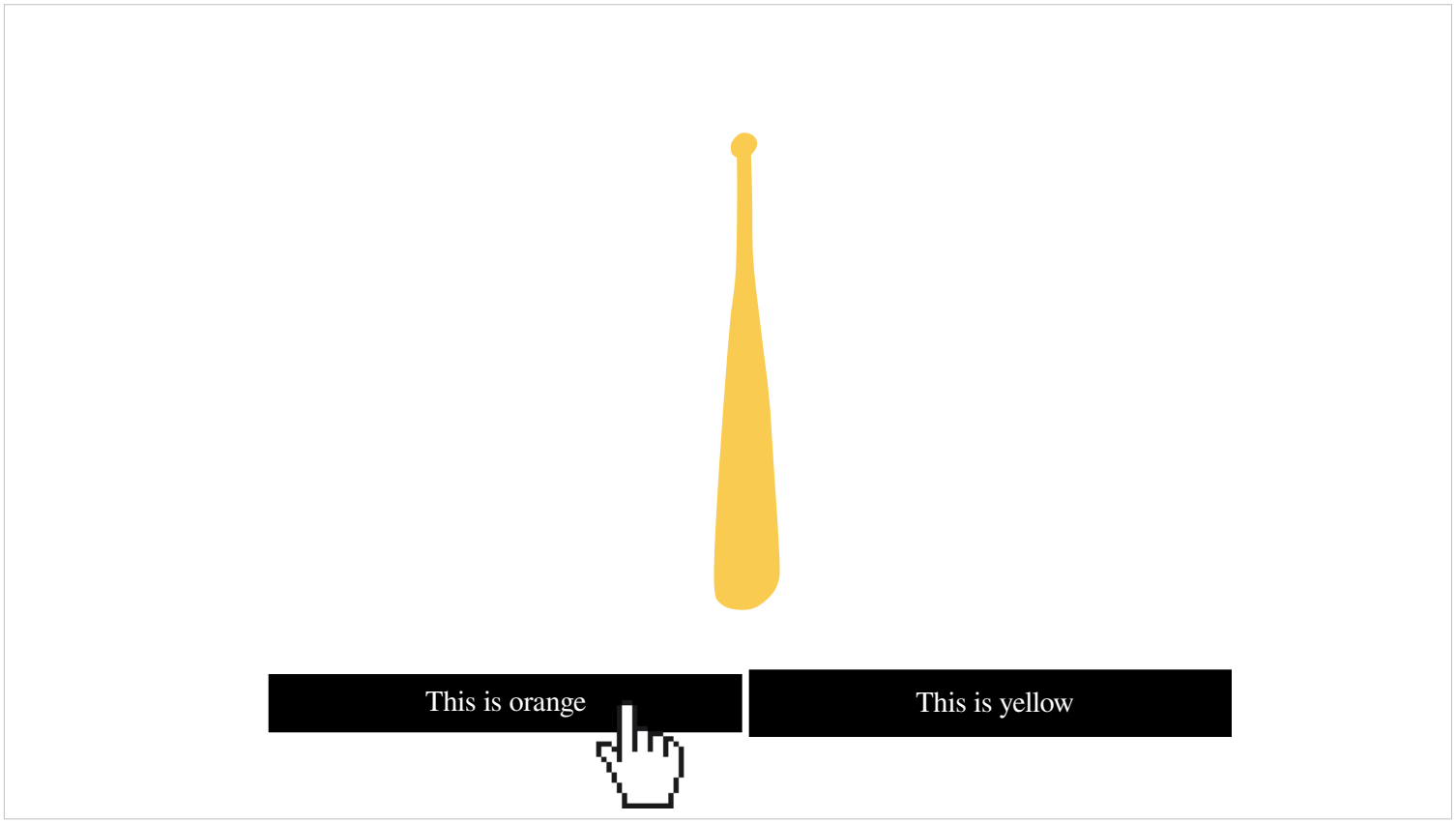
Yellow – moving closer to orange.



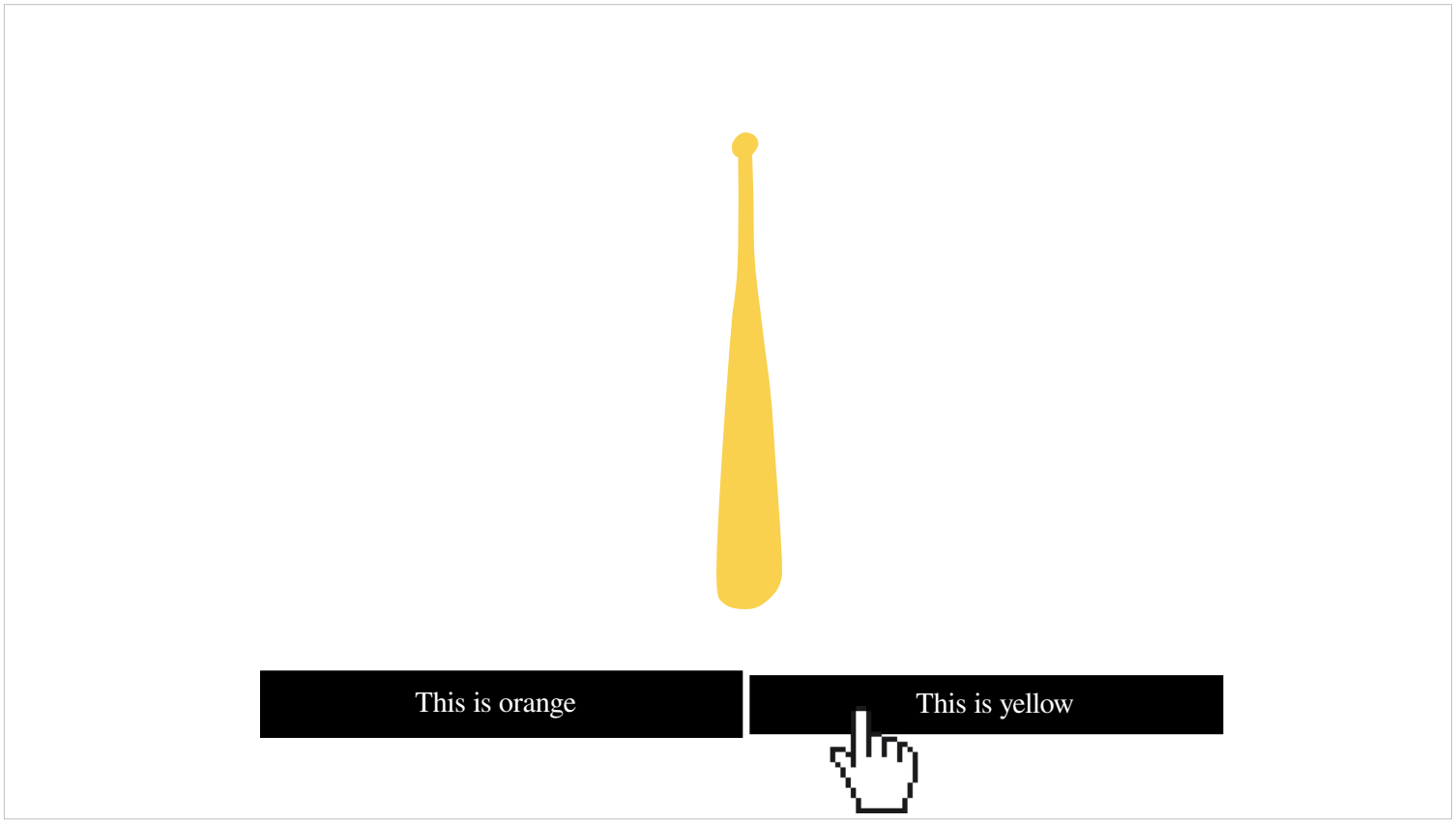
Juussst orange



Jussttst yellow

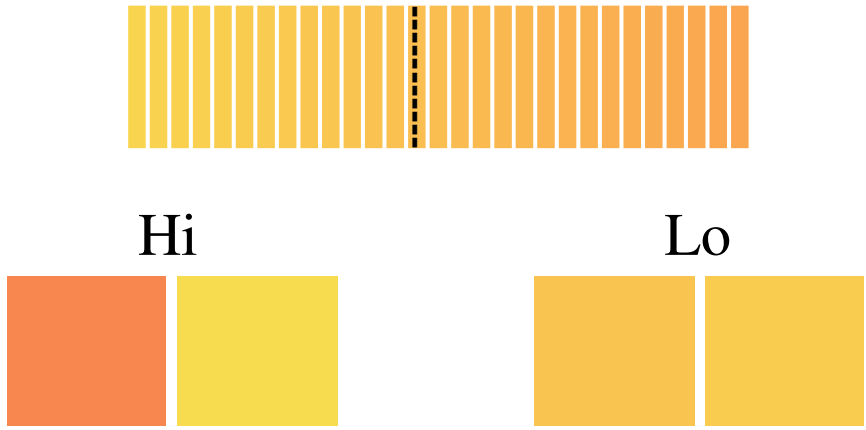


Juussst orange



And now we’re at yellow but that’s my limit.

