Why be vague?

Kees van Deemter & Matt Green

University of Aberdeen

Scotland, UK

Going back in time a few years ..

- My book, *Not Exactly: In Praise of Vagueness*, Oxford University Press 2012
  - Theories of vagueness (logic, linguistics)
  - Why and where does vagueness arise? (e.g., how might V. be useful to hearers?)
  - Implications for Natural Language Generation
- A weakness: lack of empirical evidence
- This seminar explores recent work with Matt Green (psycholinguist at Aberdeen)

Plan of the talk

- Natural Language Generation (NLG)
- When do hearers benefit from vagueness?
- Experimental work
  - E.Peters et al.
  - Mishra et al.
  - Green & van Deemter

Plan of the talk

- Natural Language Generation (NLG)
- When do hearers benefit from vagueness?
- Experimental work
  - E.Peters et al.
  - Mishra et al.
  - Green & van Deemter
- Are we asking the wrong question?

Natural Language Generation (NLG)

- Goal: generate sentences/texts

  For practical use, or
  to model human language production

- Input:

  Some **non-linguistic** information (database records, logic formulas, program code, …)

Weather forecasting

- FOG system for weather forecasting (Goldberg 2000). Operational since 1992
- From 2013, a generator by Arria NLG has been used by the MET Office

Arria NLG 2014

06:00 Sun 16 Mar 2014 - 06:00 Mon 17 Mar 2014

"Sunny intervals throughout the day. Staying dry, but becoming cloudier from nightfall and into Monday. Highest temperatures expected during the afternoon in the region of 13C with a minimum temperature towards the morning of around 8C. Moderate westerly winds throughout."

http://www.metoffice.gov.uk/public/weather/forecast-data2text/

NLG systems expressing quantities

Weather forecasting

Input: numbers (20 Knots, 11PM)

Output: *"Winds light near midnight"*
(FoG system, Goldberg et al. 2000)

Medical decision support

Input: Time-series data on babies in IC

Output: *"about 0.3 litres, … very variable, …"*
(Babytalk system, Portet et al. 2012)

Vagueness

- Vagueness as defined by most logicians/philosophers/linguists:

  Predicate is vague if it has **borderline cases** (and borderline cases of borderline cases)

- Is it ever helpful to be vague?
- Should practical NLG systems use vague language?

Why is language vague?

Barton Lipman: Why have we tolerated an apparent “worldwide several-thousand year efficiency loss”?

In A.Rubinstein, *“Economics and Language” (*2000)

Lipman's scenario

Airport scenario: *I describe Mr X to you, to pick up X from the airport. All I know is X's height; heights are uniformly distributed across people on [0,1].*
If you identify X right away, you get payoff 1; if you don't, you get payoff -1

What description would work best?

- *State X's height "precisely"* $\Rightarrow$
  If each of us knows X's exact height then the probability of confusion is close to 0.

What description would work best?

- *State X's height "precisely"* $\Rightarrow$
  If each of us knows X's exact height then the probability of confusion is close to 0.

What description would work best?

- *State X's height "precisely"* ⇒
  If each of us knows X's exact height then the probability of confusion is close to 0.

Lipman: no boundary cases, hence not vague

Theorem: under standard game-theory assumptions (Crawford/Sobel), vague communication can never be optimal

so . . .

- When is vague communication more useful than crisp communication?
- **Strategic** vagueness. This is not about situations where the speaker has no choice (e.g., where no exact metric exists)
- Focus on **collaborative** situations

- Some possible factors

  Van Deemter 2009 *J. of Philosophical Logic* **38**/6.

