

Visual memory and numerosity estimation: some data and some ideas

Declan Devlin, Eivind Kaspersen, Trygve Solstad

Norwegian University of Science
and Technology

28 July 2026

Overview of presentation

1

**What we are trying
to explain**

2

**Visual memory /
comparison
explanations**

3

**New data: can
people remember
dot patterns?**

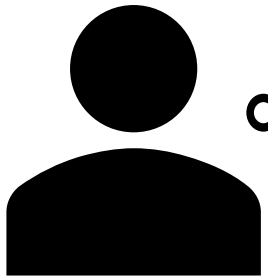
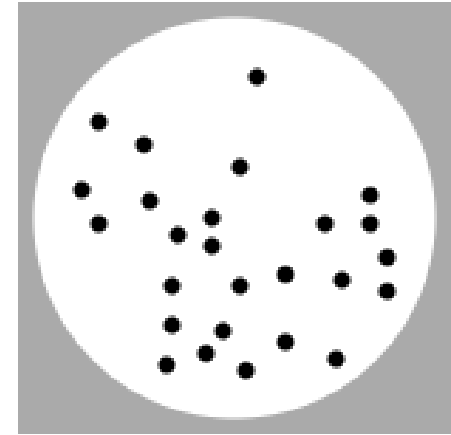
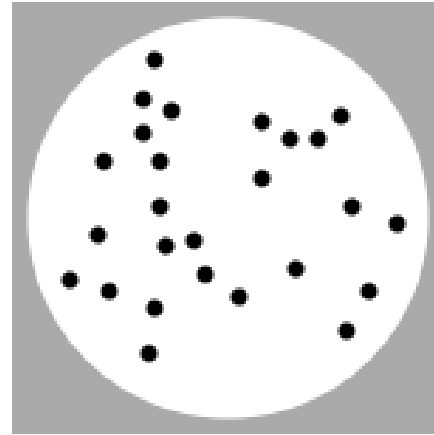
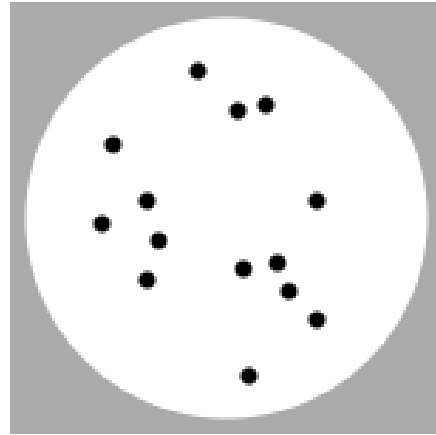
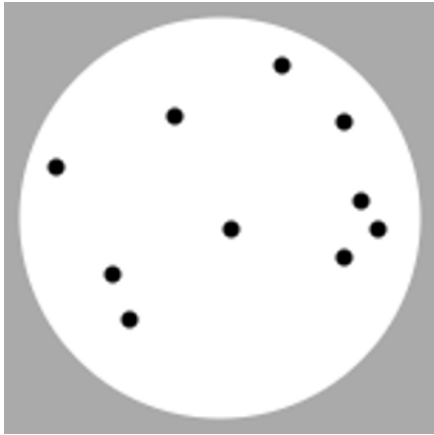
4

Questions, implications, directions for future research

1. What are we trying to explain?

What are we trying to explain?

How many dots?



We want to identify the cognitive strategies and mechanisms underlying quantity estimation...

But it is not so simple...

Preprint

Non-symbolic quantity estimation may be characterised by many signatures, yet we typically focus on only a few

Declan Devlin¹ and Natalia Dubinkina²

¹Department of Teacher Education, Faculty of Social and Educational Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway*

²Department of Communication Studies, University of Antwerp, Belgium

ABSTRACT

When a person estimates non-symbolic quantities, they typically do so in a way that produces a characteristic set of behavioural patterns. For example, small quantities may be estimated with near-perfect accuracy, whereas estimates of larger quantities become increasingly variable. Consequently, attempts to identify the mechanisms underlying quantity estimation involve explaining not only how people are able to estimate quantities, but also how they do so in a way that gives rise to these recurring behavioural phenomena. Over the past several decades, an increasing number of candidate “signatures” of quantity estimation have been identified, including effects relating to visual cues, previous estimates, and individual differences. Despite this, theoretical accounts often differ both in which phenomena they attempt to explain and in how they prioritise even shared behavioural signatures. In this sense, competing perspectives may disagree not only on what mechanisms underlie quantity estimation, but also on what aspects of estimation behaviour we should even be trying to explain. In this review, we examine how different theoretical approaches select and prioritise behavioural signatures, discuss conceptual questions concerning what should count as a “signature” of quantity estimation, and consider the implications of these issues for how competing models are developed and evaluated. In particular, we propose treating candidate signatures as a provisional “dataset to be explained” against which competing models may be compared within a shared explanatory framework. Although we do not claim to resolve these issues definitively, we hope that articulating them explicitly will encourage greater transparency regarding explanatory assumptions and contribute towards a clearer and more cumulative understanding of non-symbolic quantity estimation and its underlying mechanisms.

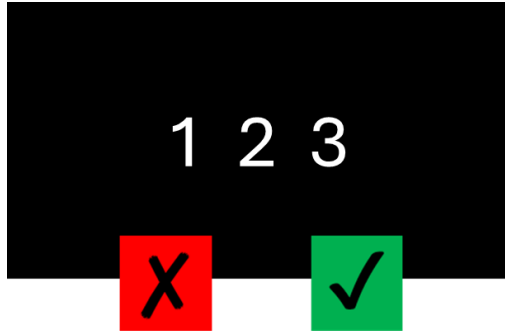
Keywords: quantity estimation, numerosity, behavioural signatures, numerical cognition

We do not simply try to explain *how* people perform a skill...

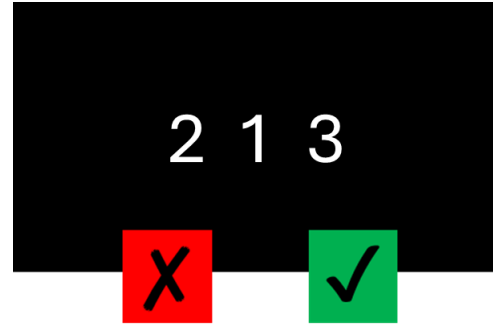
But rather how they do so in a particular way so as to produce a specific set of behavioural phenomena...

Which phenomena we focus on explaining influences our explanations...

An example from order verification



In order 😊



Not in order ☹️

Order effect

234 vs 243

Direction effect

234 vs 432

Regularity effect

135 vs 136

Single-digit effect

7-8-9 vs 17-18-19

Reverse distance effect

123 vs 135

Which effects we choose to focus on really matters....

The reverse distance effect

We only really care about this one!

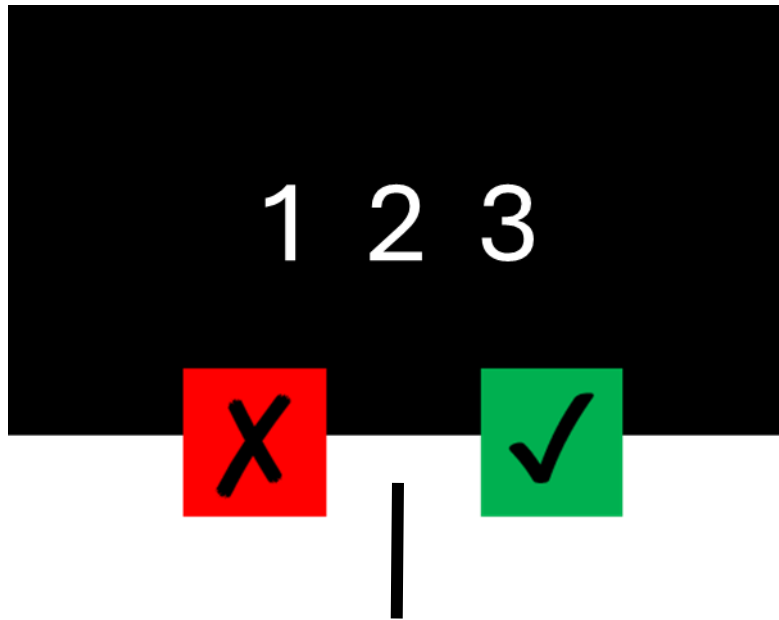
Reverse distance effect

123 vs 135

*The **RDE** has often been described as a “hallmark” or order processing (e.g., Lyons & Ansari, 2015)*

*...and some going as far as to suggest that if no **RDE** is observed, then “order processing” is not really occurring (i.e., Wong et al., 2021)*

Implications of the RDE



“Order processing”

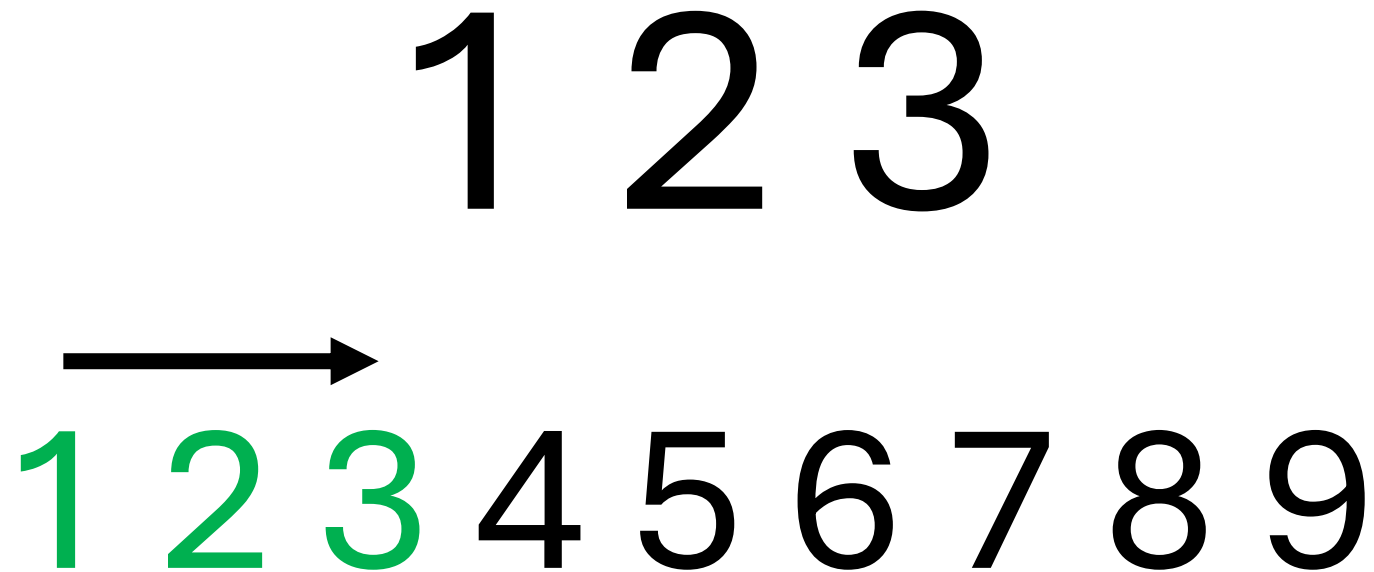
Reverse distance effect

123 vs 135

Explaining the RDE will tell us something about this latent construct of ordinality

Serial-scanning mechanism

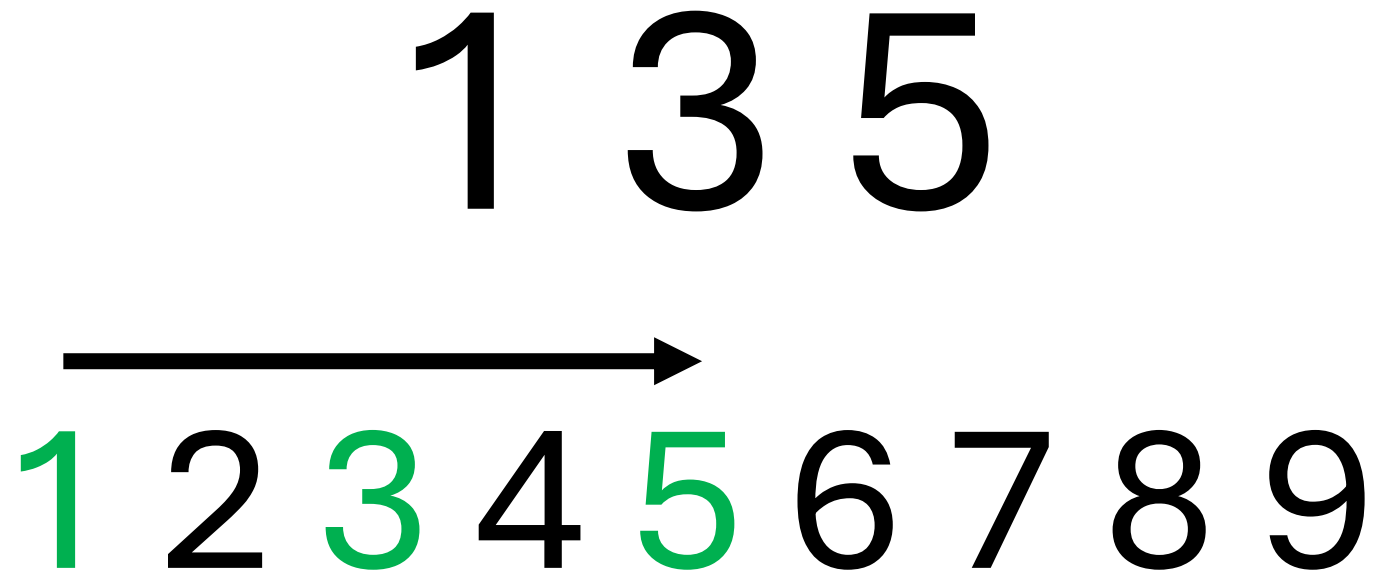
Prioritising the RDE naturally encourages distance-based accounts of order processing, such as a proposed **serial-scanning mechanism**:



(e.g., Franklin et al., 2009; Turconi et al., 2004, 2006)

Serial-scanning mechanism

Prioritising the RDE naturally encourages distance-based accounts of order processing, such as a proposed **serial-scanning mechanism**:



(e.g., Franklin et al., 2009; Turconi et al., 2004, 2006)

OK but...

How does this explain all those other effects?



Order effect

234 vs 243

Reverse distance effect

123 vs 135

Direction effect

234 vs 432

Regularity effect

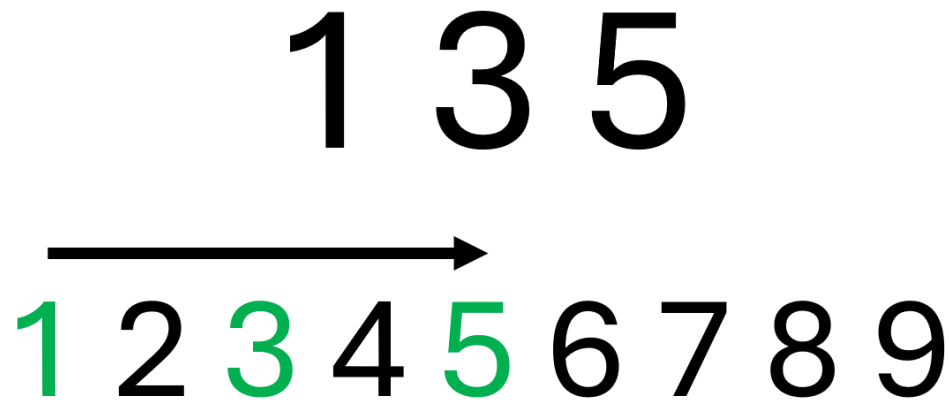
135 vs 136

Single-digit effect

7-8-9 vs 17-18-19

Is the RDE actually a RDE?

The term “**reverse distance effect**” implies a **linear relationship** whereby RTs increase as distance between digits increases



*Serial scanning mechanism can
account for linear increases in RTs*

Actually a consecutiveness effect?

The “**reverse distance effect**” may be better thought of as a “**consecutiveness effect**”

1 2 3

Distance 1

1 3 5

Distance 2

1 4 7

Distance 3

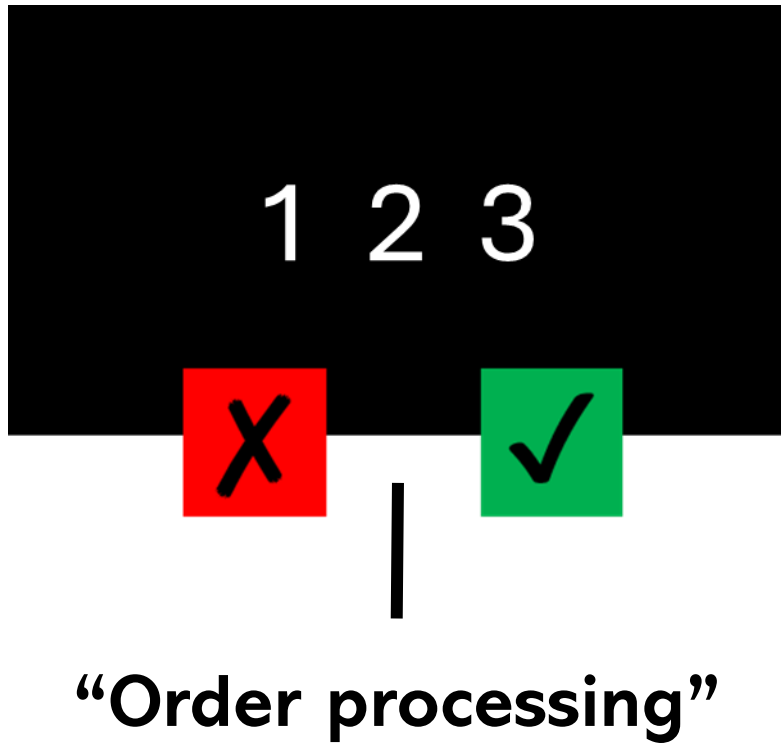
Faster processing appears **specific to distances of 1**, as greater distances produce either no distance effect or even standard distance effects (e.g., Turconi et al., 2006; Vogel et al., 2017).