What we end up with (colour)



*Your low-discriminability may differ from mine

And around that category boundary we can pull out our low-discriminability. And we take the highs from the far end for the director – thus, we end up with individualised points of high and low discrim that we can apply to our stimuli and use in our search efficiency manipulations.

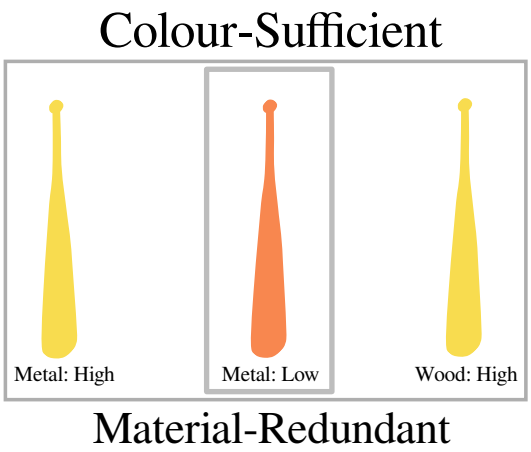
We run separate stairs for both colour and material

Does Search Efficiency drive reference for both colour and material?

…We must manipulate search efficiency!

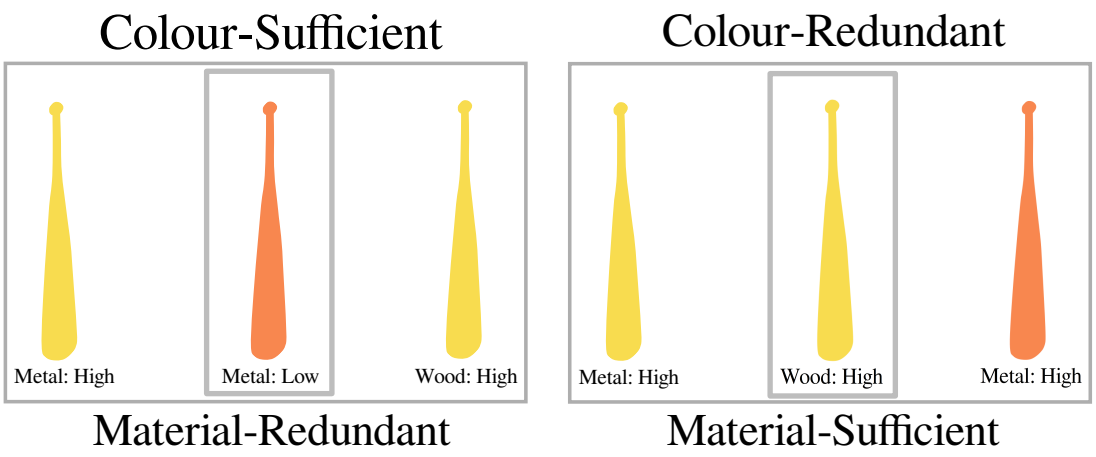
Cool, we got our stimuli and we're ready to go with the director task. Let's return to our question one more time: does search efficiency generalise? Does it drive reference for both colour AND material?

To do this we're going to take our stimuli and create an experimental manipulation that targets search efficiency.

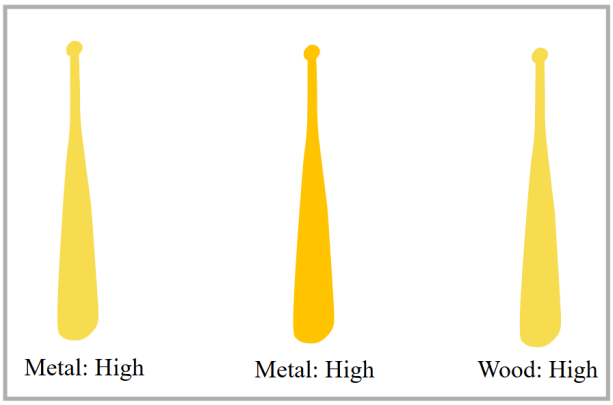


We manipulate search efficiency by playing with the distinction between SUFFICIENT and REDUNDANT.

In this case, colour is sufficient – and what we mean by colour being sufficient is that colour is all you need. Material is redundant.



In the converse, we have the colour-redundant, material sufficient condition. In this case, colour is redundant -- material is all you need.



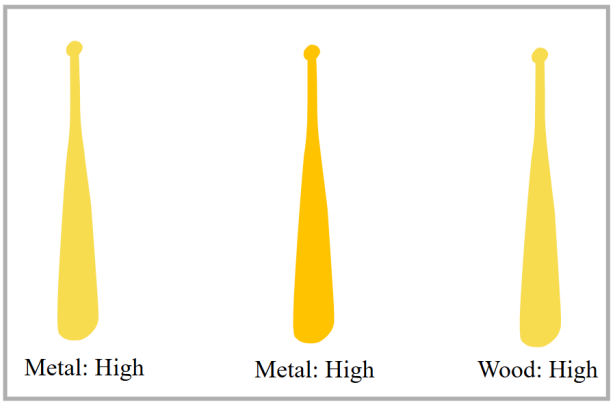
Sufficient Low, Redundant High

"The orange **metal** bat"

Difficult search → Overinform

And we play with sufficiency and redundancy to create conditions where redundancy is useful for search, and contrast these with conditions where it's NOT useful for search.

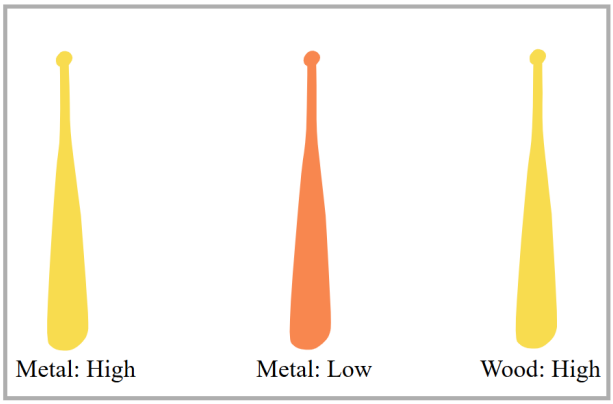
So in this display, the sufficient attribute makes search difficult. Search Efficiency says that people should overinform with the redundant, high-discriminability metal attribute to make search easier.



Sufficient Low, Redundant High

"The orange **metal** bat"

Difficult search → Overinform



Sufficient High, Redundant Low

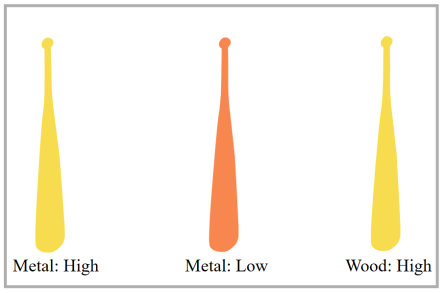
"The orange bat"

Easy search → do not overinform

And we compare this with the case where the sufficient attribute is high. In these cases the search for the sufficient attribute is easy, so search efficiency says speakers should just say what they need and leave it at that – don't overinform.