1. Production/interpretation cost

- Example: planning a trip

  1. *The temperature is 23.75 C*
  2. *It's mild*

- (2) takes fewer syllables than (1)
- Precision of (1) adds little benefit

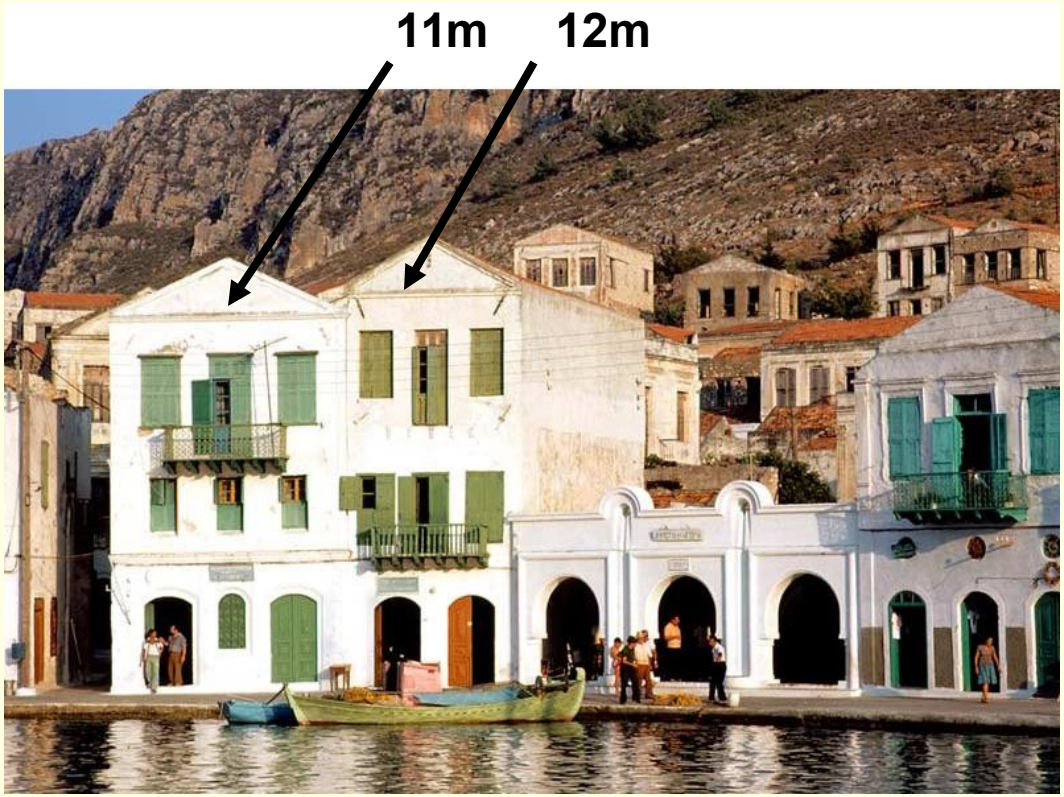
  Feasibility of an outing does not depend on whether it's 20C or 30C

2. Evaluation payoff

- Example: The doctor says
  1. *Your blood pressure is 153/92*
  2. *Your blood pressure is high*

- Version (2) offers less detail than (1)
- But (2) also offers **evaluation** of your condition (cf. Veltman 2000)
  - A link with actions (cut down on salt, etc.)
  - Crucial if metric is “difficult”
- But evaluation <> borderline cases

3. Comparing vs. matching



Comparing vs. matching

- Example: One house of 11m height and one house of 12m height
  1. *the house that's 12m tall needs to be demolished*
  2. *the tall house needs to be demolished*
- Comparison is easier and more reliable than measurement ➔ prefer utterance 2
- But comparison has nothing to do with borderline cases

4. Future contingencies

- Indecent Displays Control Act (1981) forbids display of ***indecent matter***
  - "indecent" at the time

  $\Rightarrow$ the law has been parameterised

  Waismann 1968, Hart 1994
- Again, this has nothing to do with vagueness

Experimental evidence

1. Peters et al. (2009)

Peters et al. (2009)

Hospital ratings based on numbers:

(1) survival %
(2) % recommended treatment
(3) patient satisfaction

"How attractive is this hospital to you?"