~~Reverse distance effect~~



Consecutiveness effect

Implications of a consecutiveness effect



Consecutiveness effect

123 vs 135

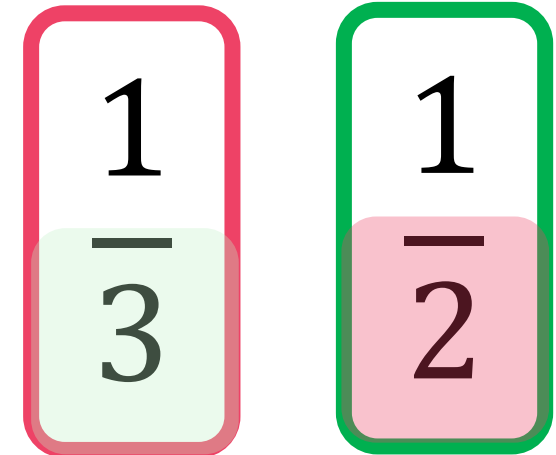
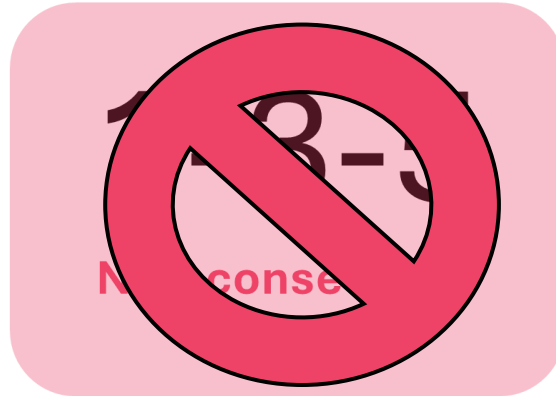
Explaining the CE will tell us something about this latent construct of ordinality

A very different mechanism...

Most studies now operationalise the RDE as a consecutiveness effect,
which has recently led to count-list-based theories:

Early experiences with order come from learning the verbal count-list
“one, two, three, four, five...”

Leading to an intuition that “in order” == “matches the count-list”



(e.g., Gattas et al., 2021; Hutchison et al., 2022)

OK, but...

How does this explain all those other effects?



Order effect

234 vs 243

Reverse distance effect

123 vs 135

Direction effect

234 vs 432

Regularity effect

135 vs 136

Single-digit effect

7-8-9 vs 17-18-19

Is it actually a consecutiveness effect?

~~“Small distance sequences are processed faster than large distance sequences”~~

~~“Consecutive sequences are processed faster than non consecutive sequences”~~

“Familiar sequences are processed faster than unfamiliar sequences”

The familiarity effect

Consecutive

1 2 3 4 5 6 5 6 7

2 3 4 3 4 5

6 7 8 7 8 9

2 4 6

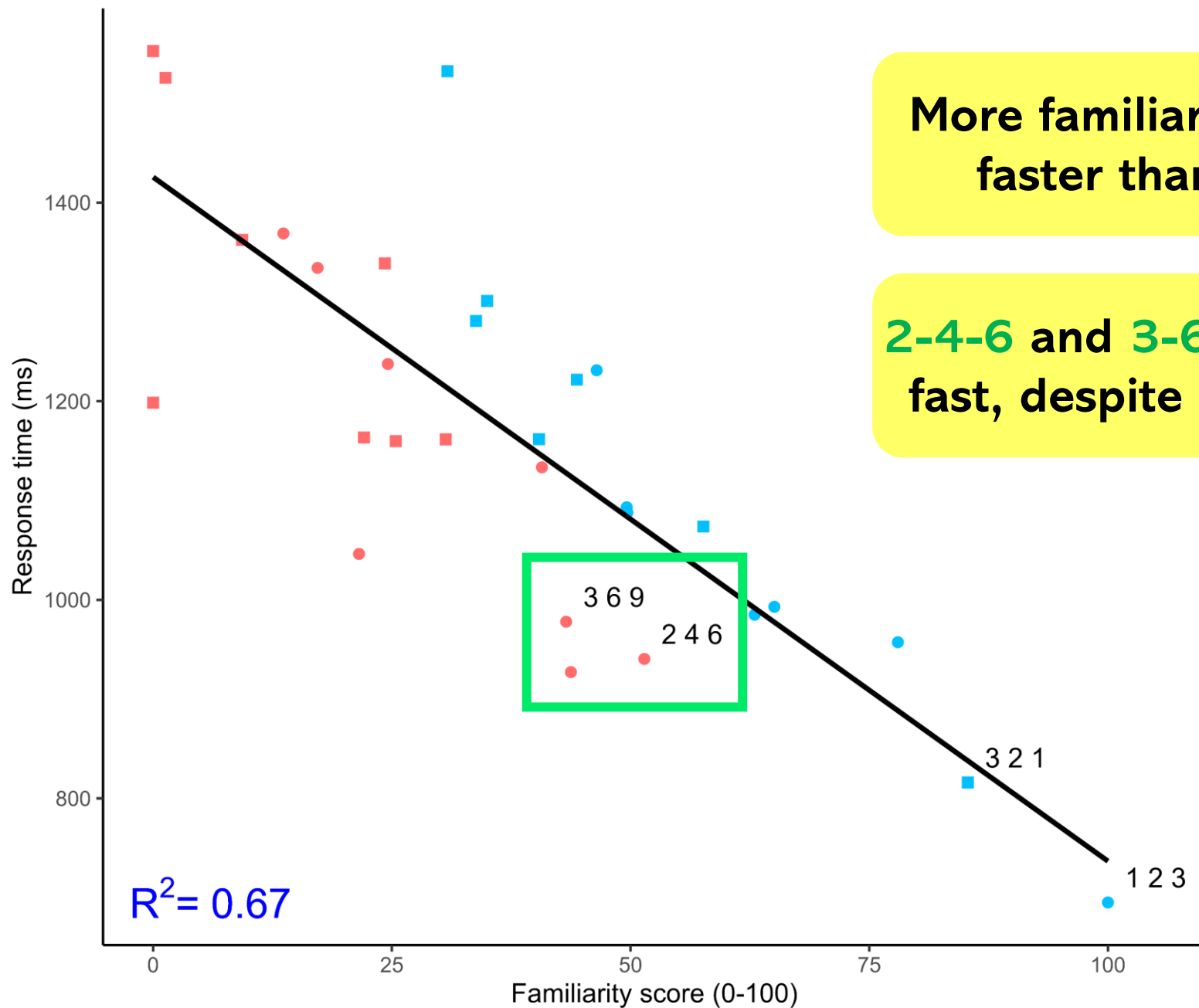
Non-consecutive

1 3 5 4 6 8 3 6 9

2 4 6 5 7 9

3 5 7 2 5 8

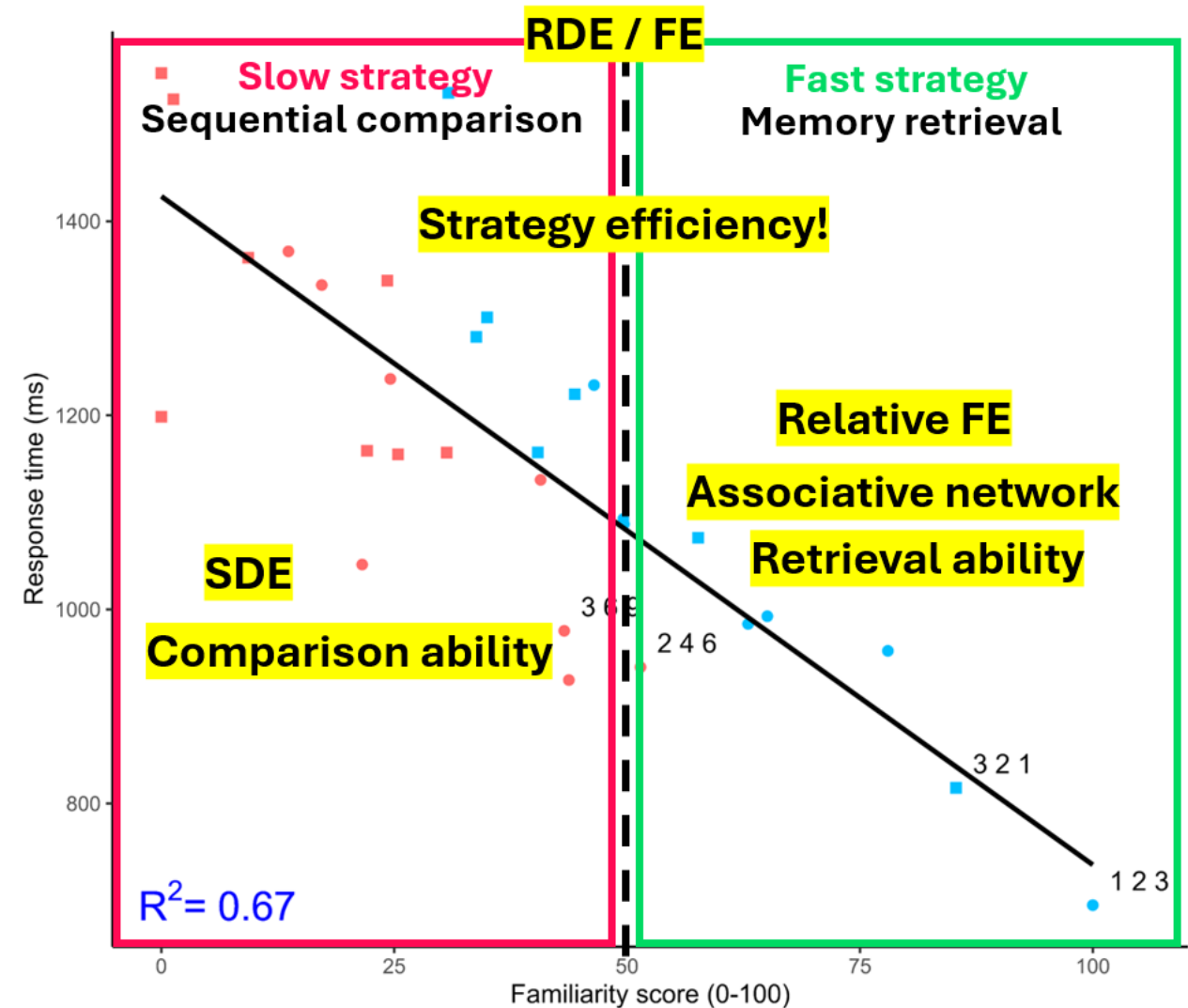
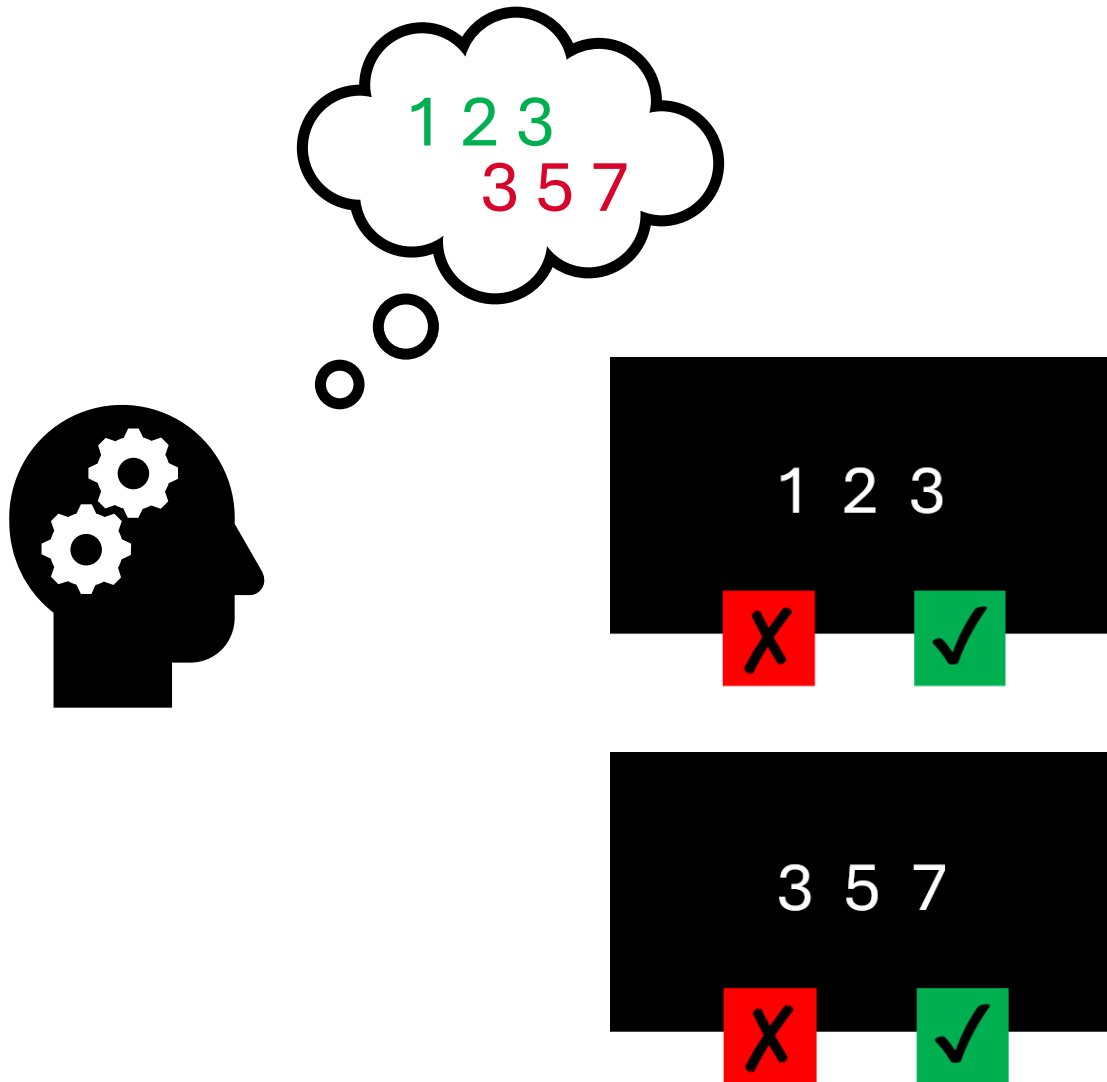
3 6 9



More familiar sequences were processed faster than less familiar sequences

2-4-6 and 3-6-9 were processed relatively fast, despite not matching the **count-list**!

Implications of the familiarity effect



OK, but...



How does this explain all those other effects?

These effects may be viewed as **specific instances** of a **general familiarity effect**

Order effect

234 vs 243



Reverse distance effect

123 vs 135



Direction effect

234 vs 432



Regularity effect

135 vs 136



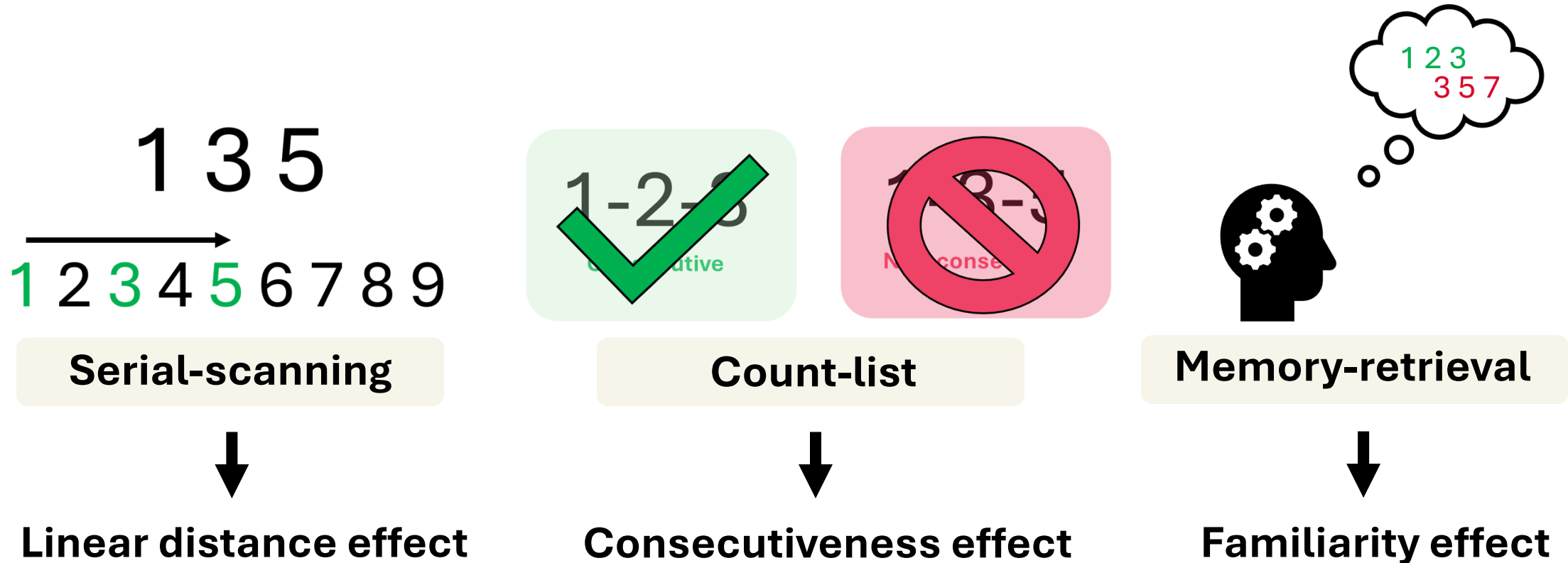
Single-digit effect

7-8-9 vs 17-18-19



What is my point!?

We may think of these as competing explanations / disagreements about what mechanisms underlie order processing:



...but they are actually disagreements about what data we are trying to explain

What makes a “good” explanation?

Preprint

Non-symbolic quantity estimation may be characterised by many signatures, yet we typically focus on only a few

Declan Devlin¹ and Natalia Dubinkina²

¹Department of Teacher Education, Faculty of Social and Educational Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway*

²Department of Communication Studies, University of Antwerp, Belgium

What we try to explain influences our explanations...

...we need to consider both (i) the *quality of an explanation* and (ii) what is actually being explained

Is explaining more effects better than explaining fewer?

Are more parsimonious explanations better than less parsimonious ones?

Are some effects more important than others?