Sufficient High, Redundant Low

"The orange bat"

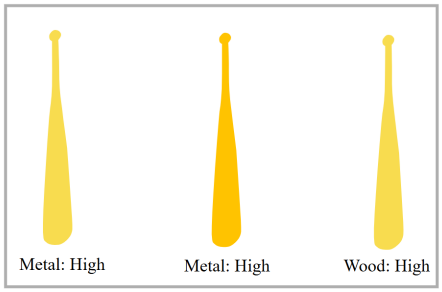


Easy search:
Do not overinform

→ → →

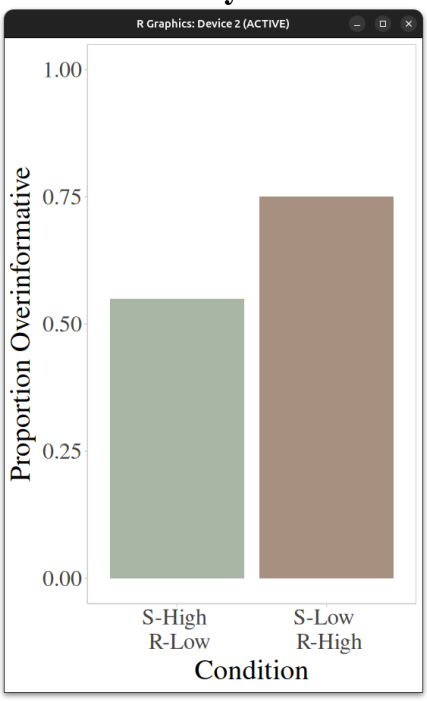
Sufficient Low, Redundant High

"The orange **metal** bat"



Difficult search:
Overinform

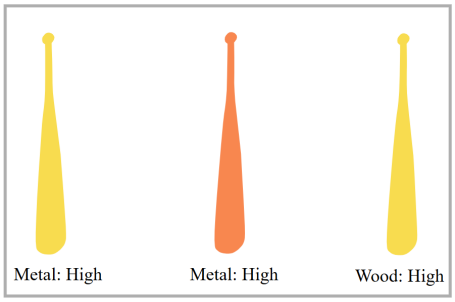
Search Efficiency Predictions



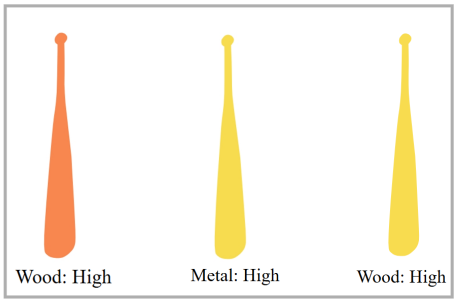
So translating this into prediction plots, we have the case where search using the sufficient attribute is easy on the green bar. Overinforming here should be low, because search is good enough as is without redundancy. And we contrast this with the brown bar in the plot, where search using the sufficient attribute is going to be difficult. So here people should overinform more often, grounding reference in the search efficient, high discriminability redundant attribute.

Everything we do for colour, we do for material.

Material Redundant

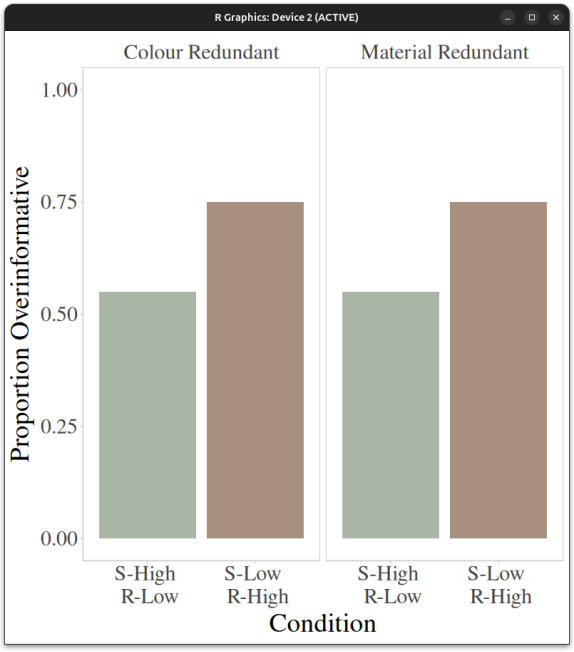


Colour Redundant



→ → →

Search Efficiency Predictions



Yet another reminder – material and colour are treated completely separate – and if the search efficiency view is general, we should see that the manipulations work the same way for colour and for material.

Recruit 120, hope for 80

Final N: 72

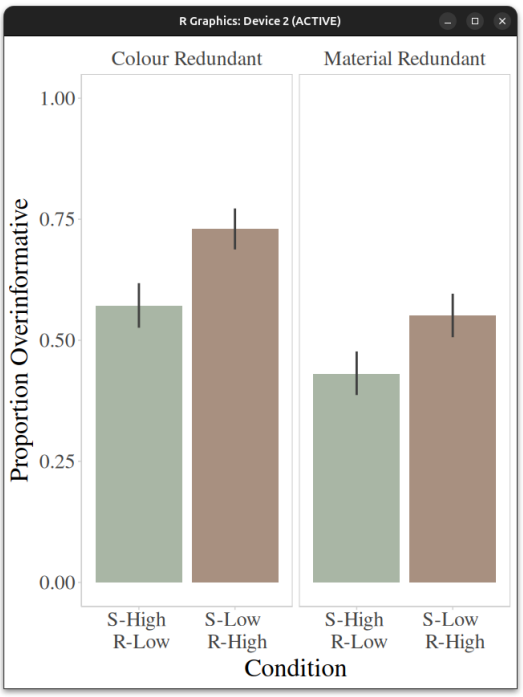
Crit Trials: 48 per person.

Final N crits: 2776

So we recruited 120 people which is about all we could in the budget for my honours project. But we knew from pilots that running the psychophysics tasks online was a real nightmare, so we aimed for 80. Everyone runs through 48 critical trials and after trial exclusions we end up with a little under 3000 critical trials.

Alright. By now I think by now we've built up enough excitement. Let's see what we got.

Whaddowegot?



There it is!

So the Search Efficiency predictions were exactly right. Speakers overinform more when it actually grounds reference in a high-discrim attribute (as shown in the brown bars). If the sufficient attribute is already high-discrim, they're more likely to ignore redundancy and just go with the standard reference.

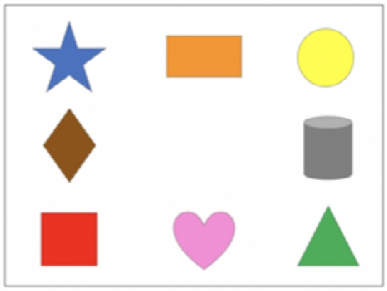
And there's no evidence for an interaction here. Search efficiency is working just as well for material as it is for colour.

A general theory of reference?

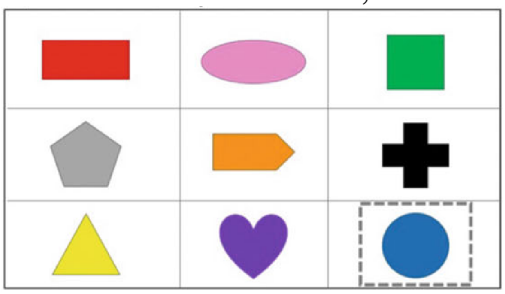
So let’s get back up to a high level and think about what’s going on.