Peters et al. (2009)

- When labels ("fair", "good", "excellent") were added, a greater proportion of variance in evaluation judgments could be explained by the numeric factors

- Without labels,
  - the most important information (survival %) was not used at all
  - less numerate subjects were influenced by mood (*"I feel good/bad/happy/upset"*)

Peters et al. (2009)

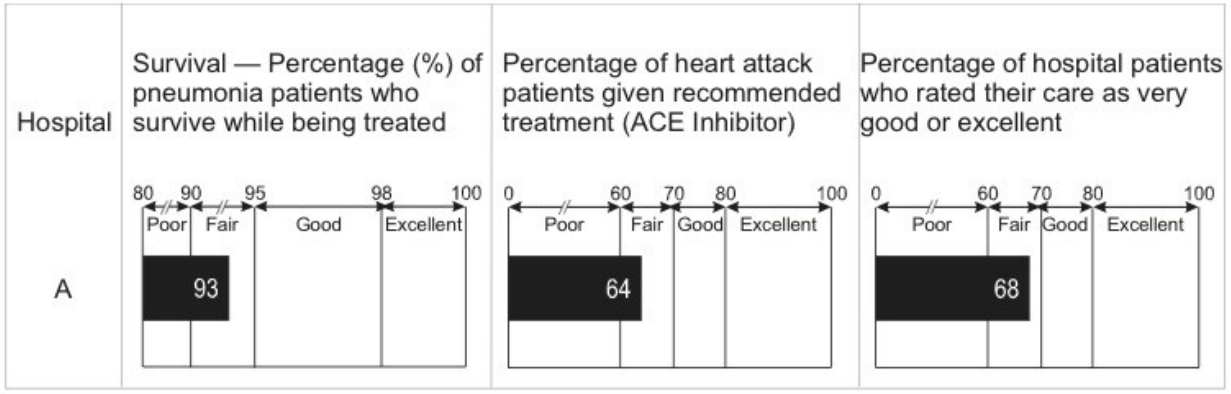
- This looks like a benefit from vague words ("fair", "good", "excellent")

Peters et al. (2009)

- This looks like a benefit from vague words ("fair", "good", "excellent")
- But …
  - The effect was caused by **evaluative** words
  - Nothing to do with borderline cases

  A vertical bar was used as a threshold

Peters et al. (2009): stimuli



Hospital	Survival — Percentage (%) of pneumonia patients who survive while being treated	Percentage of heart attack patients given recommended treatment (ACE Inhibitor)	Percentage of hospital patients who rated their care as very good or excellent
A	93	64	68

Figure 1. Study 1: Hospital information is presented in five-alternative-categories format (top) or with numbers only (bottom).

2. Mishra et al. 2011

- How does feedback affect behaviour?
- Subjects wanted to loose weight

Group 1 were given precise feedback:
BMI = x

Group 2 were given a range:
$x \leq BMI \leq y$

Algorithm a says BMI=x
Algorithm b says BMI=y

2. Mishra et al. 2011

- Results: more weight loss in **Group 2**
- Mishra's explanation: using a range allows subjects to feel optimistic about their progress
  - Wishful thinking is common (≥ Marks 1951)
  - Feeling near one's target helps performance

- Interesting!

2. Mishra et al. 2011

… but:

Vagueness didn't play a role:

- Mishra's range had no borderline cases
- Explanation didn't involve borderline cases
- The real issue was **low granularity** (cf. Hobbs 1985)