Order effect

234 vs 243

Reverse distance effect

123 vs 135

Direction effect

234 vs 432

Regularity effect

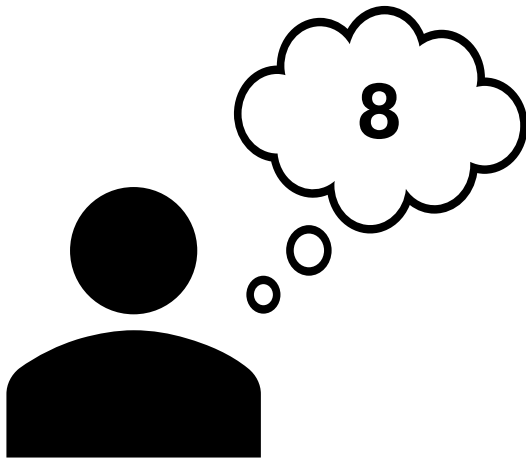
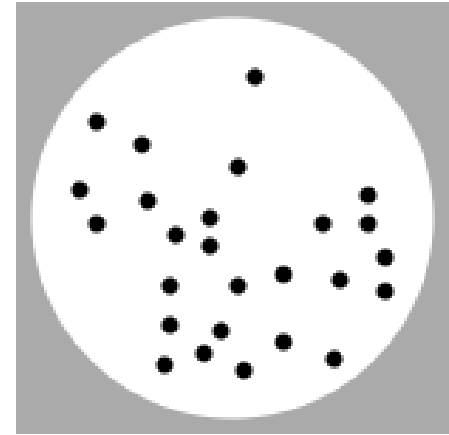
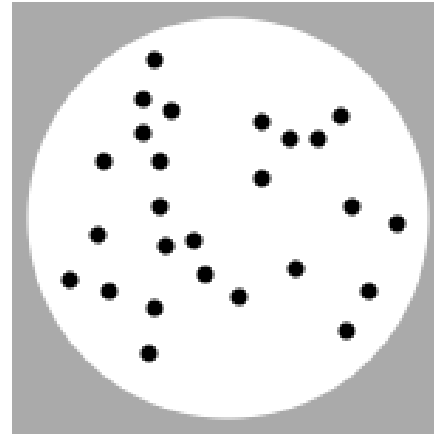
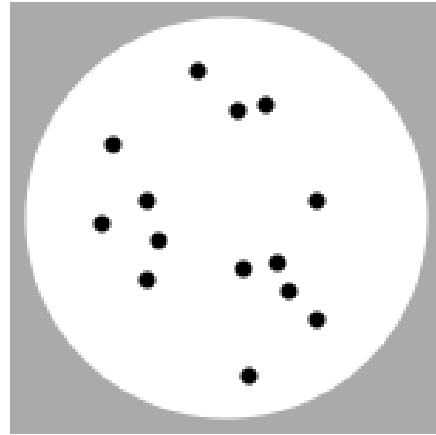
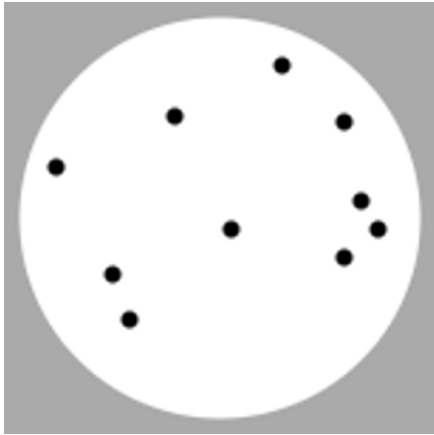
135 vs 136

Single-digit effect

7-8-9 vs 17-18-19

So what about quantity estimation?

How many dots?



People are able to estimate quantities reasonably well...

What are we trying to explain?

Estimation produces a set of systematic behavioural patterns:

Exact range

Small numerosities are estimated with near perfect accuracy

Over/under-estimation

Medium numerosities tend to be underestimated, larger numerosities tend to be overestimated

Scalar variability

Errors become increasingly variable as numerosity increases

The many effects of quantity estimation

Exact estimation
range

Influence of
non-numerical
visual cues

Overestimation of
medium numerosities

Numerosity-specific
error reductions

Estimation bias

Feedback-dependent
estimation bias

Underestimation of
large numerosities

Scalar variability

Serial dependence

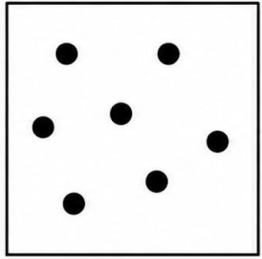
Stimulus-duration
effects

Sigmoidal response
time functions

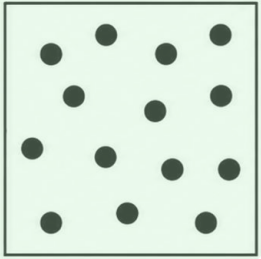
See Devlin & Dubinkina (preprint) for full list!

What we try to explain influences our explanations

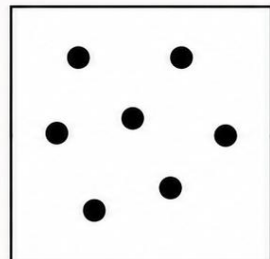
The ratio effect



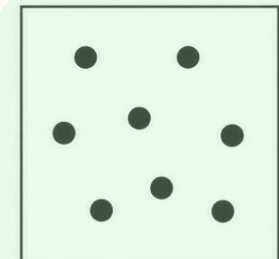
vs.



Large ratio = easier

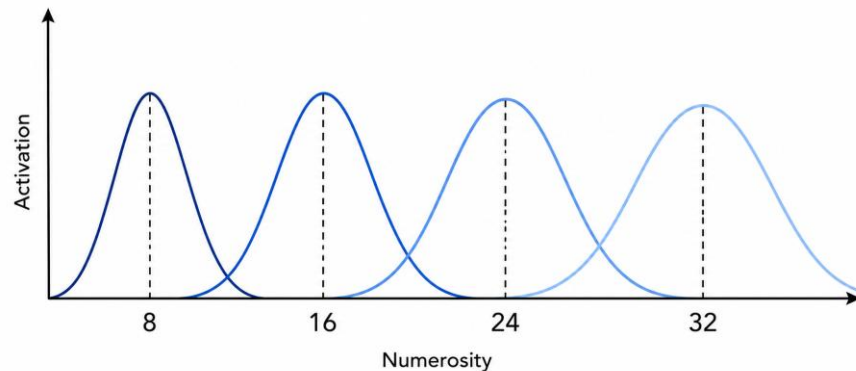


vs.

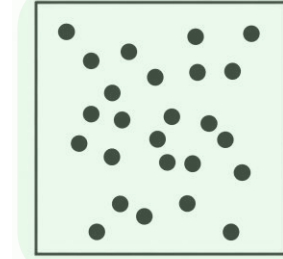


small ratio = harder

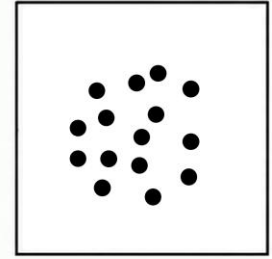
Naturally leads to ANS-type explanations



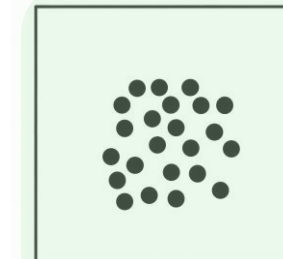
Visual cue effects



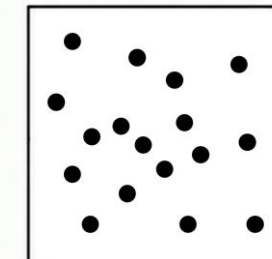
vs.



Congruent = easier



vs.



Incongruent = harder

Implies a sensory-integration account

Some points to consider

Preprint

Non-symbolic quantity estimation may be characterised by many signatures, yet we typically focus on only a few

Declan Devlin¹ and Natalia Dubinkina²

¹Department of Teacher Education, Faculty of Social and Educational Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway*

²Department of Communication Studies, University of Antwerp, Belgium

Exact estimation
range

Influence of
non-numerical
visual cues

Overestimation of
medium numerosities

Numerosity-specific
error reductions

Estimation bias

Feedback-dependent
estimation bias

Underestimation of
large numerosities

Scalar variability

Serial dependence

Stimulus-duration
effects

Sigmoidal response
time functions

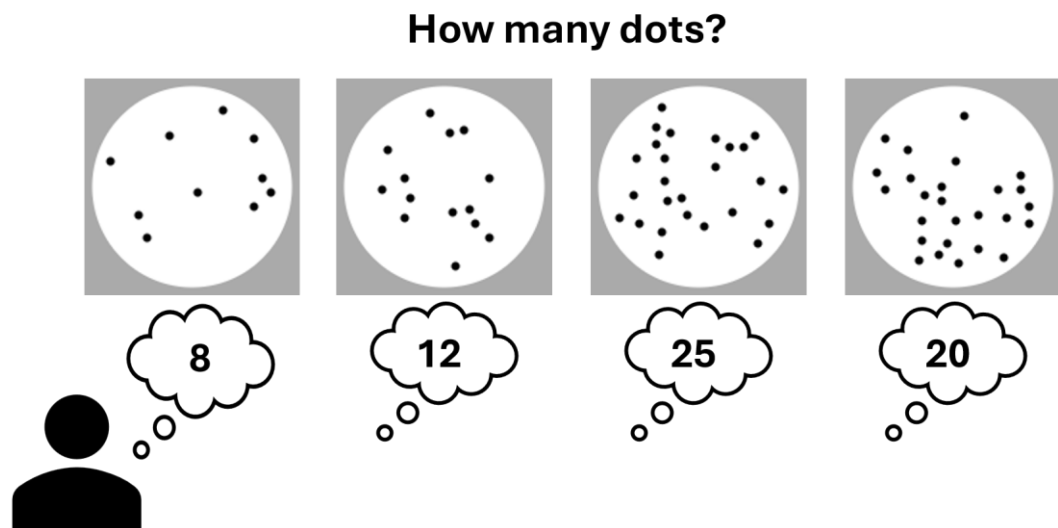
We try to explain estimation by explaining its effects

There are lots of effects, so we have to be selective...but what we focus on may influence our explanations

What look like competing explanations of the same data may actually be disagreements about what data we should be trying to explain

What makes a “good” explanation is a bit complicated...

What are we trying to explain?



Preprint

Non-symbolic quantity estimation may be characterised by many signatures, yet we typically focus on only a few

Declan Devlin¹ and Natalia Dubinkina²

¹Department of Teacher Education, Faculty of Social and Educational Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway*

²Department of Communication Studies, University of Antwerp, Belgium

Please read!

We want to identify the cognitive strategies and mechanisms underlying quantity estimation...

Exact estimation range

Influence of non-numerical visual cues

Overestimation of medium numerosities

Numerosity-specific error reductions

Estimation bias

Feedback-dependent estimation bias

Underestimation of large numerosities

Scalar variability

Serial dependence

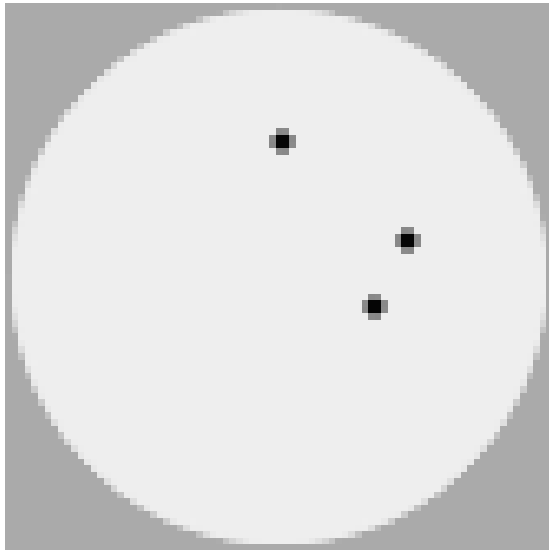
Stimulus-duration effects

Sigmoidal response time functions

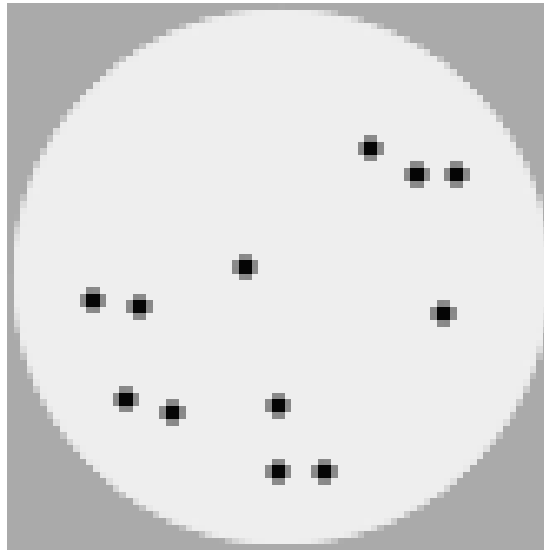
Which effects are we trying to explain?

Static vs. Contextual Effects

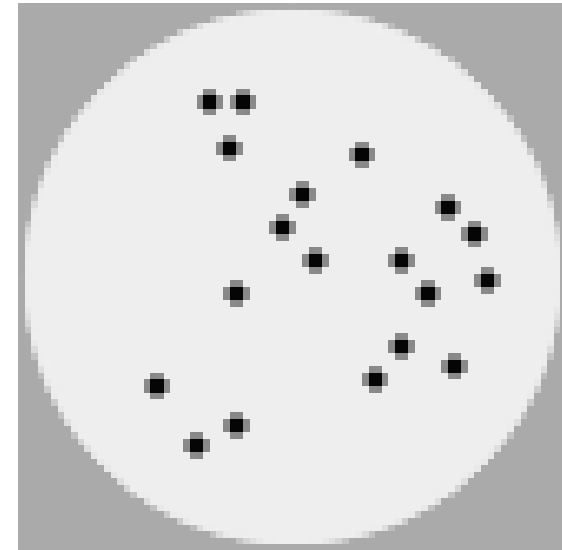
There are perhaps two kinds of effects we observe in estimation: static effects and contextual effects



Exact range



Underestimation



Overestimation

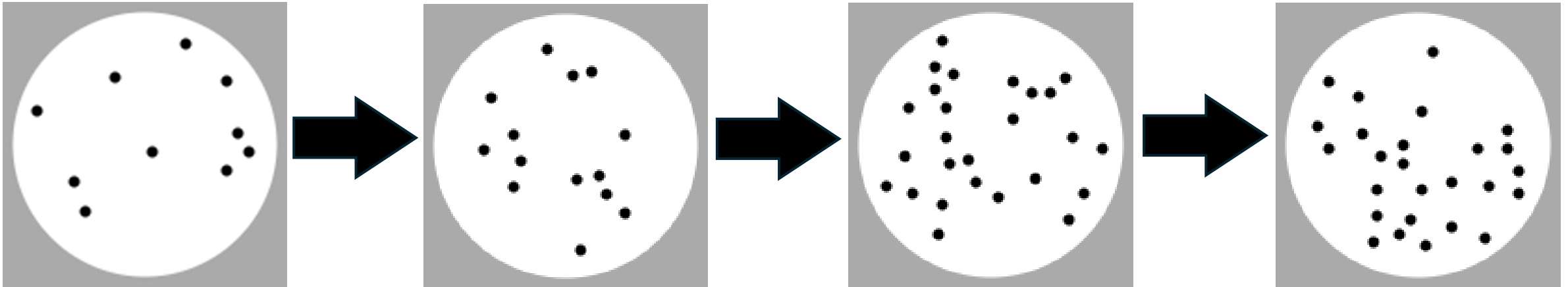
Static – because they concern a single instance

Contextual Effects

Some effects are **contextual** in nature because they involve the use of external information

Serial dependence

Previous estimates can influence subsequent estimates

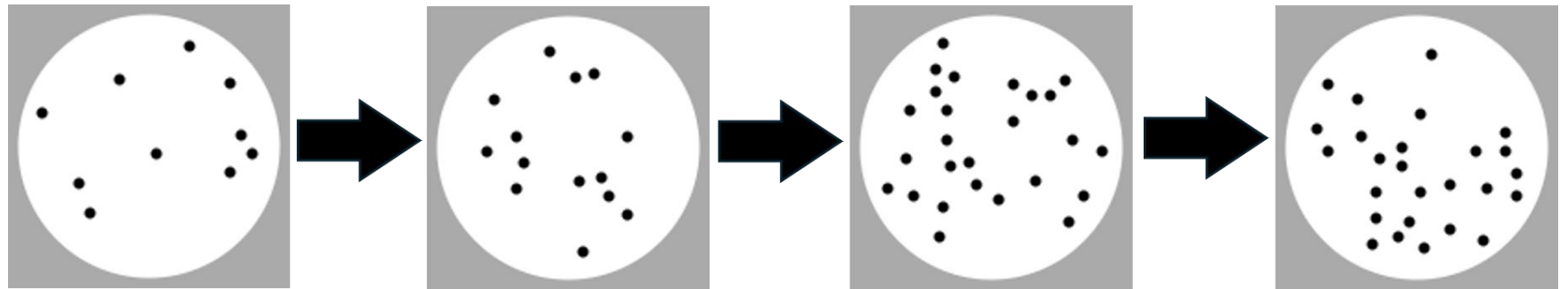
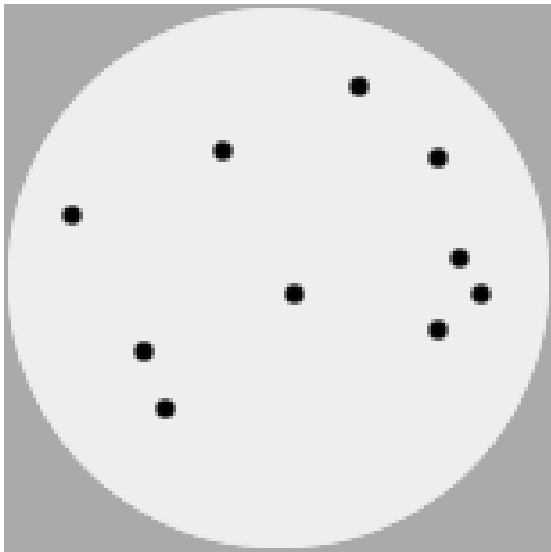


Contextual Effects

Estimates can also be influenced by providing feedback/references:

***Feedback-dependent
estimation bias***

This is 10:



What are we trying to explain?