A general theory of reference

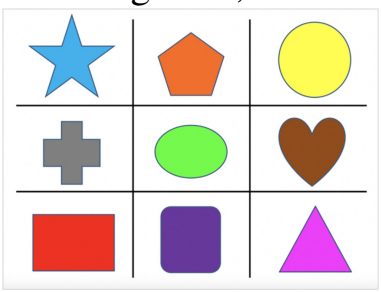
Rubio-Fernandez, 2019



Rubio-Fernandez, 2021



Long et al., 2021

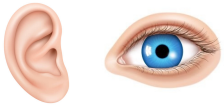


Kursat & Degen, 2021

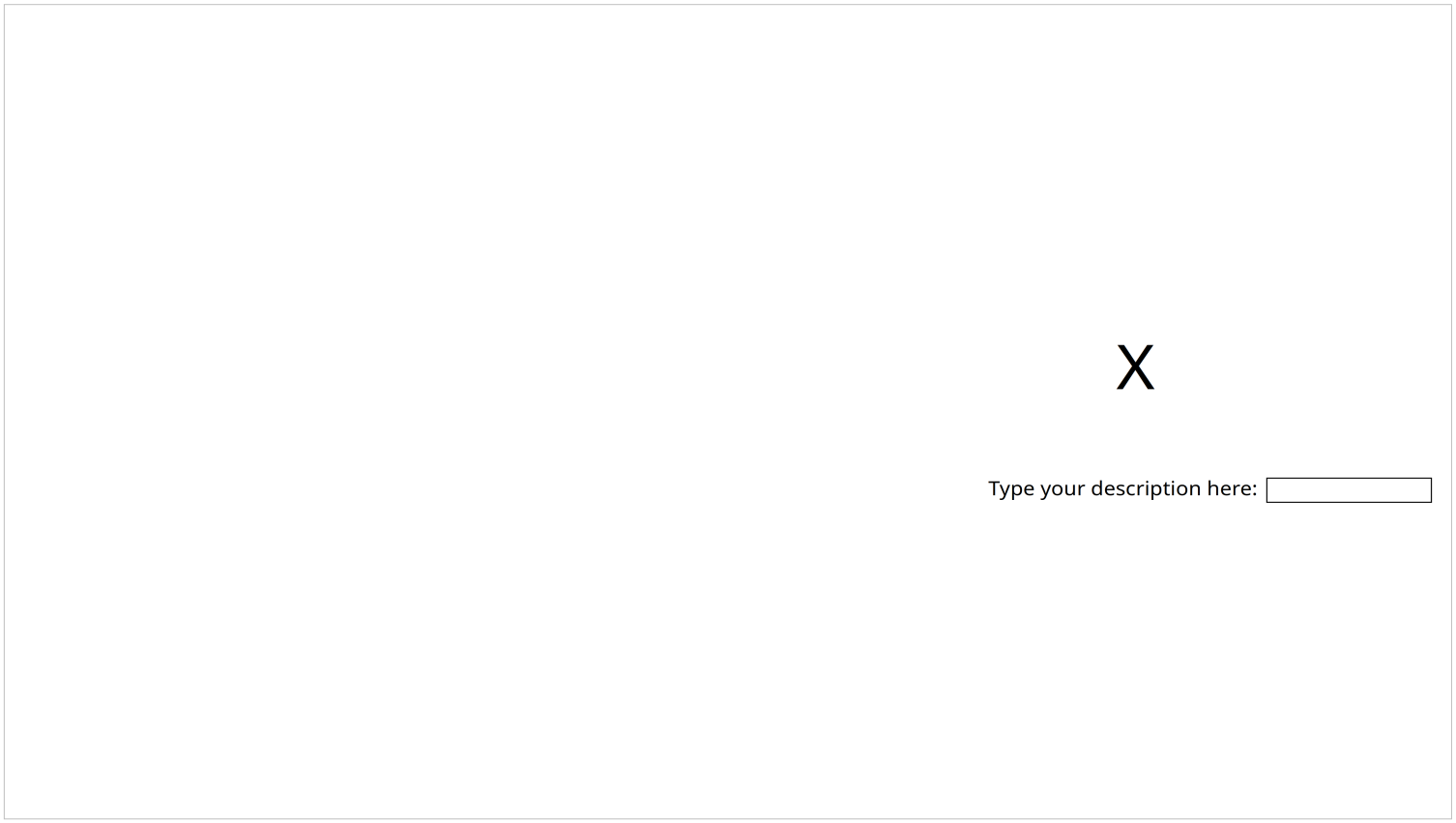
Jara-Ettinger & Rubio-Fernandez, 2022

So remember when I was motivating the study I pointed out that this Search Efficiency thing was really promising, but we just don't really have any evidence for it beyond colour.

Well, as it turns out, search efficiency generalise just fine – works well for material.



And it even works well across modalities to the auditory modalities – speakers even make auditory search more efficient.



And the principle of search efficiency holds even when speakers refer to items from memory – nothing was perceptually available when our participants were referring. Everything was done from memory.

A general theory of reference 🎉

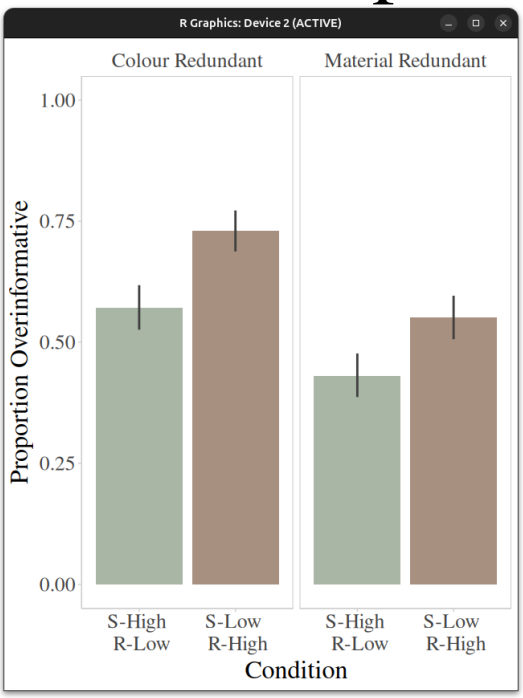
👂👁

X

Type your description here:

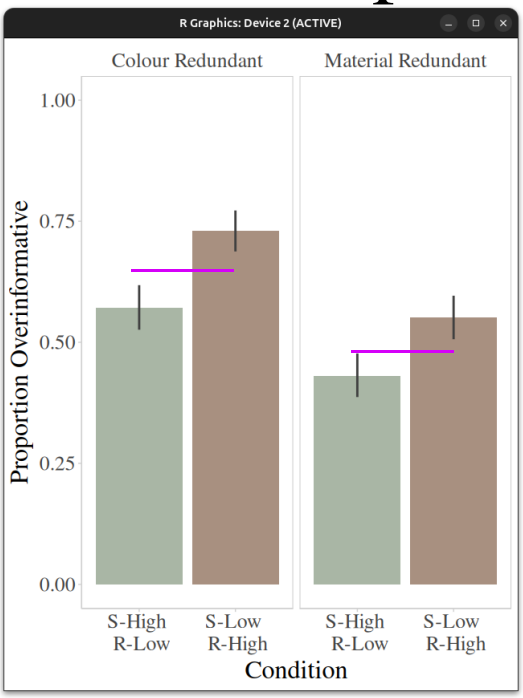
So cool. At this point there's a pretty stong case to be made for search efficiency as a general driver of reference.

But is colour special…?



But – you don’t have to look very closely to notice something about the results.

But is colour special…?



Speakers are referring to redundant colours much more than materials

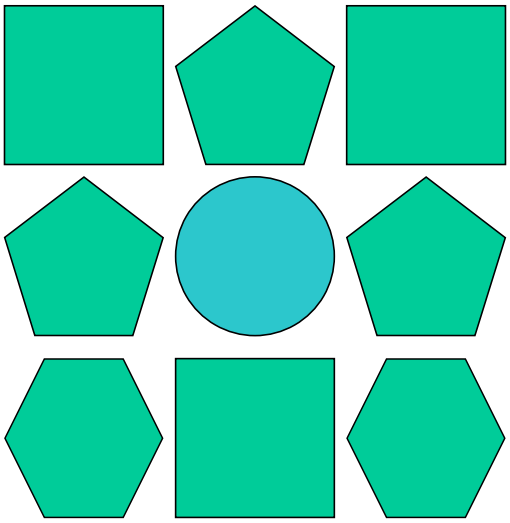
But is colour special…?

Foreshadowing Intensifies

Now of course, experiment one can’t quite settle this.

It might be the case that speakers really like grounding reference in *visual* attributes, and are reluctant to use auditory properties.

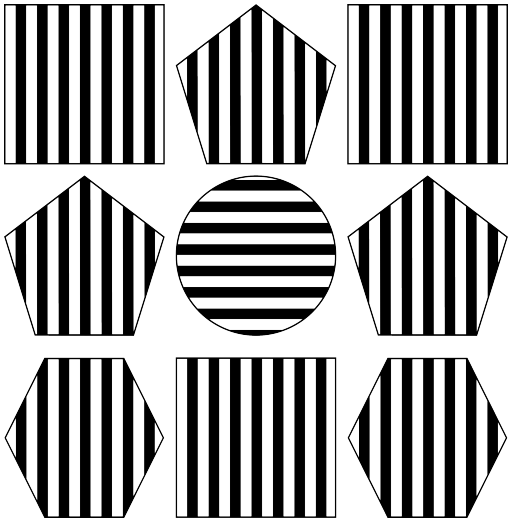
Colour captures visual attention in a special way



Wolfe & Horowitz, 2017 for a wonderful review of attention-guiding attributes

Visual search researchers will say that colour is actually kinda special – it captures visual attention and makes search really fast. Material doesn't.