3. Green & van Deemter

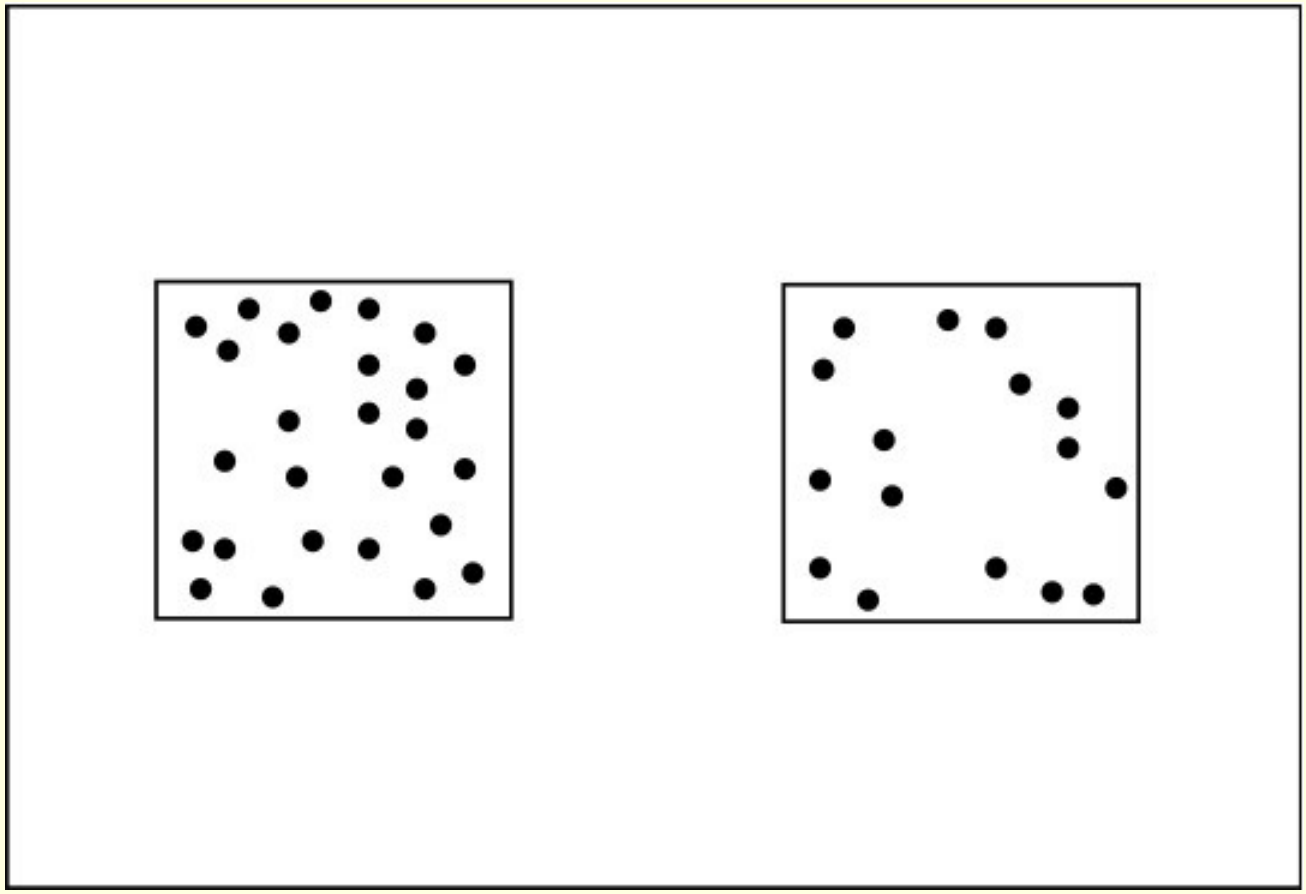
- Focus on referring expressions
- Hypothesis: interpretation cost is crucial

- Compare readers' Response Times, e.g.

  Choose the square with four dots

  Choose the square with many dots
- Are RTs smaller for vague instructions?