We want to identify the cognitive strategies and mechanisms underlying quantity estimation...

Exact estimation range

Influence of non-numerical visual cues

Overestimation of medium numerosities

Numerosity-specific error reductions

Estimation bias

Feedback-dependent estimation bias

Underestimation of large numerosities

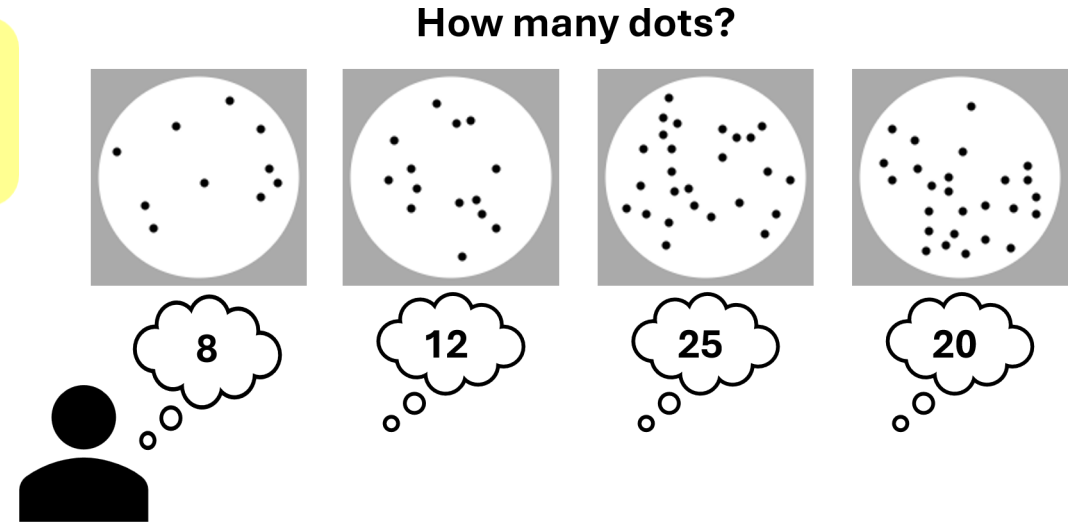
Scalar variability

Serial dependence

Stimulus-duration effects

Sigmoidal response time functions

There are lots of effects... and what we try to explain influences our explanations



I am interested in contextual effects:

Serial dependence

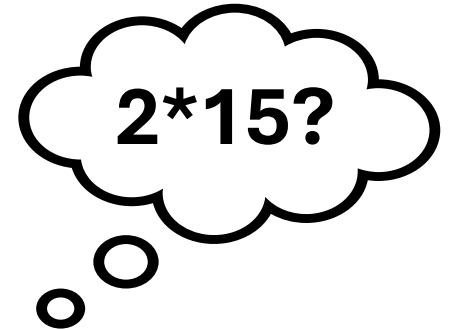
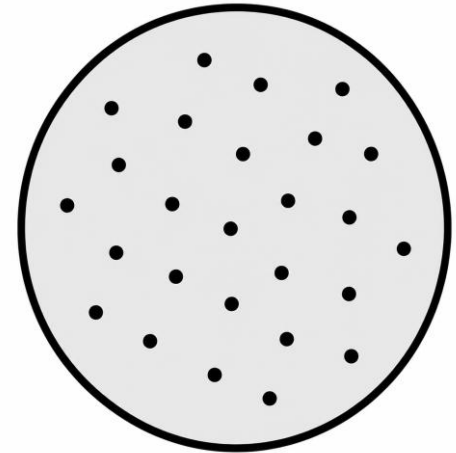
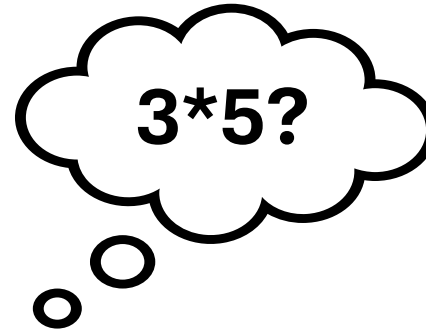
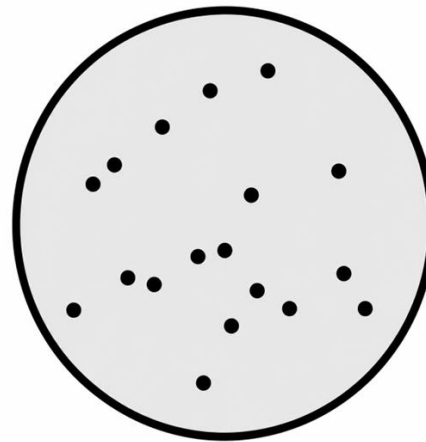
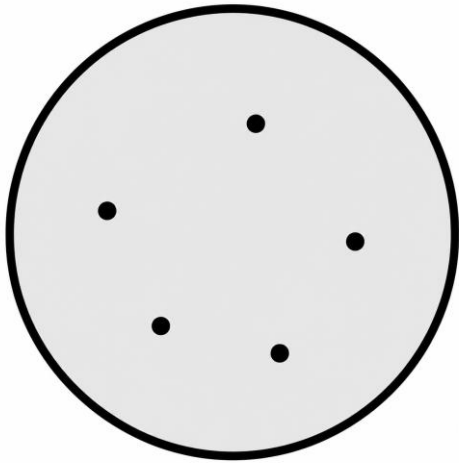
Feedback effects

2. Visual memory // comparison explanations

Estimation by mental comparison

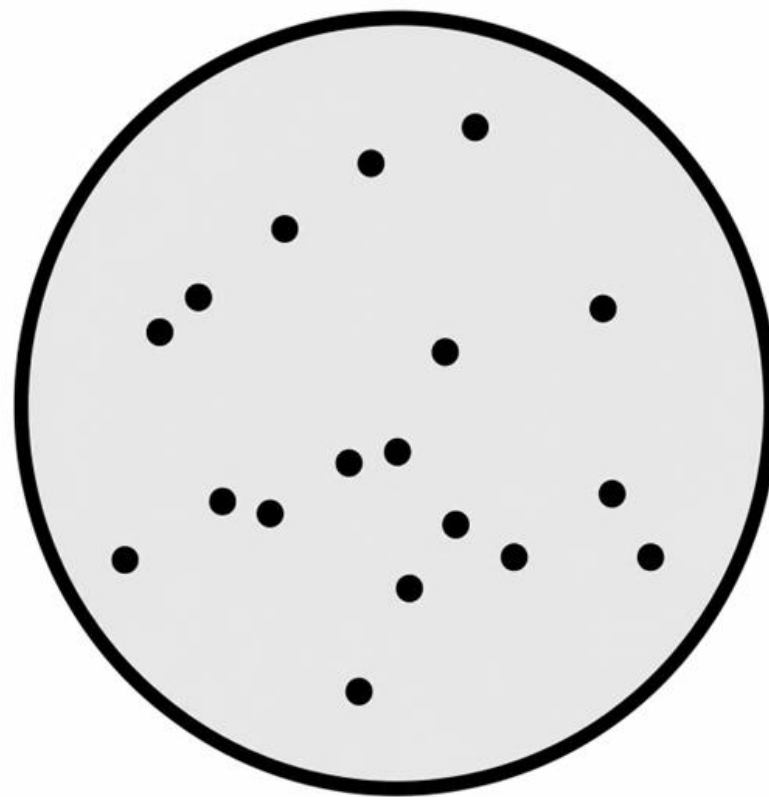
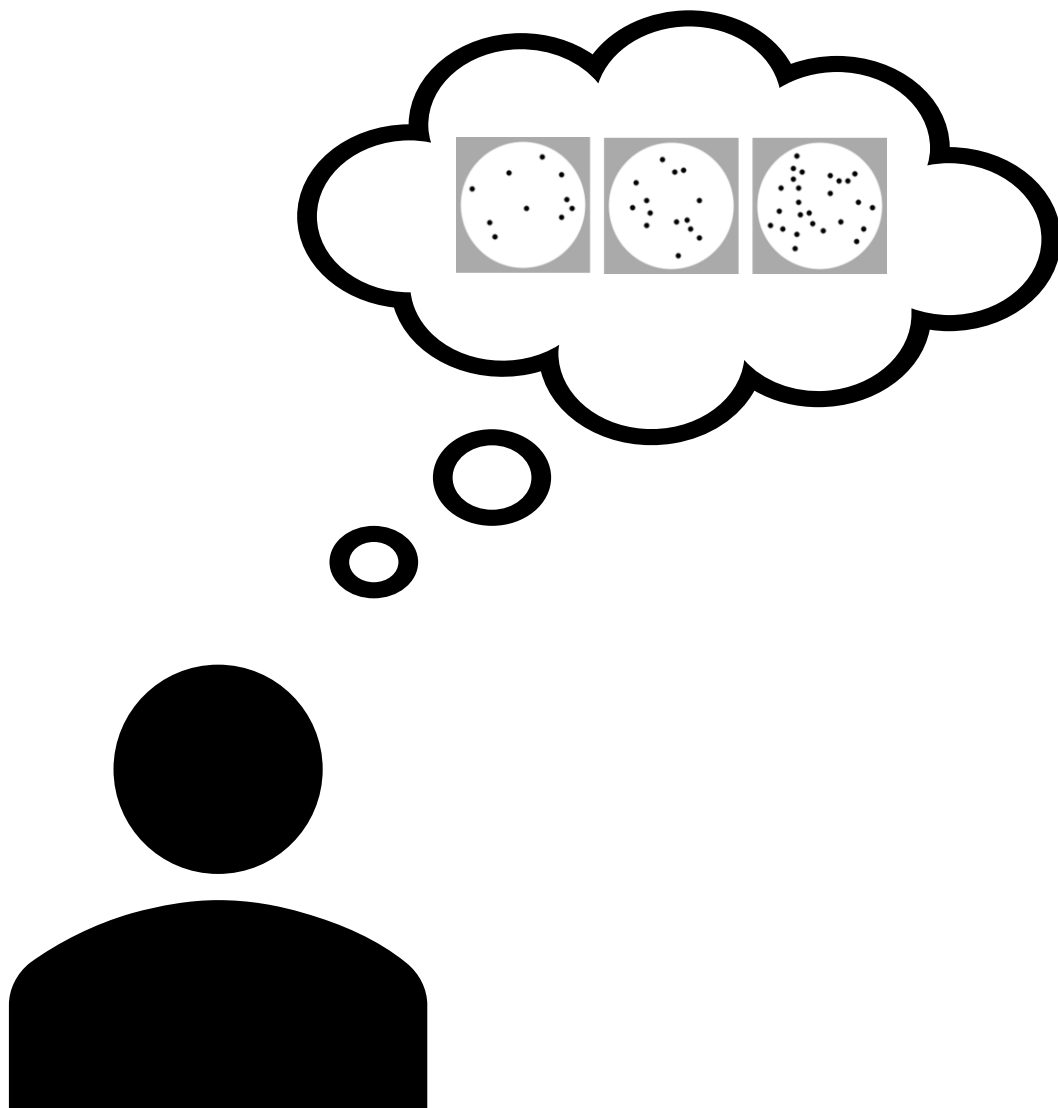
Serial dependence

Previous estimates can influence subsequent estimates



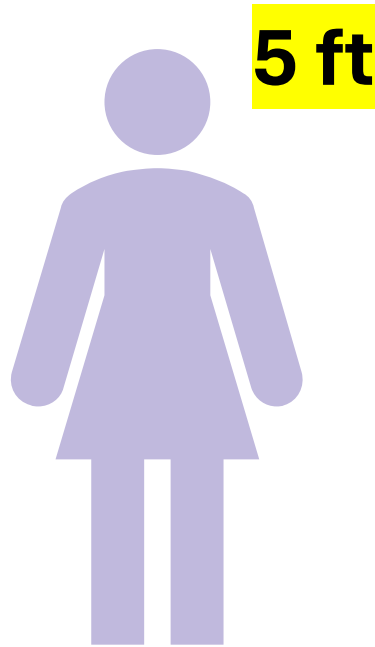
External reference images?

How many?



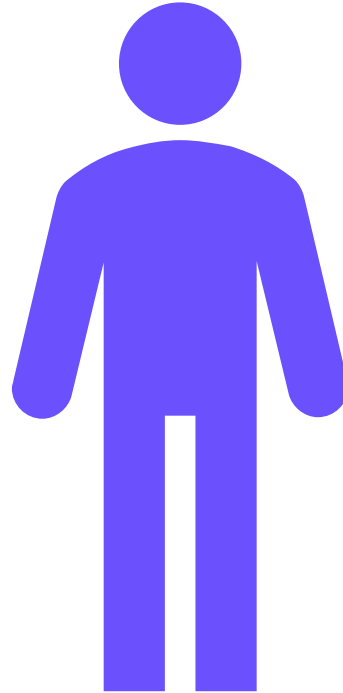
Estimating heights

How tall is Declan?



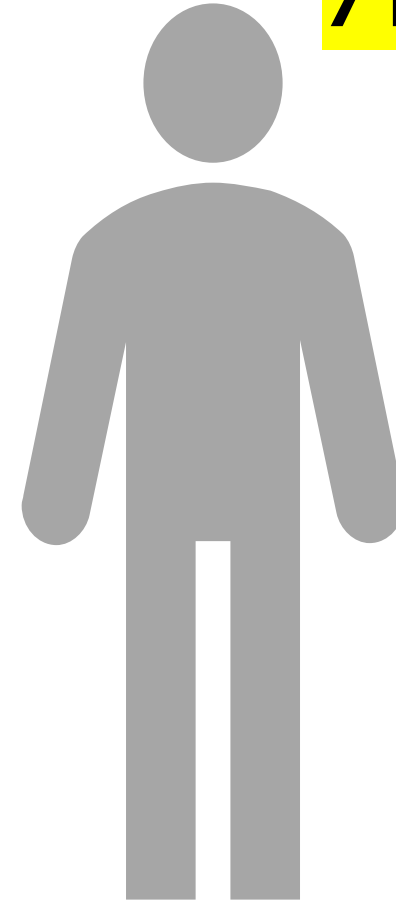
5 ft

Camilla



5' 6"?

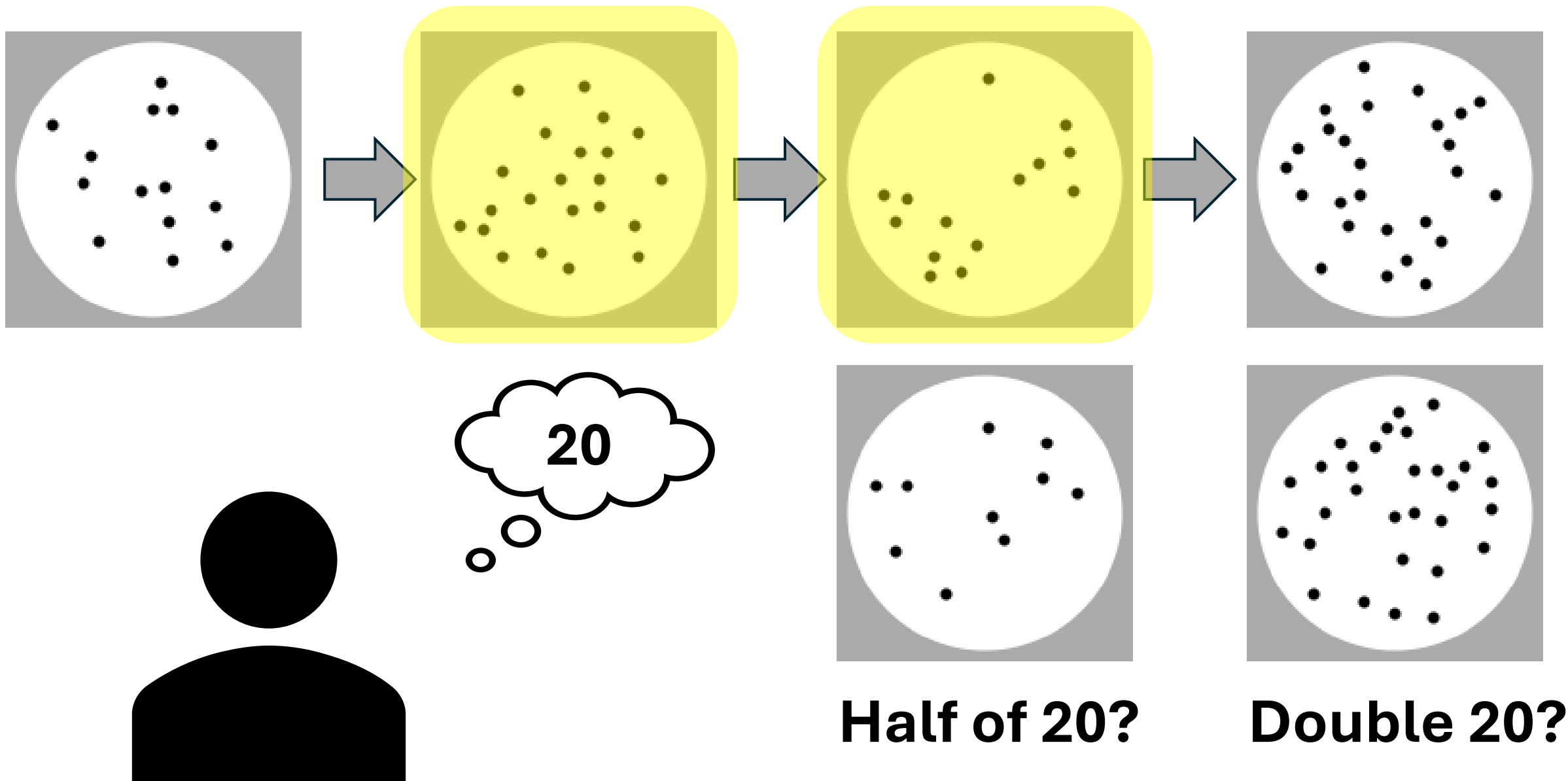
Declan



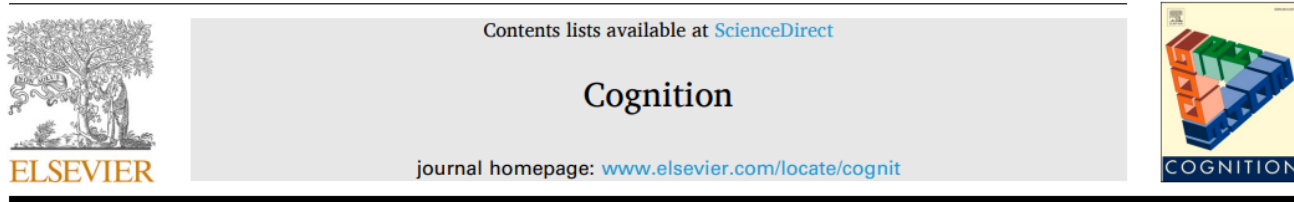
7 ft

Matthew

Internal reference images?