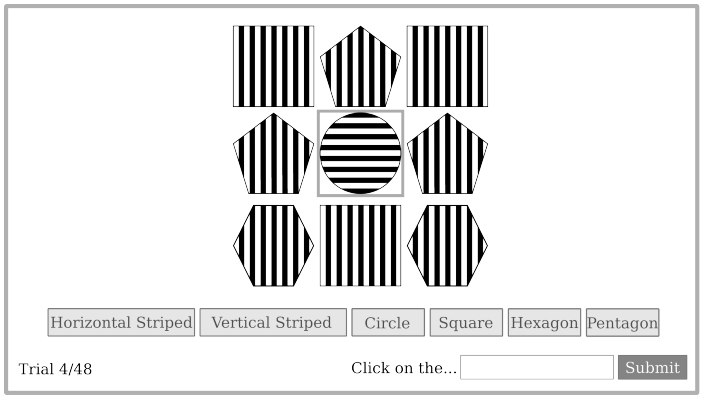
Orientation also captures visual attention in a special way

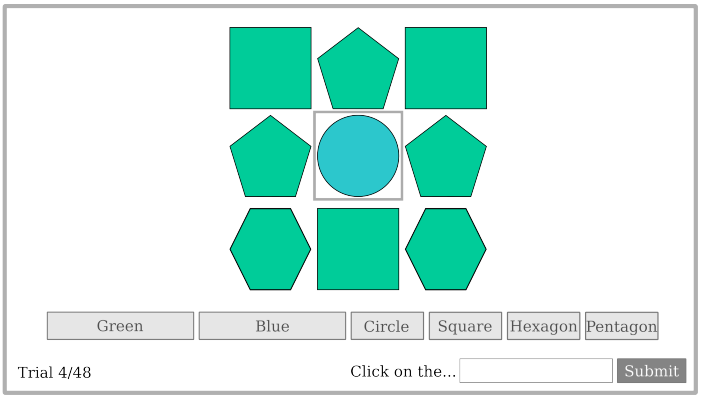


Wolfe & Horowitz, 2017 for a wonderful review of attention-guiding attributes

But colour isn't the only one. Orientation is also an attentional guidance attribute.

Compare



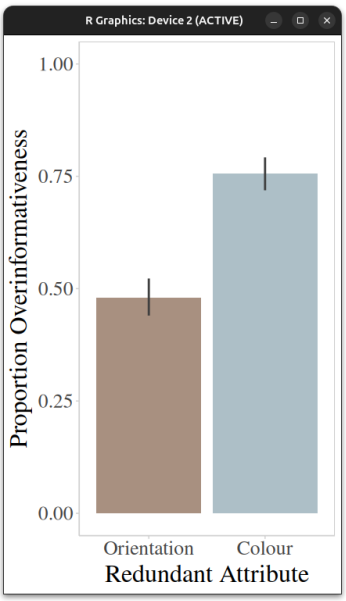


So what we did is ran a follow up experiment comparing colour and orientation in a director task, where you use buttons to construct expressions – equalising production cost.

Colour is Special.

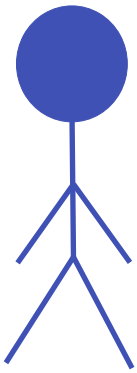
Between-Subjects

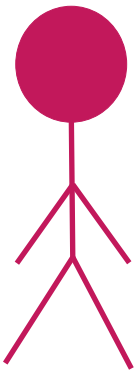
$N = 97$



Here are the results, and sure enough. Colour is special. It's way higher than orientation – so attention capturing isn't it. The proportion of colour overinforming in general is way high, so salience isn't it, and production effort and word frequency made no difference.

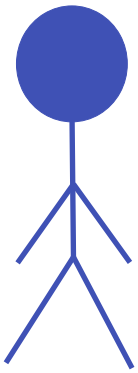
So what's going on? Why is colour special?



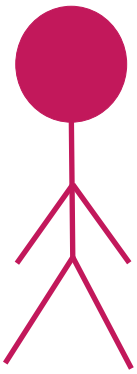


Well, I'll outline a general picture of what might be happening when people refer, and how this could give rise to a colour privilege.

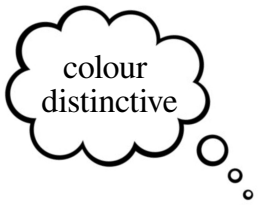
Imagine you (blue) find yourself needing to point something out using words for the first time ever.

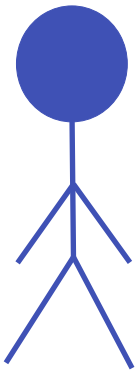


Strategy Choice	Score
The **[object]**	0
The **[colour] [object]**	0
The **[material] [object]**	0

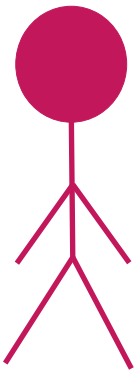


You're presented with a problem: they have a set of possible strategies and have to choose a successful one.

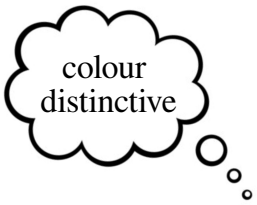


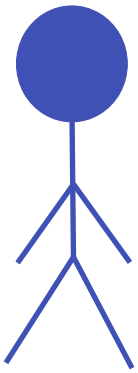


Strategy Choice	Score
The **[object]**	0
The **[colour] [object]**	0
The **[material] [object]**	0

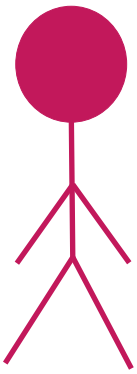


You use your eyes and notice that colour is really distinctive of the thing they want to point out.

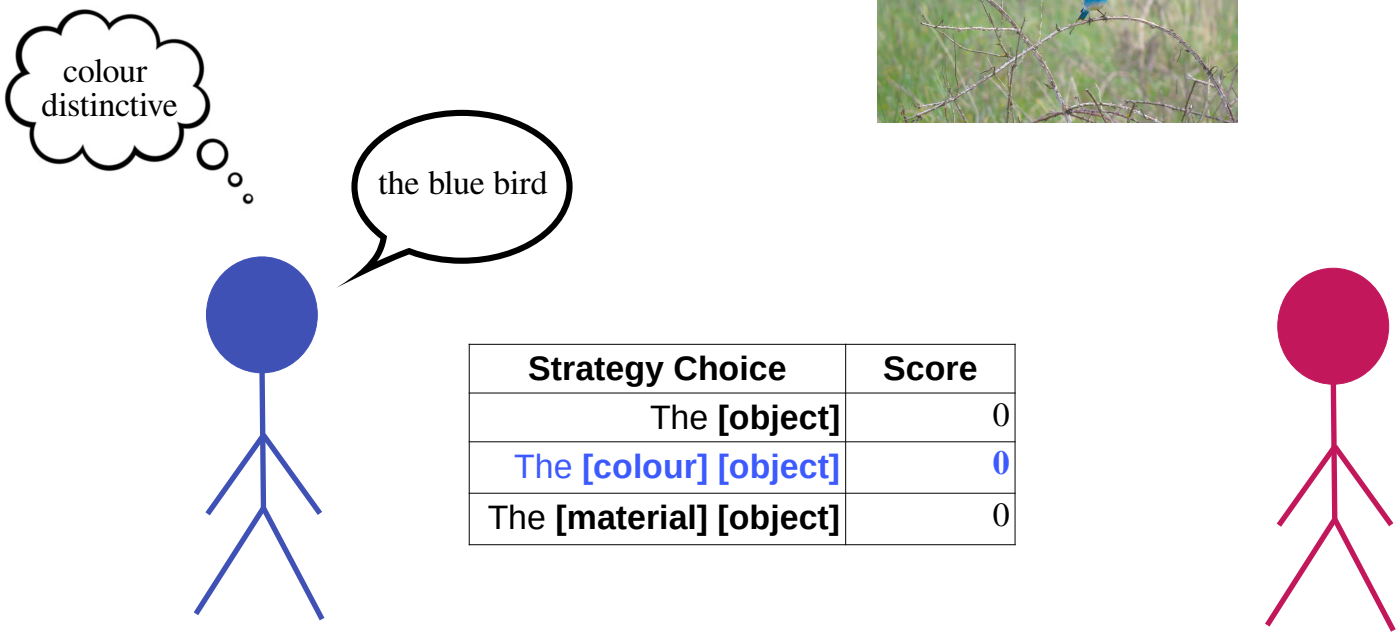




Strategy Choice	Score
The **[object]**	0
The **[colour] [object]**	**0**
The **[material] [object]**	0