Evidence for this perspective



Decision-level processes in rapid numerosity estimation

Trygve Solstad^{a,1,*}, Eivind Kaspersen^{a,1}, Elisabeth I. Romijn^a, Jeremy Hodgen^{a,b,c}

^a Department of Teacher Education, Faculty of Social and Educational Sciences, NTNU – Norwegian University of Science and Technology, Trondheim, Norway

^b IOE, UCL's Faculty of Education and Society, University College London, London, United Kingdom

^c Observatory for Mathematical Education, School of Education, University of Nottingham, Nottingham, United Kingdom

ARTICLE INFO

Keywords:

Numerical cognition
Relative judgement
Number sense
Strategies

ABSTRACT

How do people rapidly judge numerosity—the number of objects in a scene? Theories of human numerical judgements are constrained by the behavioural patterns emerging when people estimate numerosity, such as exact enumeration of small sets, underestimation of large sets, and scalar variability. Current computational models capture how properties of stimulus representations can lead to characteristic behavioural patterns under rapid viewing conditions. Here, we examined which decision-level processes people engage during such tasks, and whether these processes contribute to behavioural patterns. We combined qualitative, correlational, and experimental studies of participants estimating brief (100 ms) and backward-masked dot patterns. In interviews, participants reported relying on visual memory to mentally compare a stimulus to previously estimated stimuli, count dots, and approximately add or multiply groups of dots. Quantitative multiple-choice surveys confirmed that the reported processes were systematically associated with numerosity, response time, and estimates. Pre-registered experiments further showed that three decision-level processes—comparing, partitioning, and counting—contribute to several behavioural patterns reported in prior research. These results suggest that many estimation patterns may reflect the flexible use of decision-level processes that depend both on individual factors and task demands. Formal models that integrate these decision-level processes with representation-level mechanisms may substantially enhance our understanding of rapid numerical judgements.

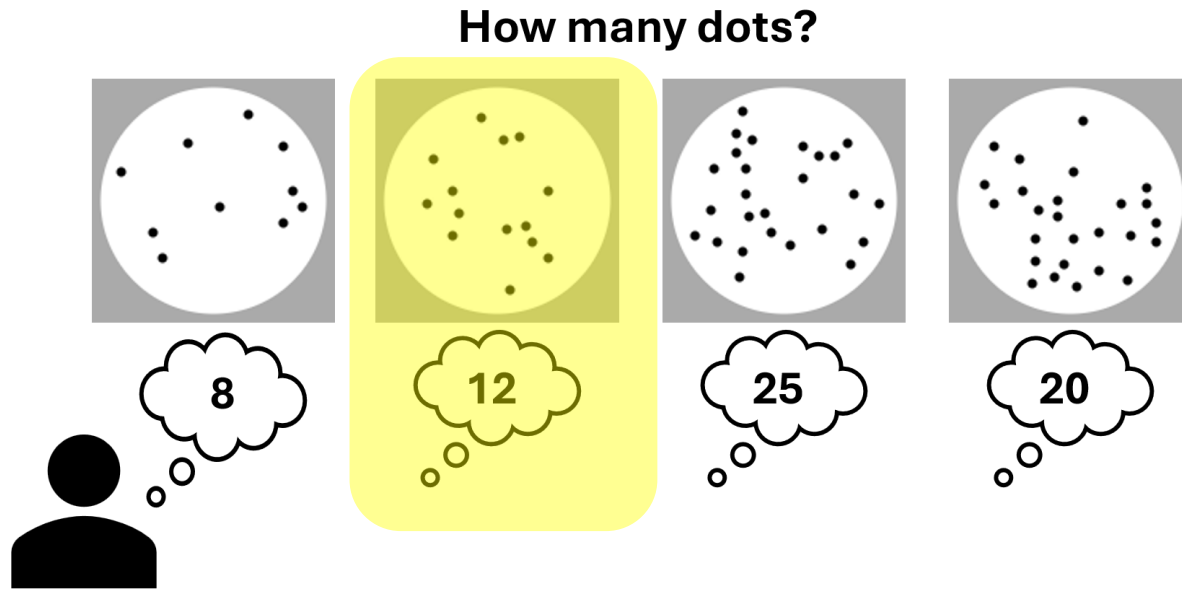
People frequently report making comparisons, either to previous trials or self-selected reference quantities:

“...it was a little more than the one of nine”.

“...it’s not like 50. It’s not 10. More than 20”.

Solstad et al. (2026)

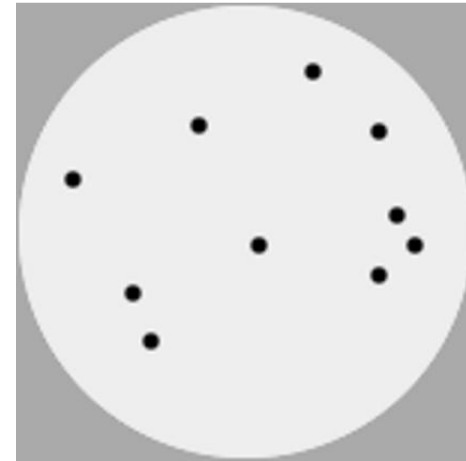
How this could explain contextual effects



Feedback effects

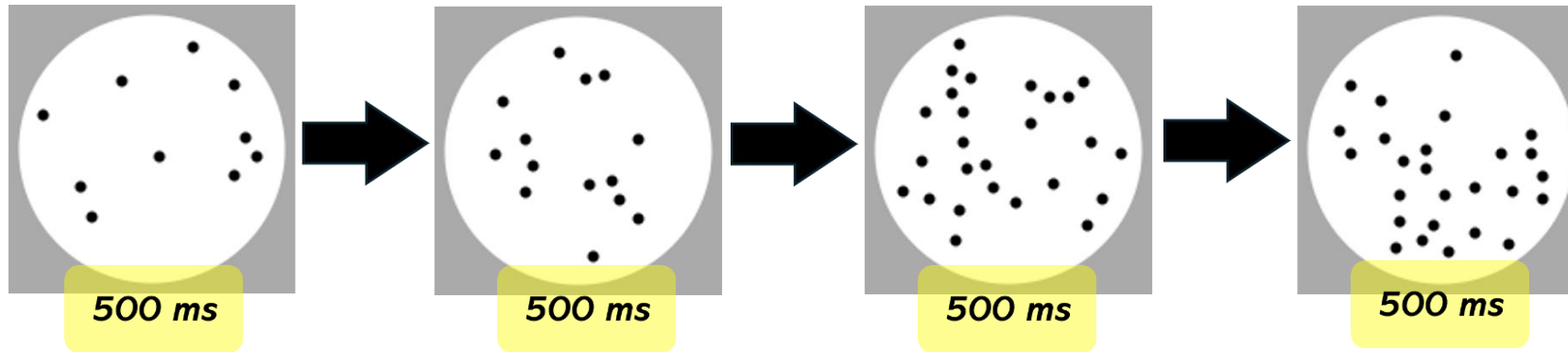
Serial dependence

This is 10:



But does it actually work like this?

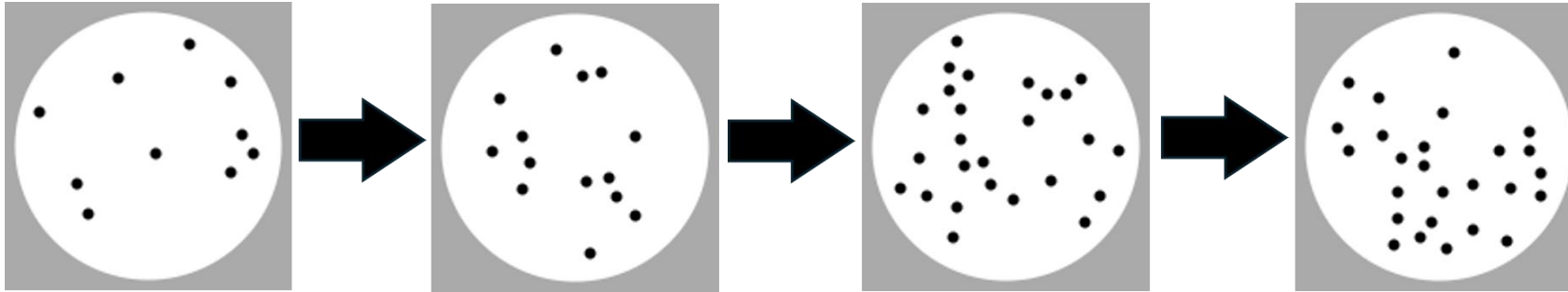
There are a few reasons to question whether this account is plausible...



1. Arrays are presented sequentially and each only very briefly (e.g., 100–500 ms)
2. Even if possible, do people naturally do this during estimation?
3. Do people remember the exact spatial patterns or something more abstract?

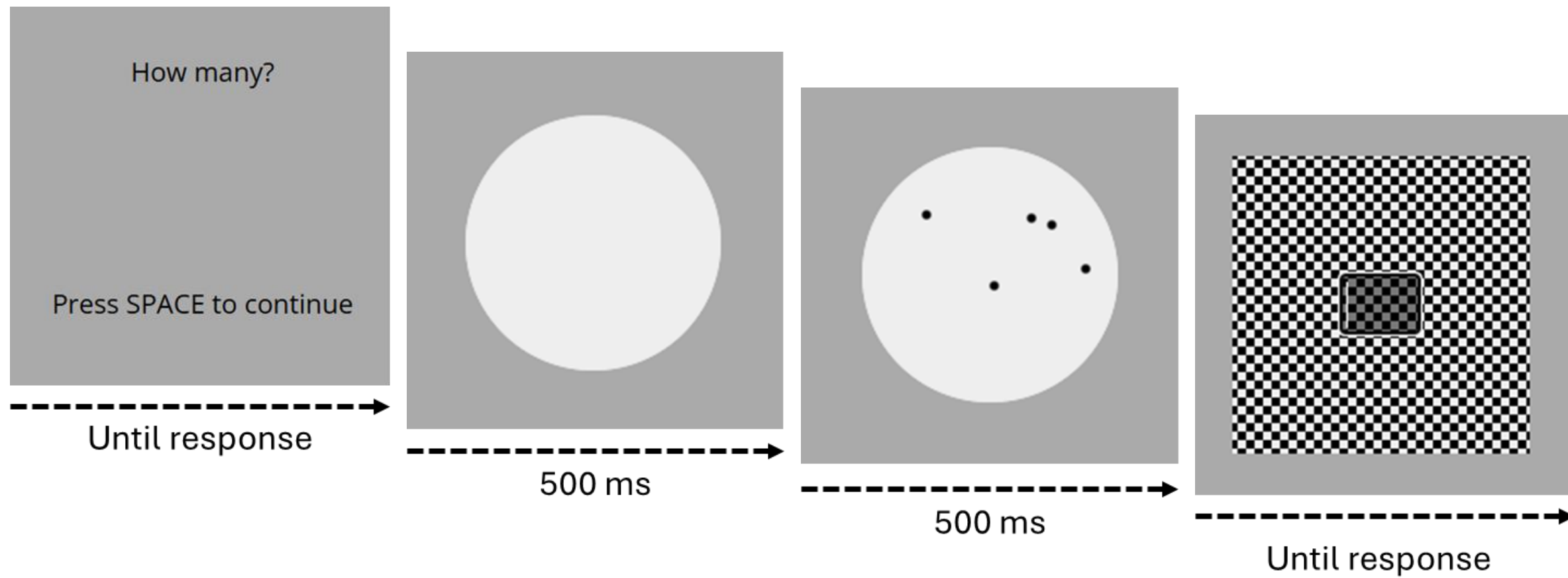
3. Some data to answer these questions! :)

Research questions



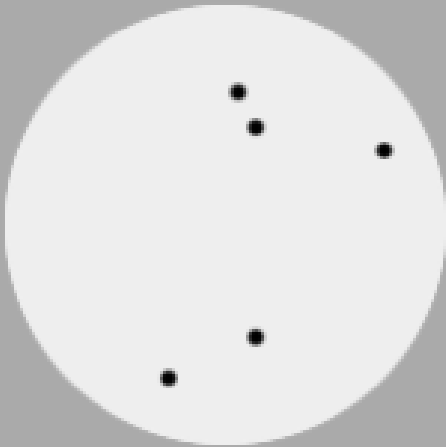
1. Can people remember dot patterns during a quantity estimation task?
2. Do they naturally do this, without actively trying to remember them?
3. Do people have conscious access / insight into their memories of the patterns?

1. Estimation task



2. Memory task

Which image did you see?



Numerosities: 1, 2, 5, 10, and 20

3. Confidence rating

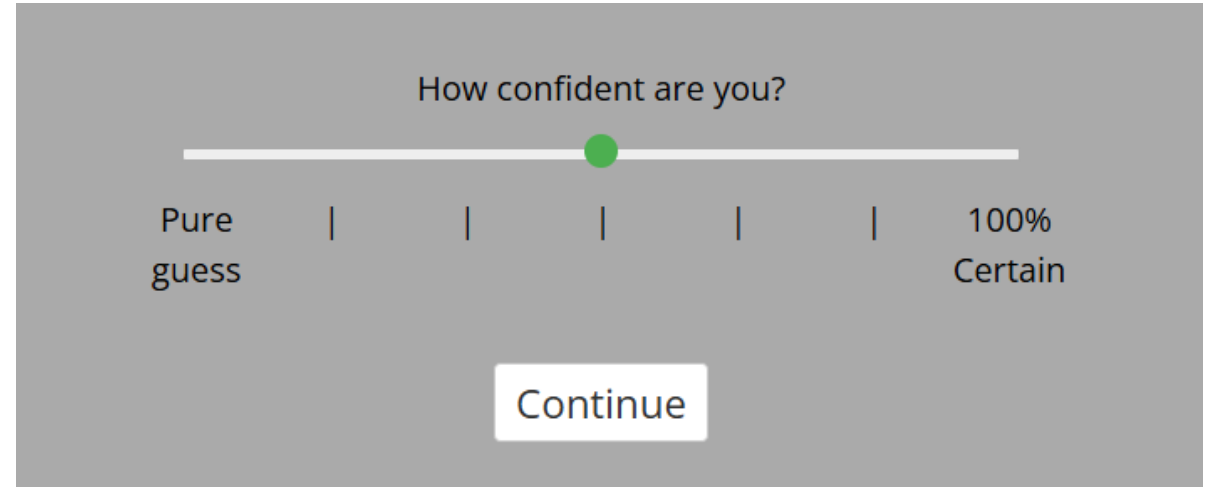
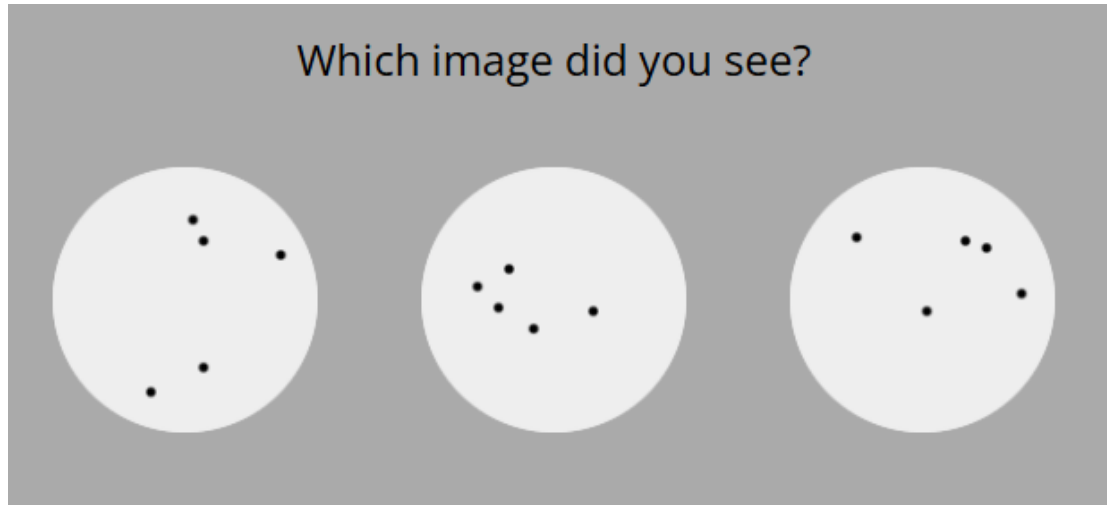
How confident are you?



Pure guess | | | | 100% Certain

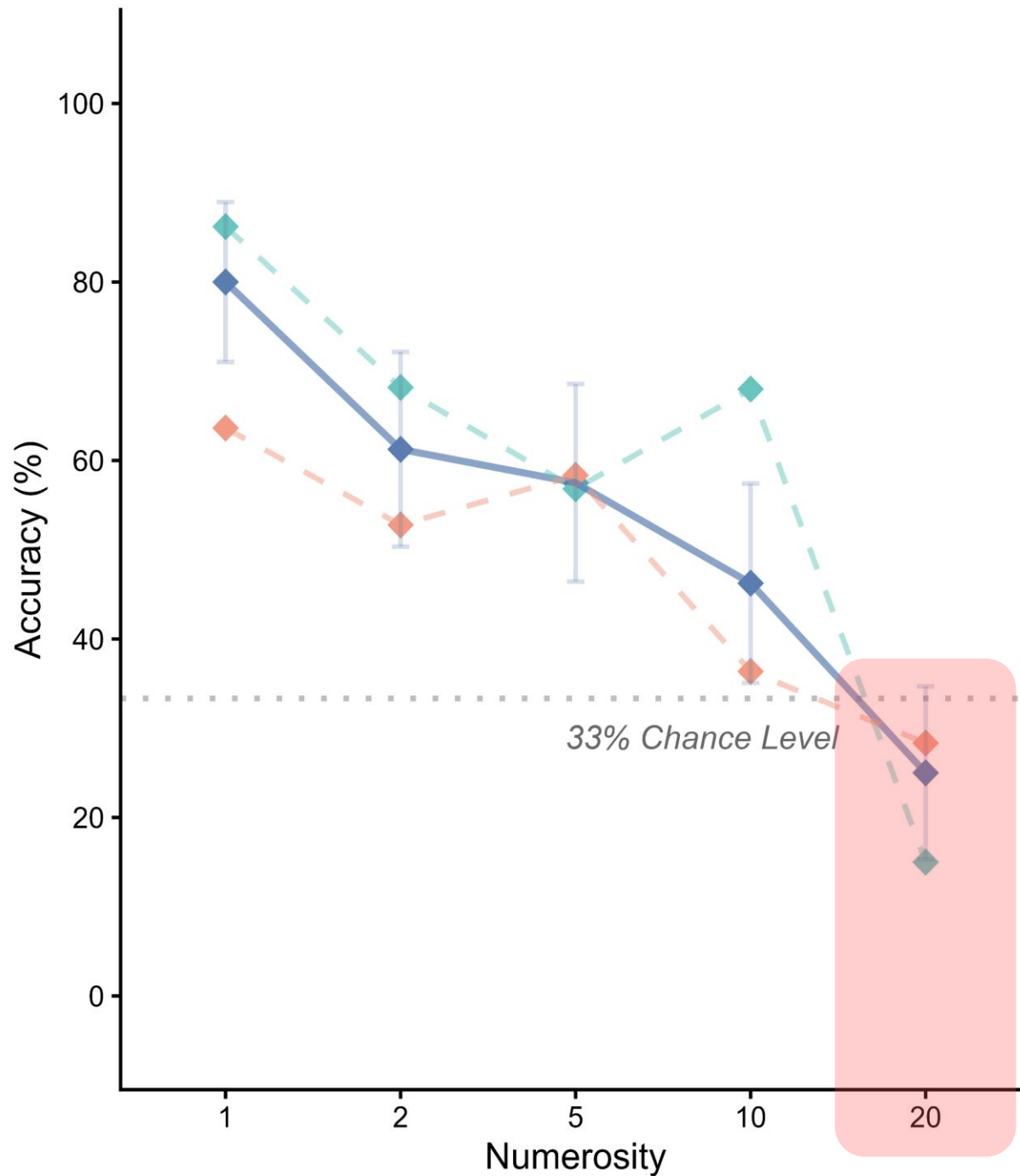
Continue

Predictions



1. Participants will show above-chance (33%) accuracy across all numerosities
2. Memory accuracy will be associated with confidence ratings

80 participants



Accuracy: 54%

Confidence for correct responses: 53.8

Confidence for incorrect responses: 36.8

Confidence

High Confidence

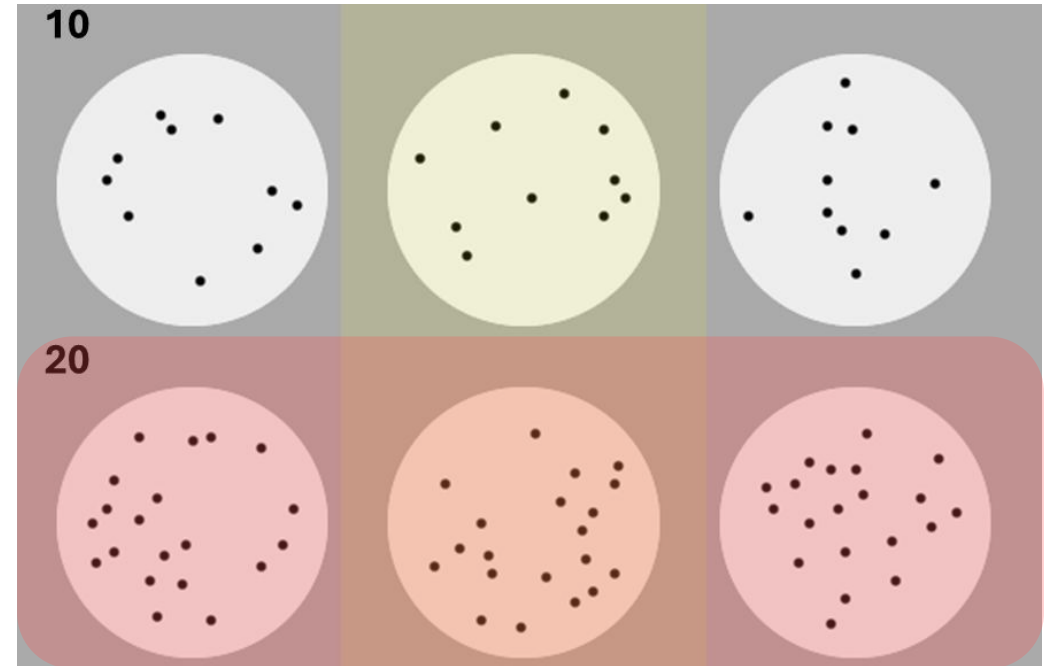
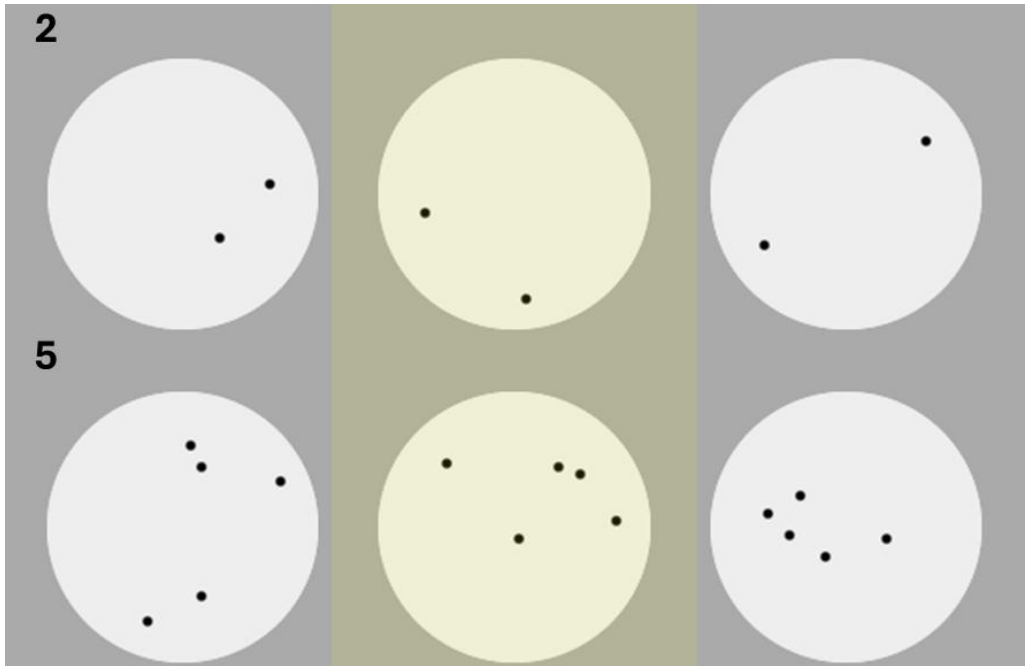
Low Confidence

Overall

Why low accuracy on numerosity 20?

Hypothesis A: people simply cannot retain large numerosities in memory

Hypothesis B: Image similarity makes recognition harder for larger numerosities



Experiment 2 – larger numerosities

We aimed to differentiate between these two hypotheses by (i) using a range of larger numerosities, and (ii) manipulating image similarity.

Exp1 numerosities: 1, 2, 5, 10, and 20

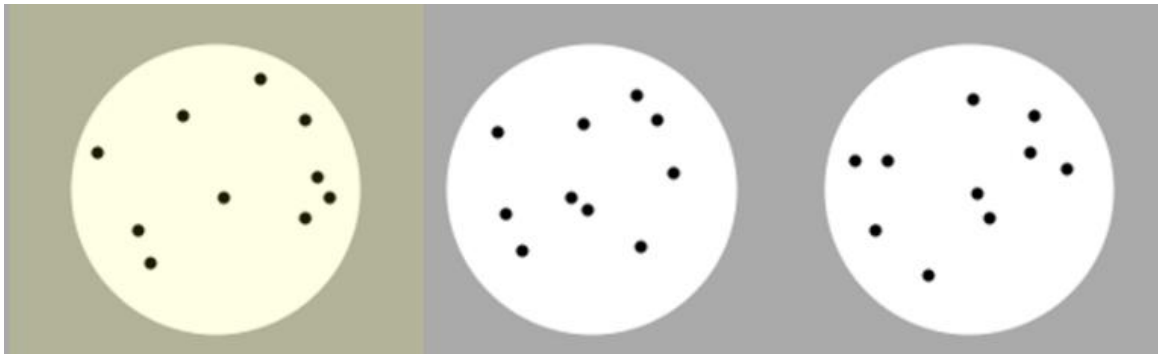
Exp2 numerosities: 10, 14, 18, 22, 26, and 30

Range centred around where we saw decreased accuracy in Experiment 1 (i.e., numerosity 20)

Experiment 2 – image similarity

For each numerosity, we had a high similarity and a low similarity trial:

High similarity



Low similarity



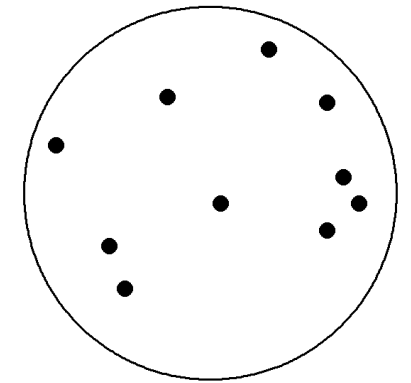
Each participant was only tested on each participant once. High/low similarity trials were counterbalanced across participants

	Group A	Group B
10	High	Low
14	Low	High
18	High	Low
22	Low	High
26	High	Low
30	Low	High

Experiment 2 – image similarity

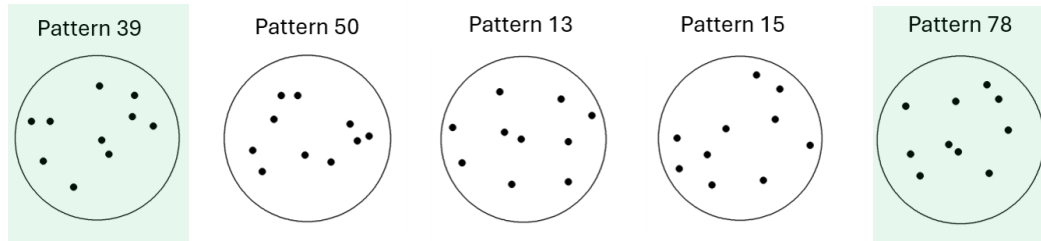
We combined human and algorithmic judgement to select high/low similarity patterns

1. Generated 100 random patterns for each numerosity
2. Randomly selected one as the target pattern
3. Used the *Hungarian algorithm* to identify 5 most and 5 least similar patterns

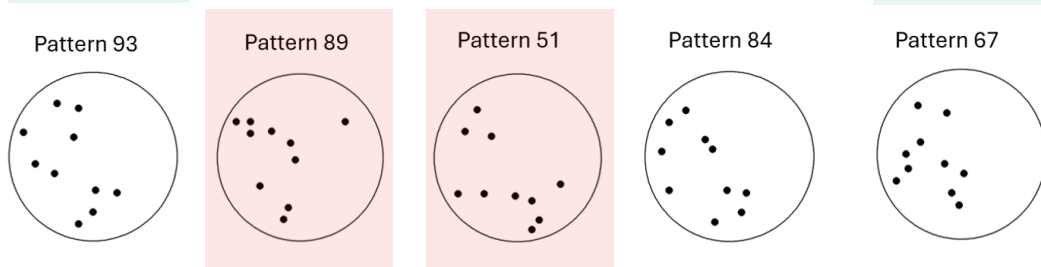


Target (10)

High

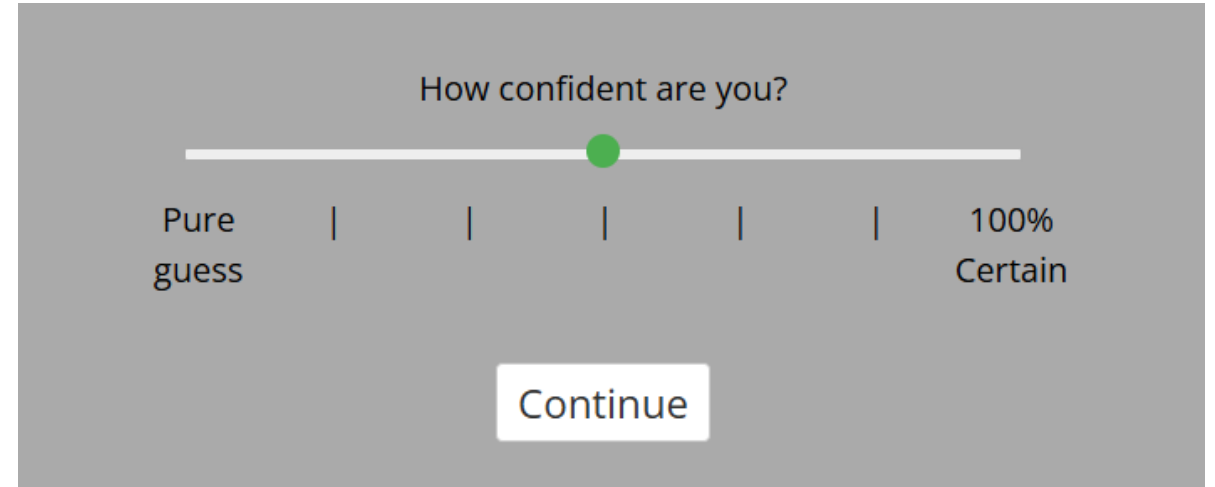
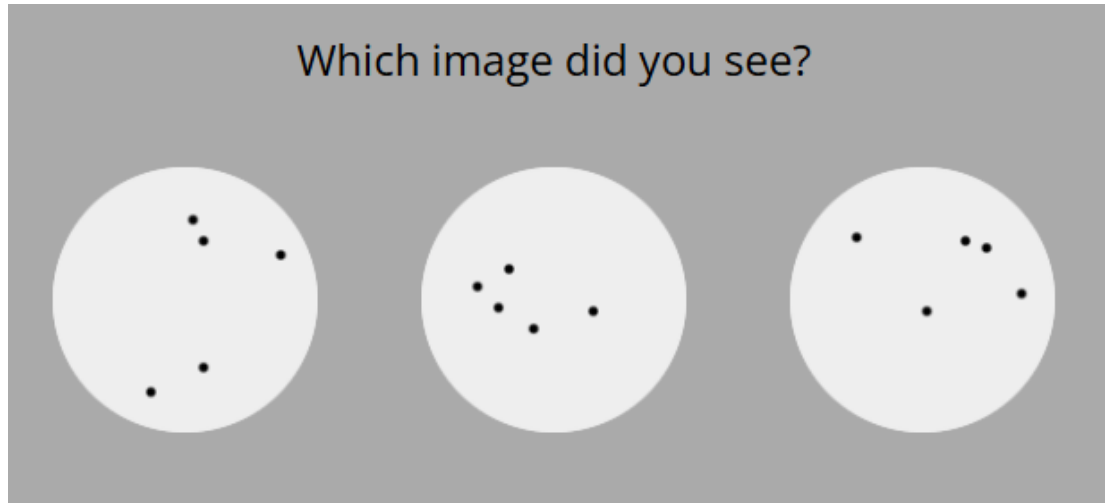


Low



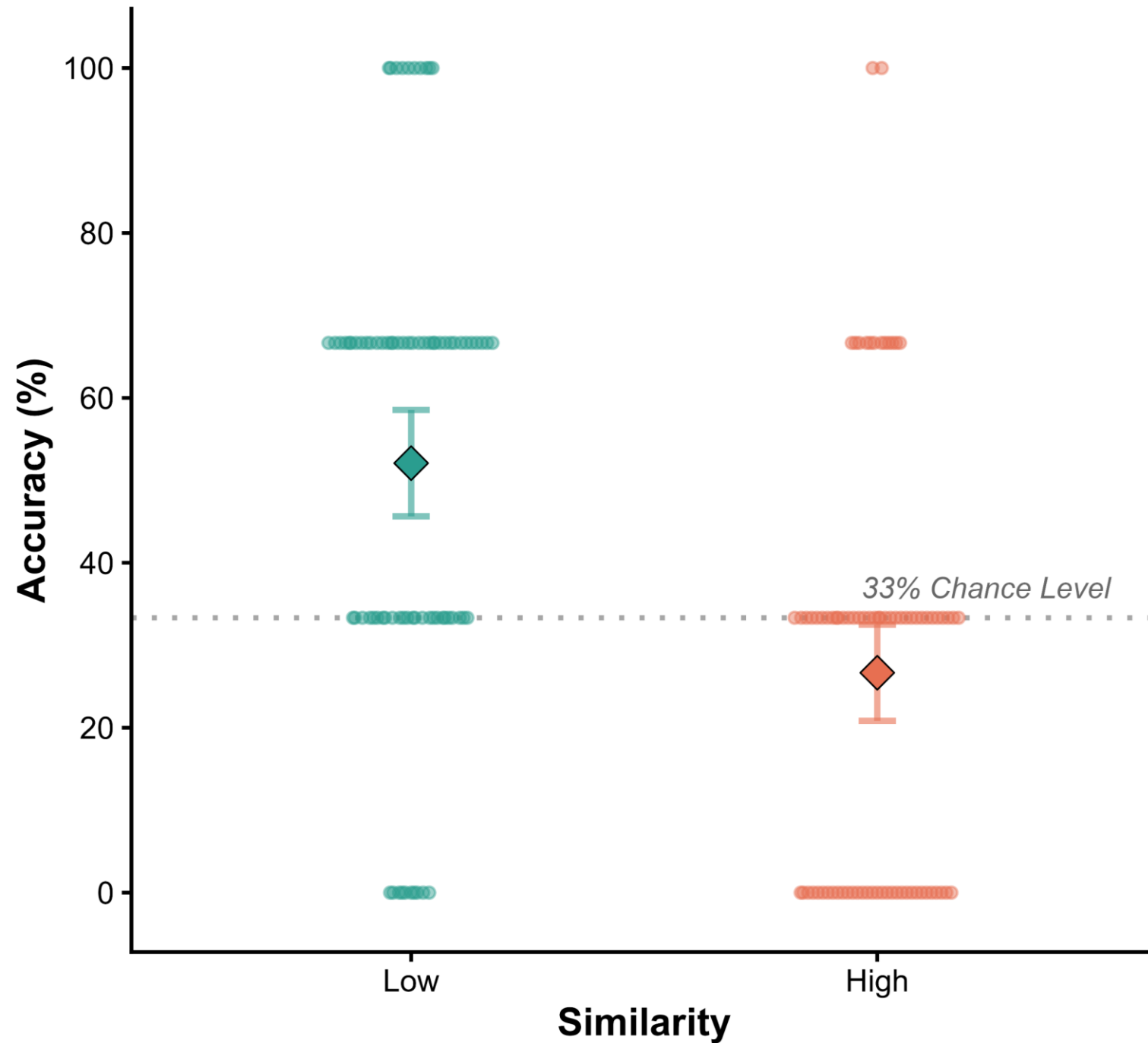
4. The three researchers ranked each of these, and selected the two highest ranked patterns for each condition