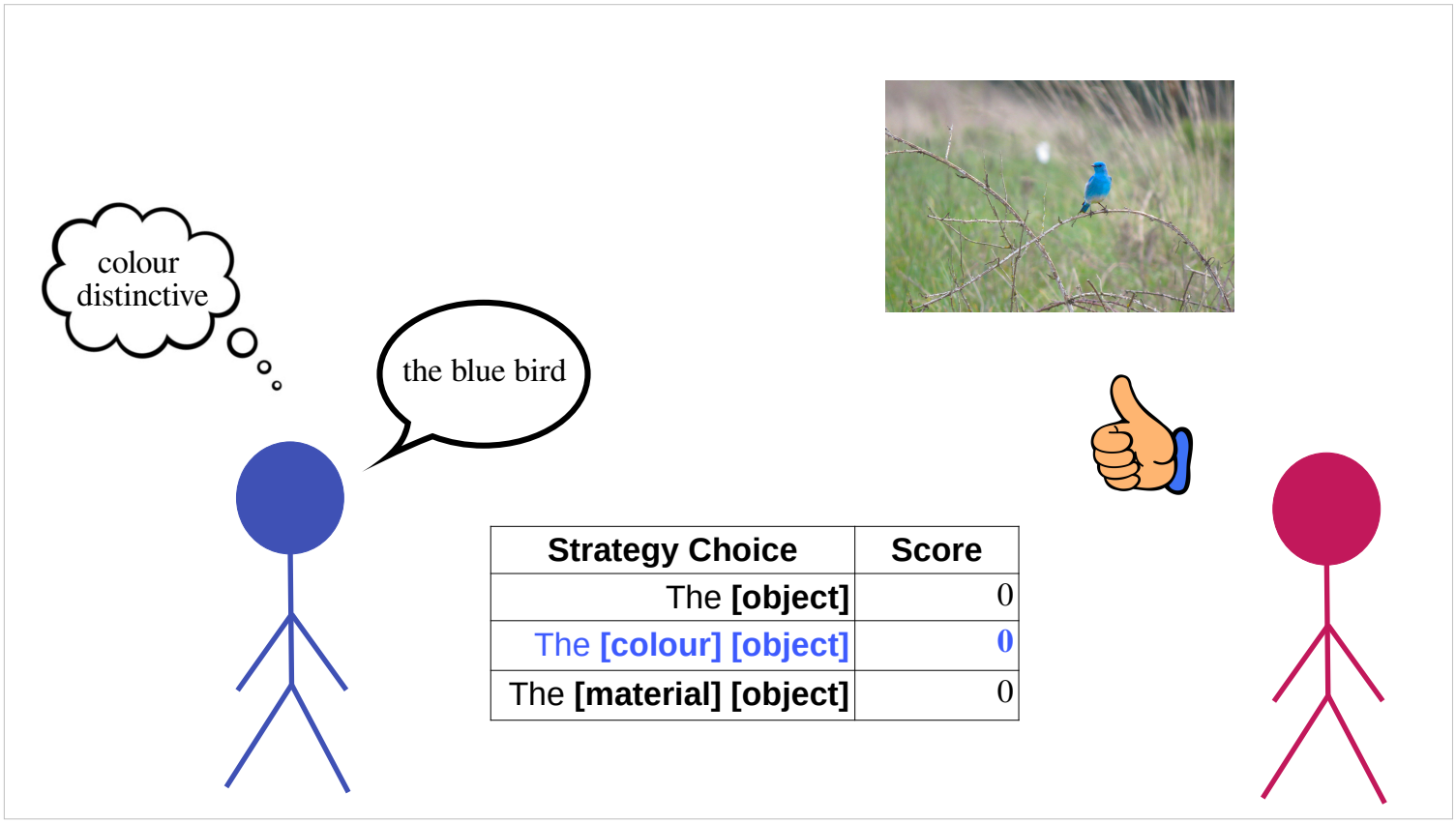


So you use it select the overinformative colour strategy.

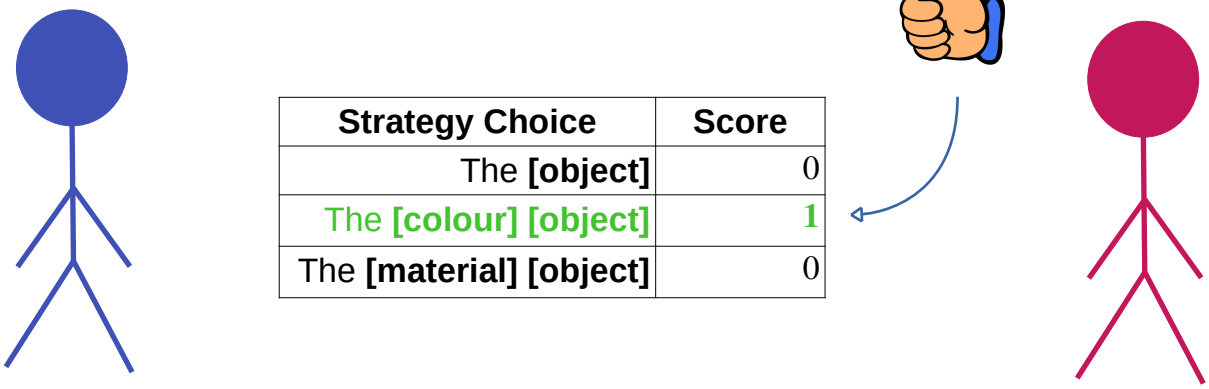


Strategy Choice	Score
The **[object]**	0
The **[colour] [object]**	**0**
The **[material] [object]**	0

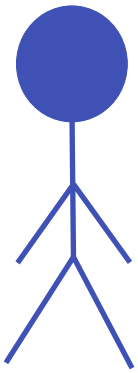
they run with it and produce the overinformative colour expression.



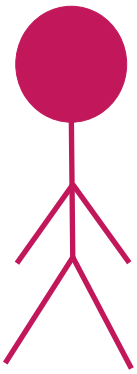
And it turns out that worked really well. They found the thing, and you receive this feedback



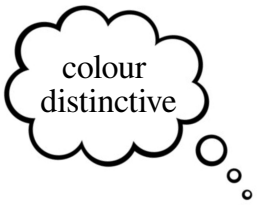
And you use the feedback to tune your strategy choices.

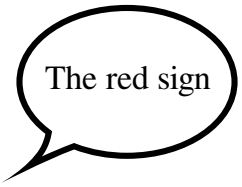


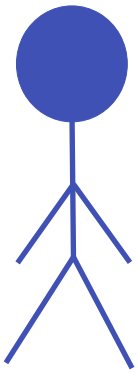
Strategy Choice	Score
The **[colour] [object]**	**1**
The **[object]**	0
The **[material] [object]**	0



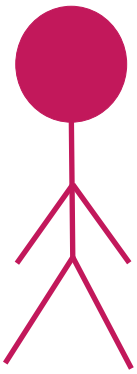
So you find yourself in another environment, and
The successful strategy moves to the top of the pile –
it's more readily available.



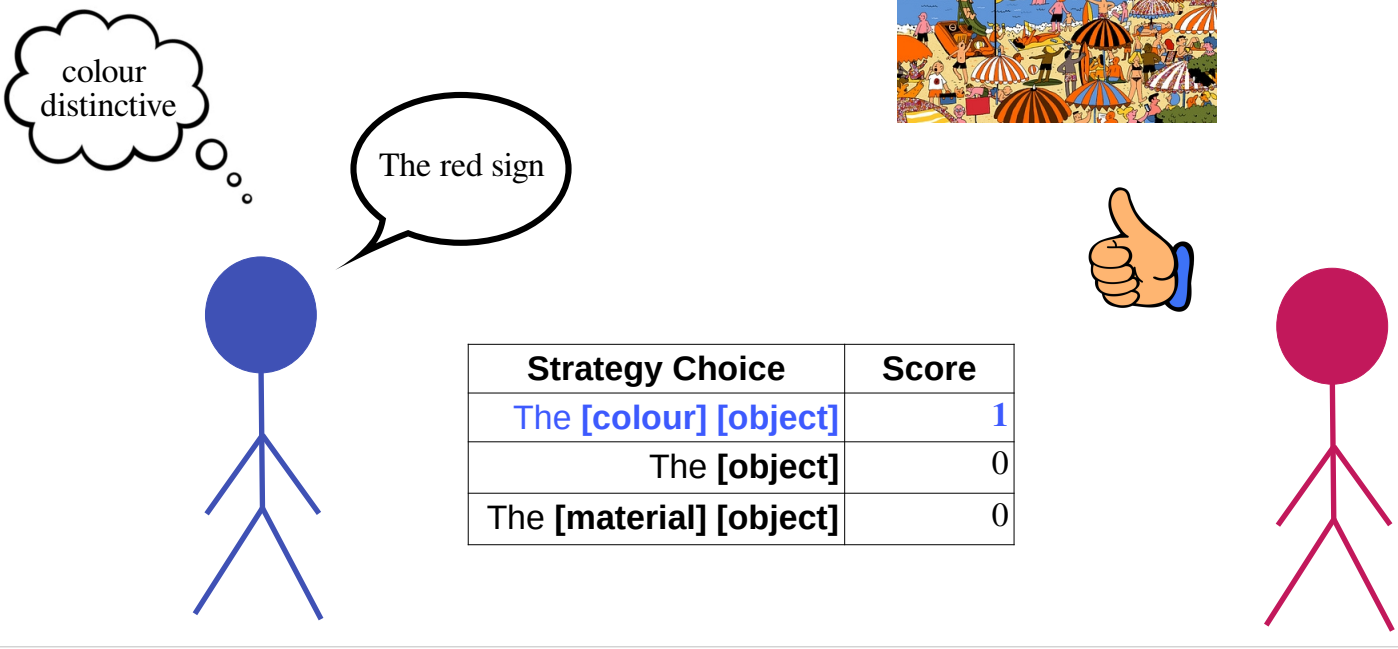




Strategy Choice	Score
The **[colour] [object]**	**1**
The **[object]**	0
The **[material] [object]**	0