Predictions



1. Participants be able to recognise patterns above chance even at larger numerosities
2. But this will depend on image similarity of the distractor patterns
3. Memory accuracy will be associated with confidence ratings

80 participants



Overall accuracy: 39.4%

Low similarity accuracy:
52.1%

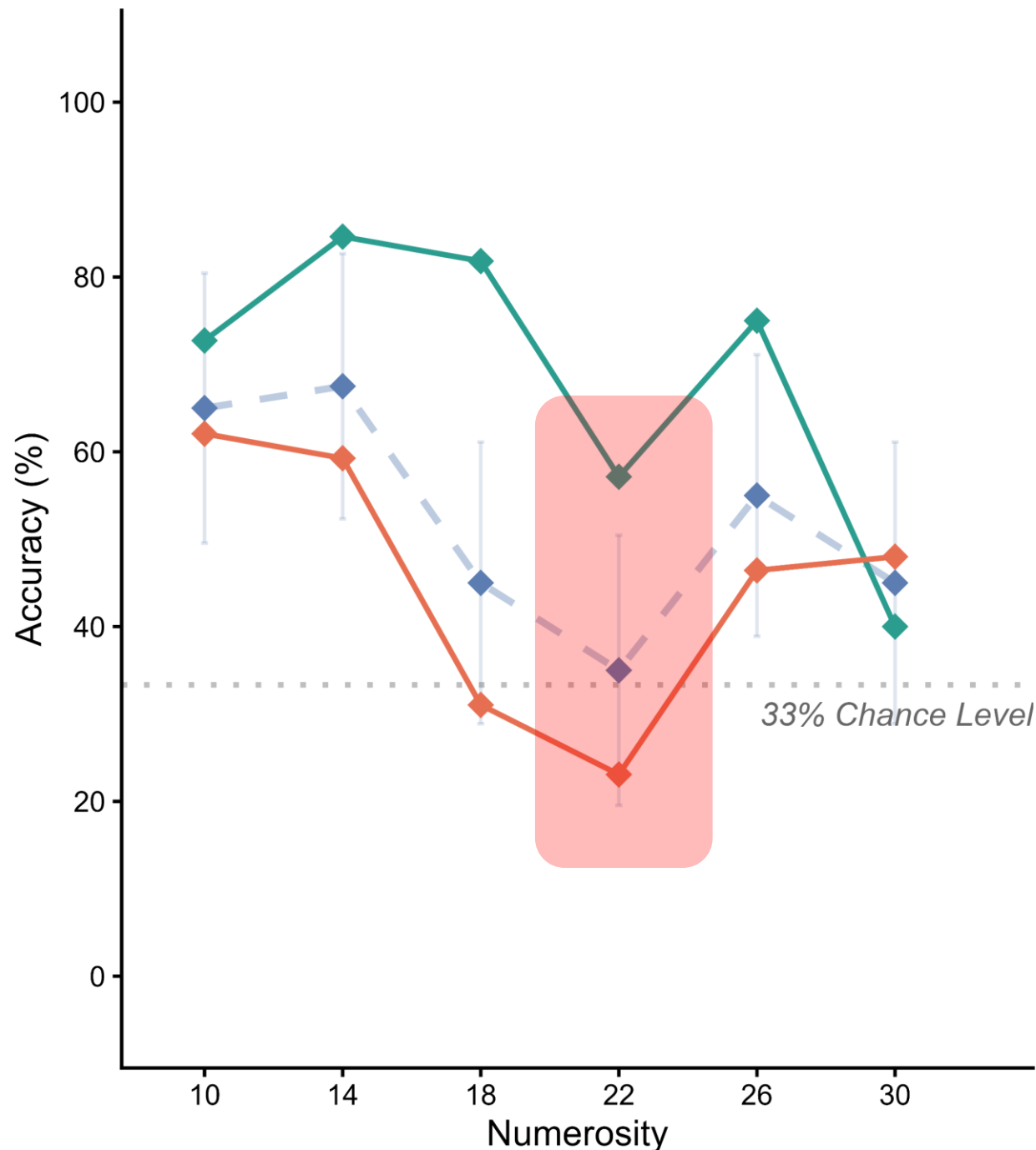
High similarity accuracy:
26.7%

Low similarity trials only

Accuracy: 52.1%

Above chance, even for this larger numerosity range

Significant effects of similarity and confidence, but not numerosity



Why lower accuracy on 22?

Were the low similarity patterns actually low similarity?

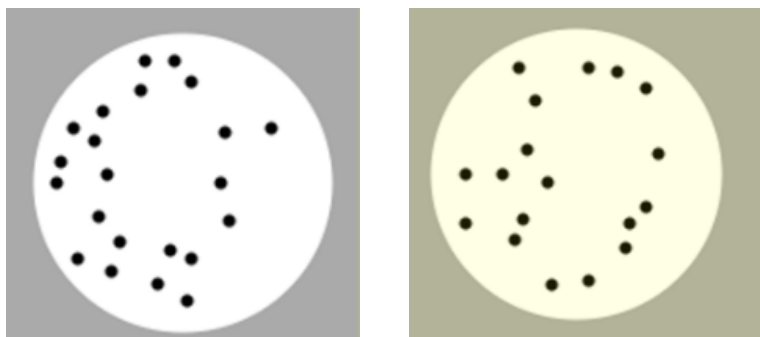
Numerosity 22



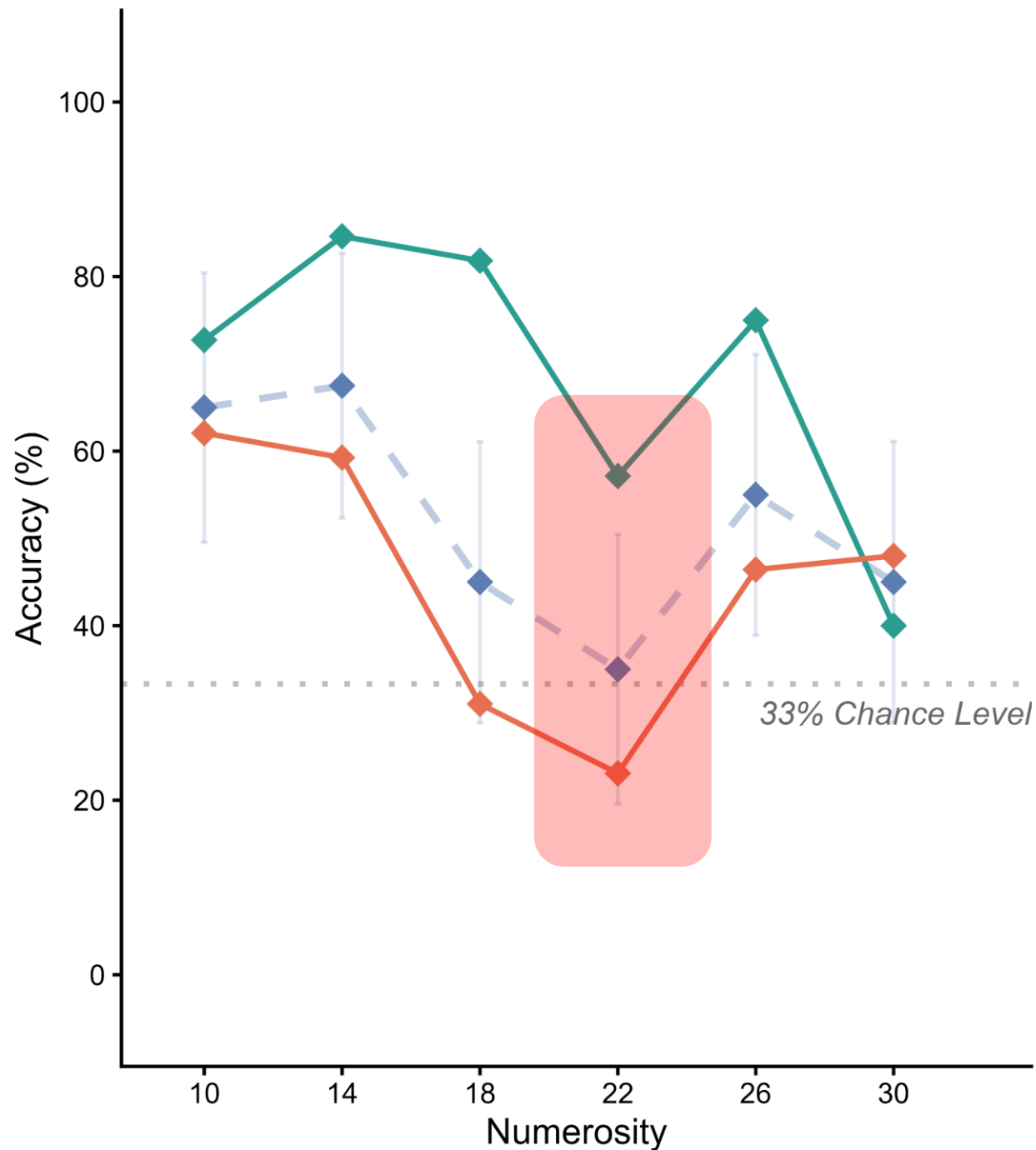
Maybe the distractor patterns were similar to some

False 22

True 18



Low similarity trials only



Whatever the reason for the low accuracy for 22, it is not because it is a larger numerosity

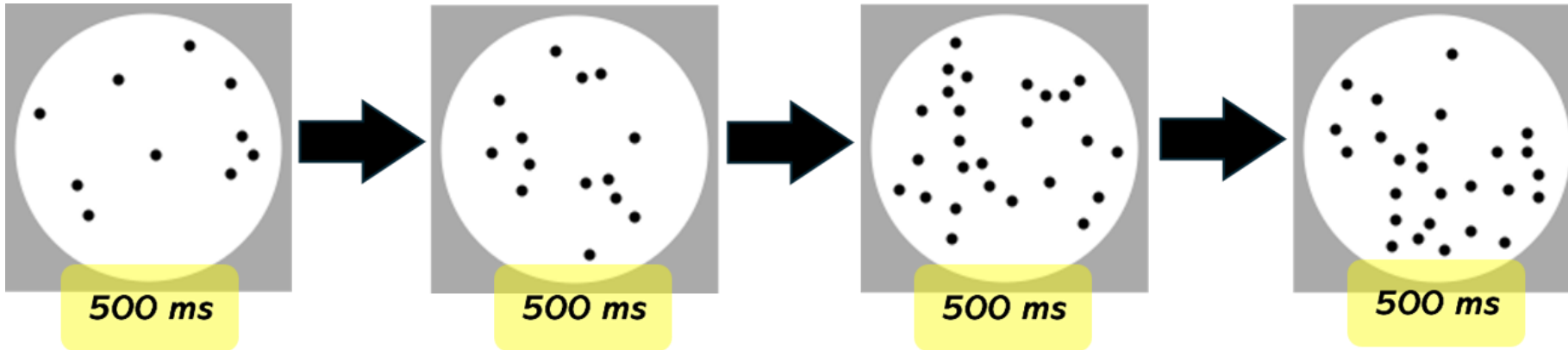
Confidence

- High Confidence
- Low Confidence
- Overall

4. Summary, implications, & directions for future research

What this study tells us...

People can remember dot patterns – even when presented briefly & sequentially!



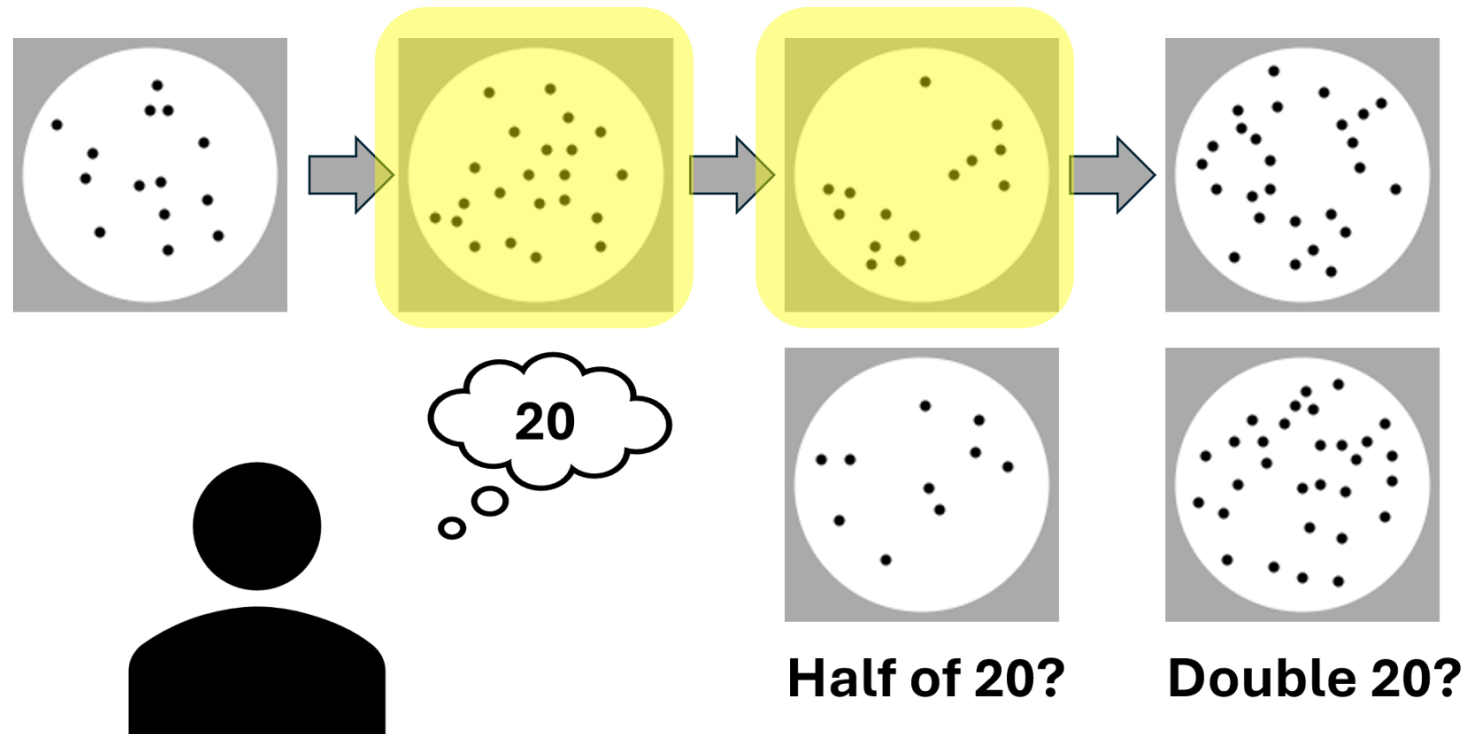
People do this naturally / spontaneously

For both small and large numerosities (at least up to 30)

And they have some degree of conscious insight into this

What this study *doesn't* tells us...

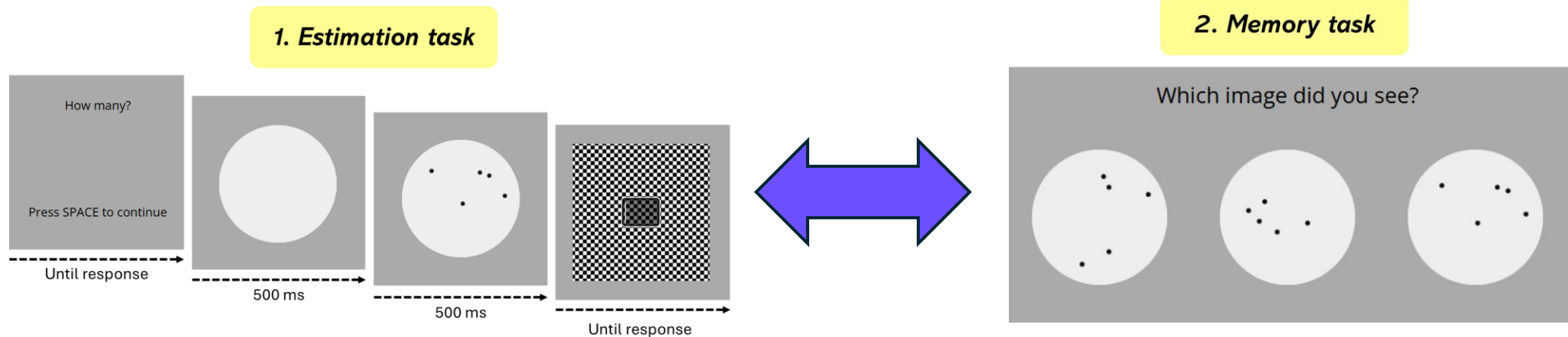
1. We don't know whether people actually *use* these memories during estimation...