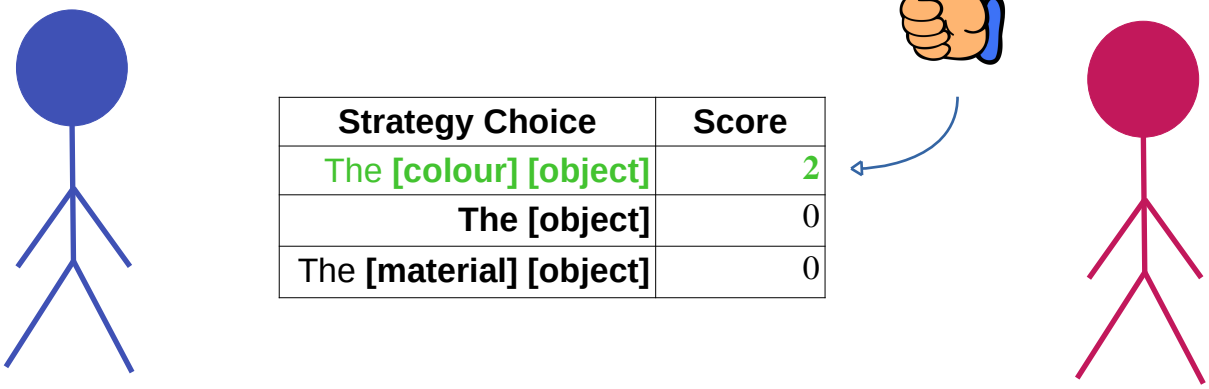


So you select it and use it again.



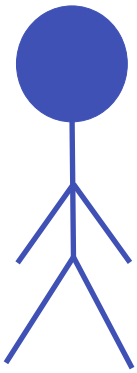
Strategy Choice	Score
The **[colour] [object]**	**1**
The **[object]**	0
The **[material] [object]**	0

And it works again.

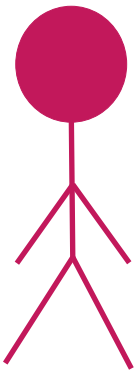


Strategy Choice	Score
The **[colour] [object]**	2
The [object]	0
The **[material] [object]**	0

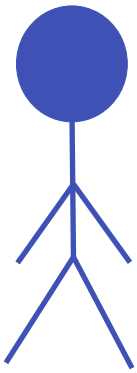
So the strategy is reinforced again



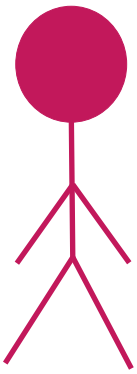
Strategy Choice	Score
The **[colour] [object]**	3
The [object]	0
The **[material] [object]**	0



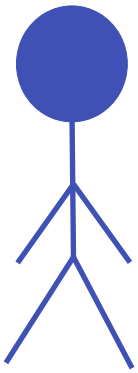
And imagine this keeps happening



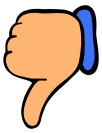
Strategy Choice	Score
The **[colour] [object]**	**4**
The [object]	0
The **[material] [object]**	0

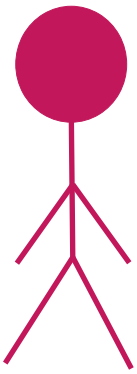


You get wins

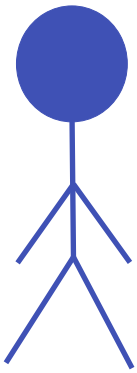


Strategy Choice	Score
The **[colour] [object]**	**5**
The [object]	**-1**
The **[material] [object]**	0

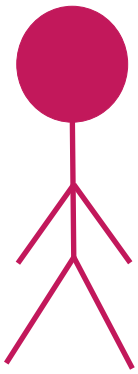




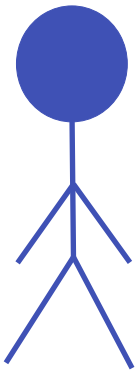
And you get losses



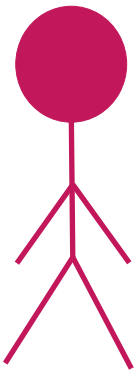
Strategy Choice	Score
The **[colour] [object]**	**10**
The [object]	4
The **[material] [object]**	2



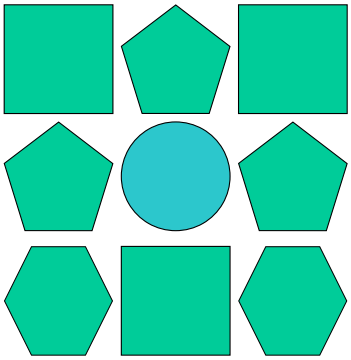
And you start to get a cross situational impression of effective strategies, and this guides your strategy selection

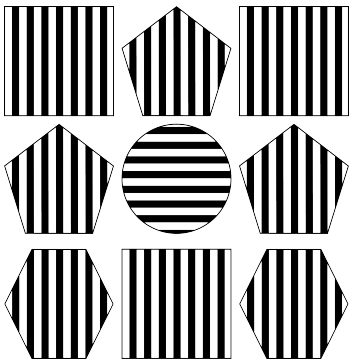


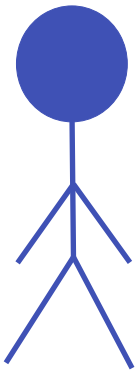
Strategy Choice	Score
The **[colour] [object]**	90
The [object]	40
The **[material] [object]**	30



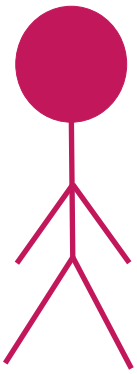
So learning over time which strategies are quick and reliable you build up these really strong habits of reference.





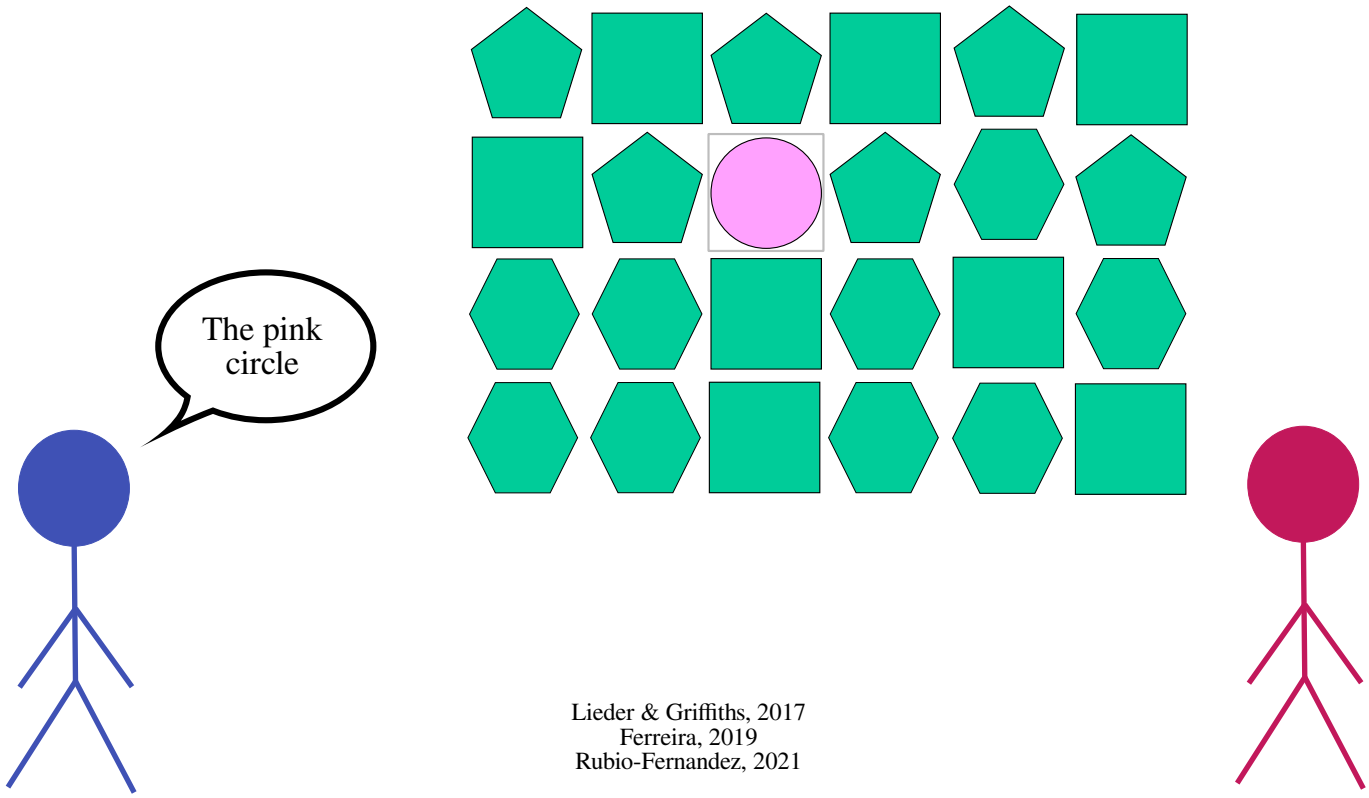


Strategy Choice	Score
The **[colour] [object]**	90
The [object]	40
The **[material] [object]**	30



Then you get put in our experiment. But you've already built up a lifes worth of habits guided by which strategies are quick and effective.

So even when we equalise colour and orientation in terms of search – most people don't just ignore this prior evidence in favour of certain attributes.



And we can still explain strategies like grounding reference in distinctive attributes or providing more information in dense displays. You have strategies that use general perceptual cues as well as certain properties.

Search Efficiency is promising as a general theory of reference

Colour is special.

We should incorporate prior experience in strategy selection.

Okay! Here are our key takeaways. We've shown search efficiency is promising – it extends across attributes, modalities, and to reference from memory – but search efficiency in the immediate scene can't quite capture the specialness of colour. We interpret this as a hint that theories of reference strategy should further incorporate prior experience.

Thank you!

Francis Mollica

University of Melbourne
Verified email at unimelb.edu.au - Homepage

Paula Rubio-Fernandez

Visiting researcher, MIT
Verified email at ifikk.uio.no - Homepage
Experimental pragmatics Theory of Mind reference

Thank you to James Traer for sharing the material impact recordings, and sharing the generative model of impact sounds.

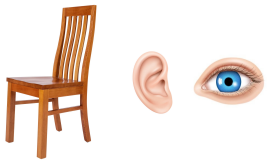
Thank you to Paul Garrett and Andy Perfors for valuable advice on interpreting the experiments and on the design of experiment two.

Thank you all for listening.

Thanks to collaborators and all the people that helped us out

Thank you!

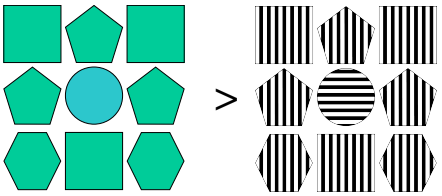
Search Efficiency is promising as a general theory of reference



X

Type your description here:

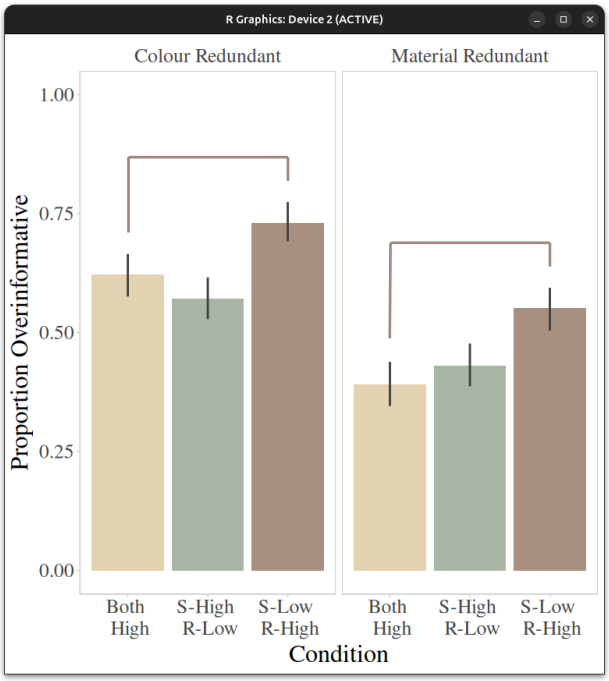
Colour is special.



There may be a role for prior experience in referential strategy selection.

PsyArxiv preprint for the full paper (Exp 2 is *not* in the proceedings)

Everything high

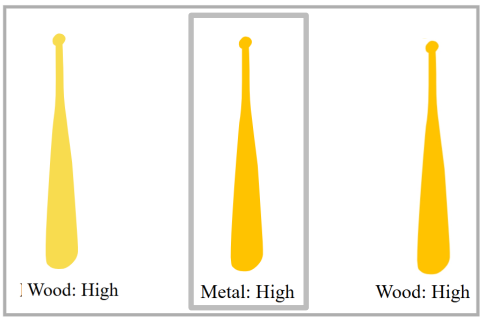


Supplementary Slides

We also ran a condition where everything was high – so both the sufficient AND the redundant attribute were high-discrim. Still, overinformativeness was lower than in the condition where the sufficient attribute was low. So this means that participants are more likely to ignore the redundant dimension when the sufficient attribute is good enough for search.

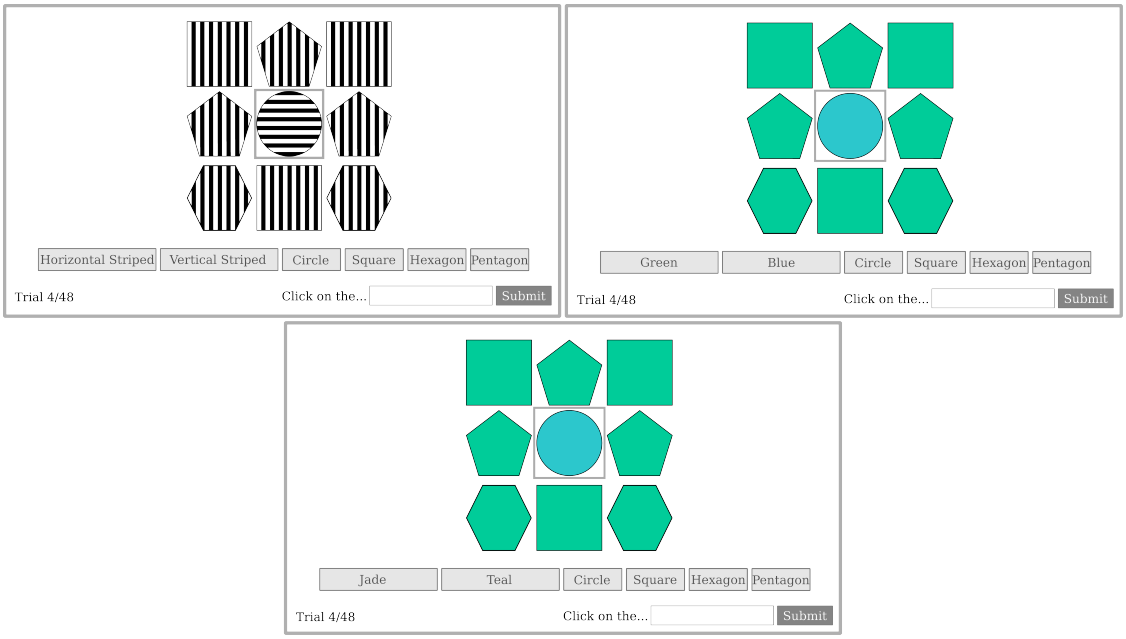
It also means that participants aren't just blurting out whatever is high discriminability in a scene – otherwise in this condition they'd mention both sufficient and redundant attributes – it means that speakers are sensitive to heightened listener search costs on some level. Maybe they follow the strategy of mentioning at least one search-efficient attribute.

Could colour dominance in E1 be due to search efficiency?



So the colour is special result holds in E1 even when we focus on cases where search efficiency is completely removed, like in the case on the left. Speakers *could* just refer to metal and ignore the colour attribute, but overinforming with colour remains higher than material.

Controlling production effort and word frequency.



Controlling production effort and word frequency.