2. We don't know whether people remember the exact spatial patterns, or something more abstract...

Idea for study 2

If people actually use these memories, we would expect memory accuracy to be associated with quantity estimation performance



...but not everyone is expected to use this strategy

...so, we may only expect an association in individuals who report using a comparison strategy

What strategy did you use?

Comparison

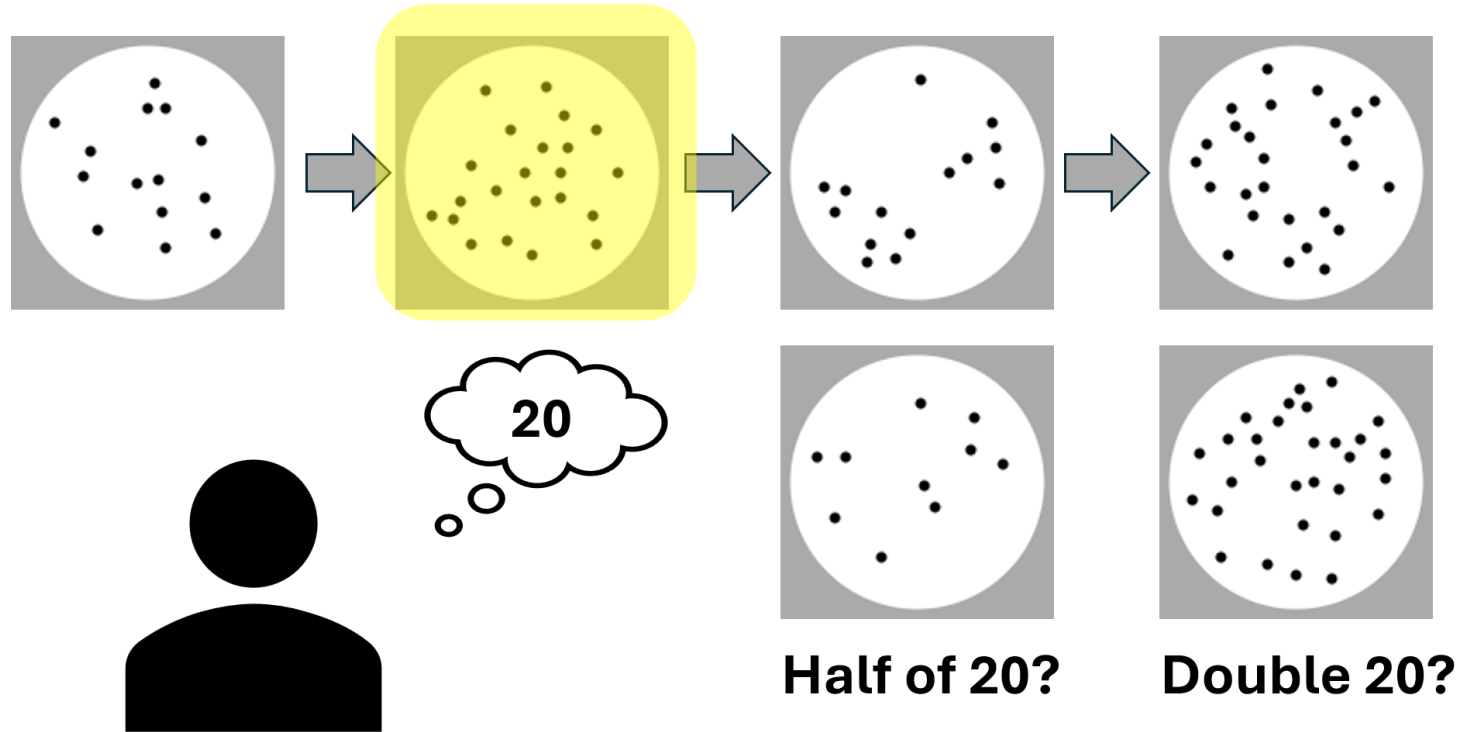
Other

Other

Other

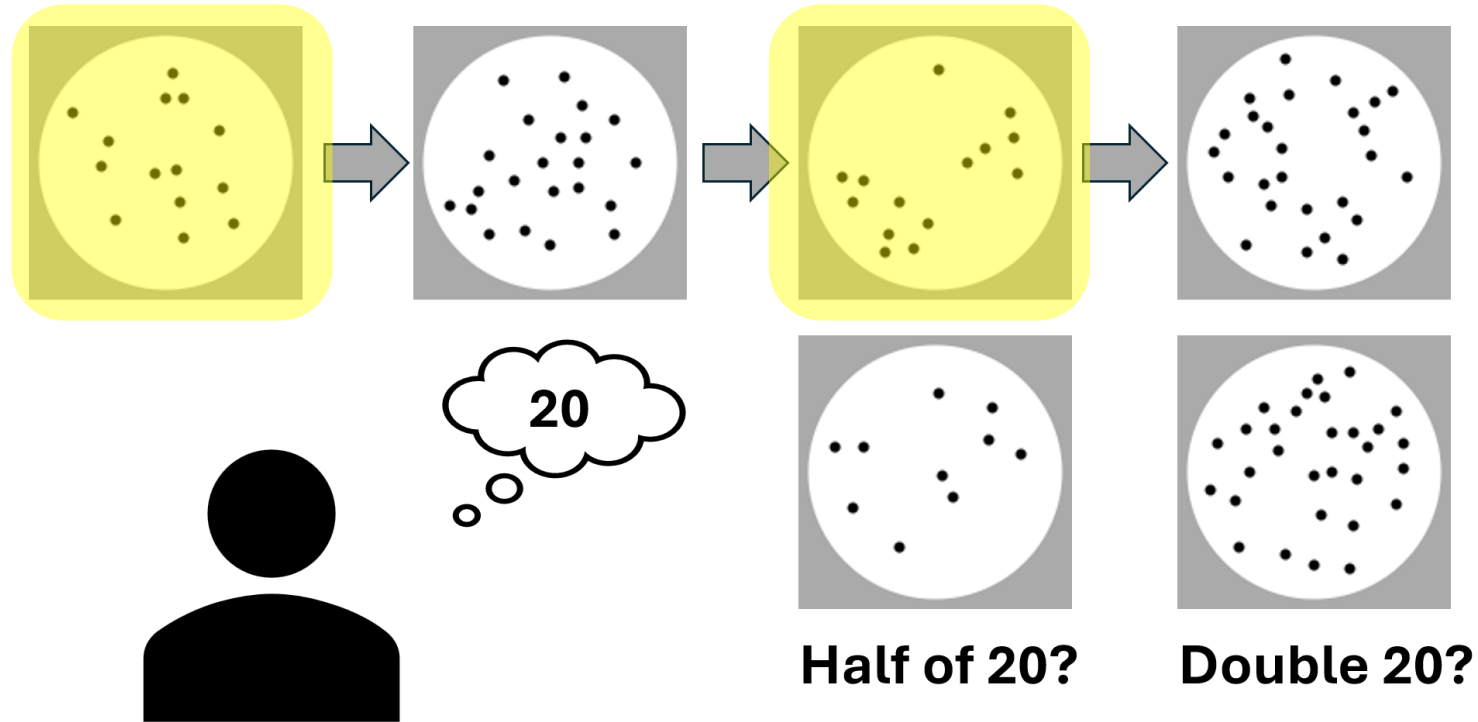
Idea for study 3

Even if people do use these memories, figuring out precisely how they do so is challenging:



Idea for study 3

Even if people do use these memories, figuring out precisely how they do so is challenging:

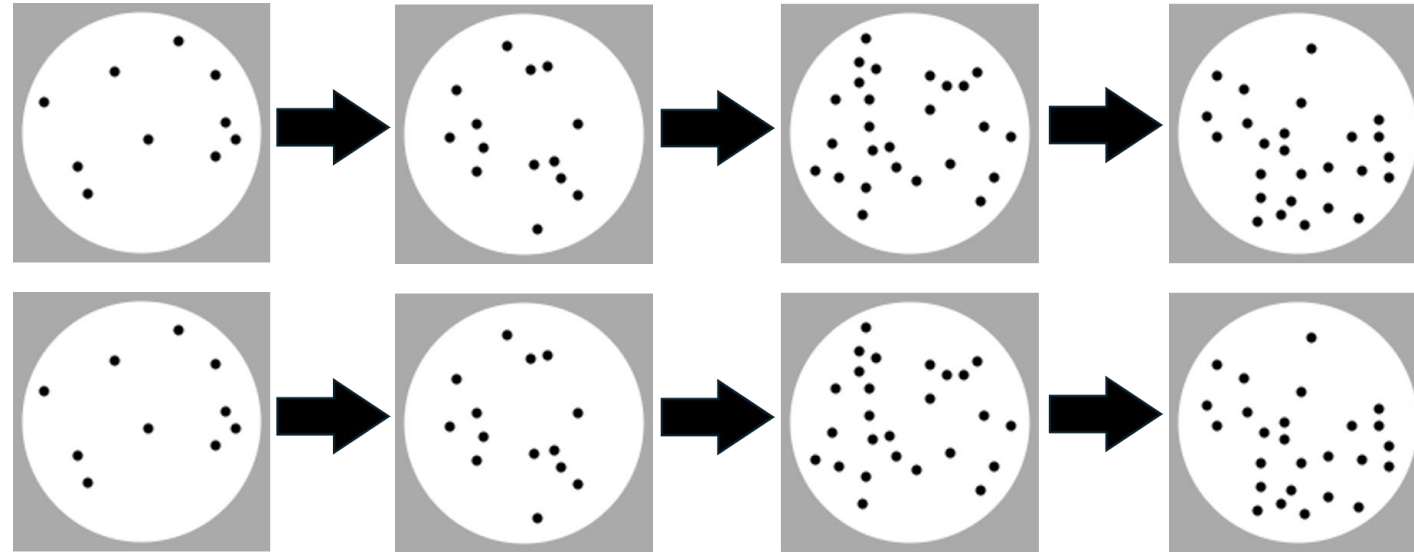
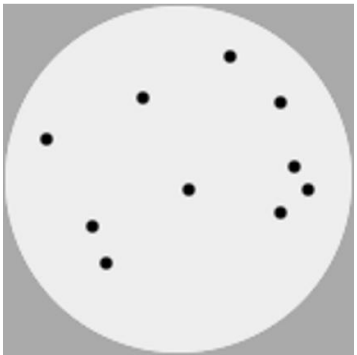


Which patterns people hold in mind, and how many of them, is expected to vary between participants....

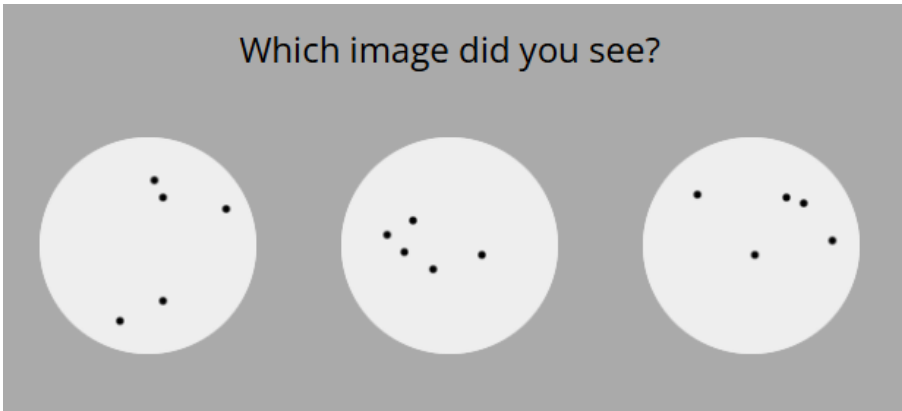
Idea for study 3

We could use reference images as a proxy for this:

This is 10:



Which image did you see?

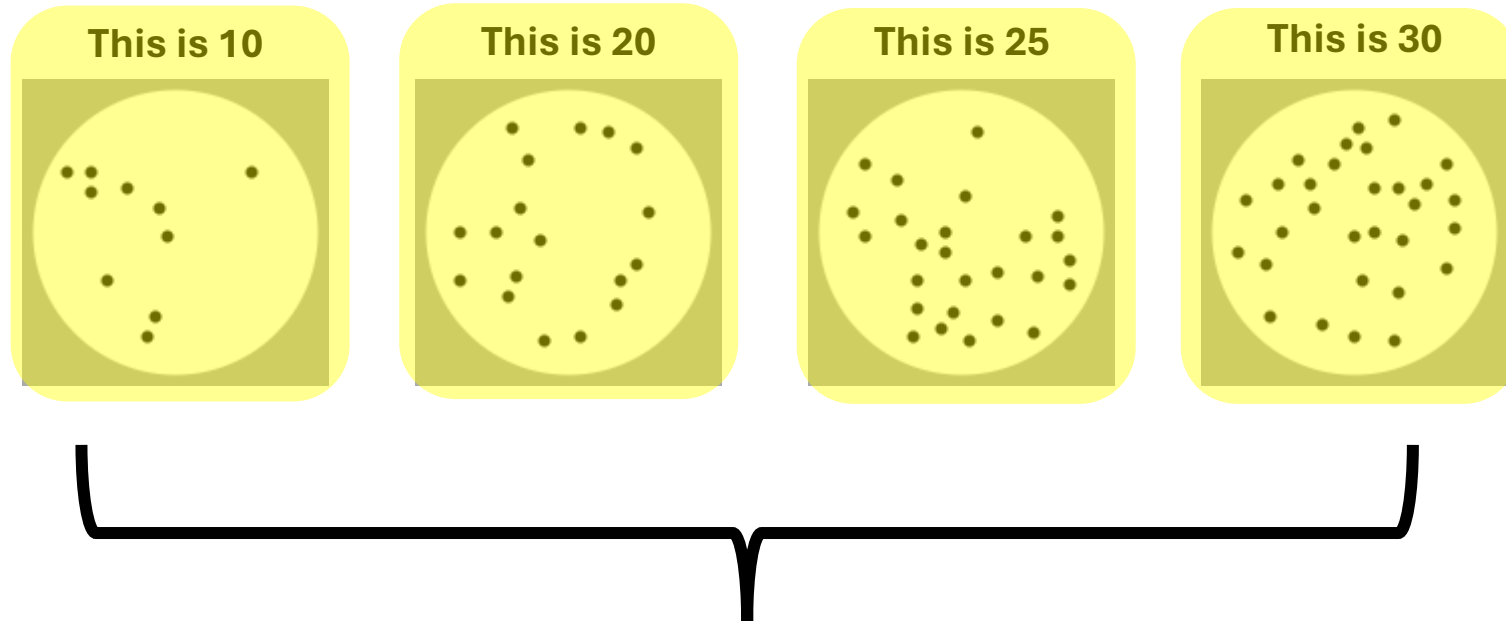


Predicted higher memory for reference patterns

We would only expect participants to be biased by the reference images when they actually remember them!

Idea for study 3

But people may hold in mind multiple patterns, and vary in their capacity:

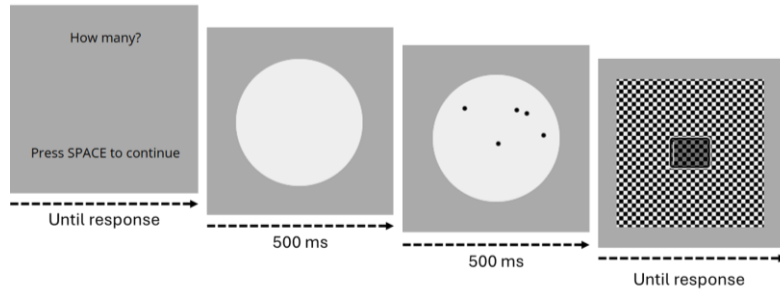


So, maybe it is this **capacity** rather than general memory accuracy that predicts performance, because the more patterns you hold in mind, the more reference points you have to refine your estimates

Other ideas & conclusion

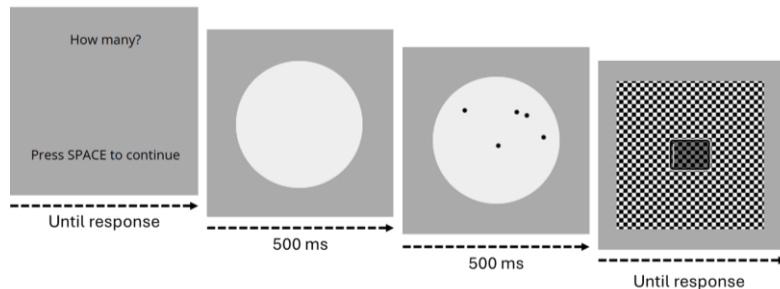
This memory approach provides an opportunity to investigate various things:

For example: intraindividual differences



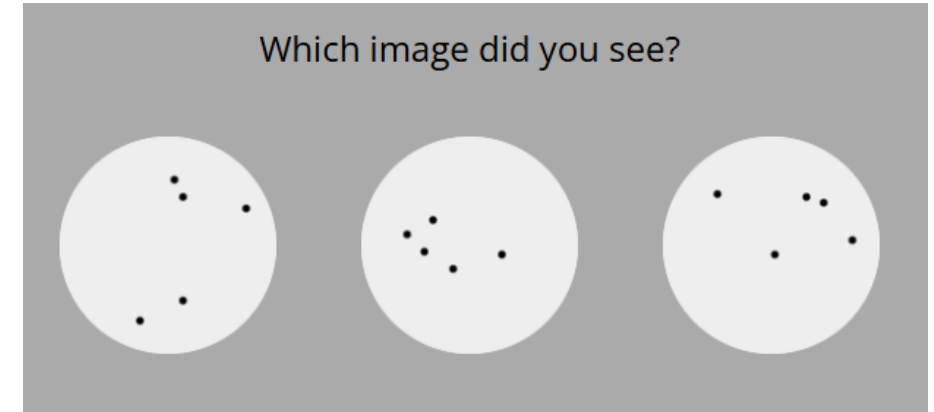
Serial dependence

Feedback effects



Serial dependence

Feedback effects



We cannot yet make claims about what people actually do during estimation. But...this methodological contribution provides exciting opportunities for the future! 😊

Takk for meg! 🙊

And thanks to my co-authors:



Natalia Dubinkina

Review - preprint



Trygve Solstad

Experimental studies



Eivind Kaspersen

declan.devlin@ntnu.no