"Choose the square with … dots"



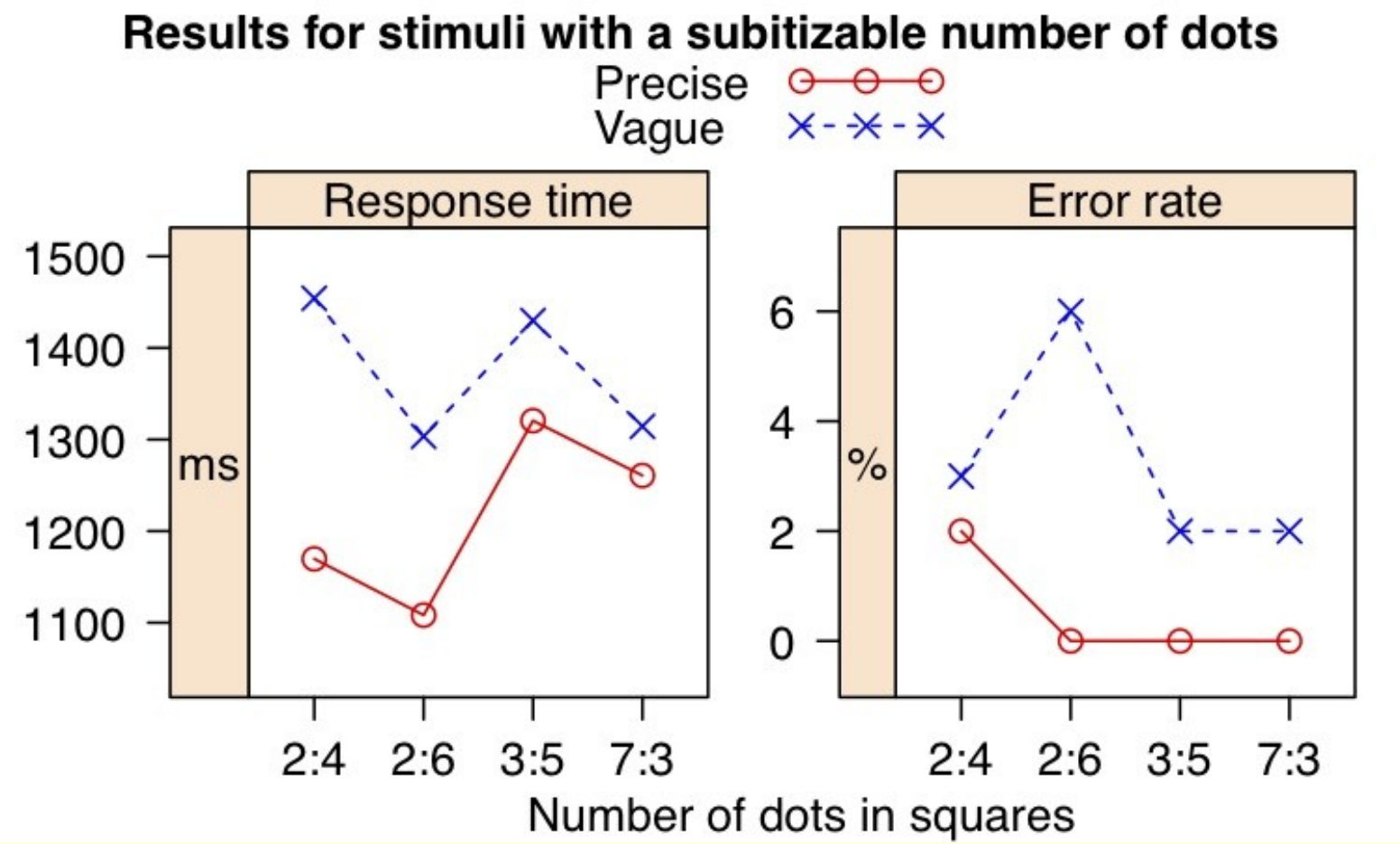
G&vD: Experiment 1

- No significant effect of vagueness
- **Subitisable** numbers followed the opposite pattern
- **Numbers below 5** play special role in visual perception, e.g.,

  Kaufman et al. (1949)

  Trick & Pylyshyn (1994)

G&vD: Experiment 1



G&vD: Experiment 2

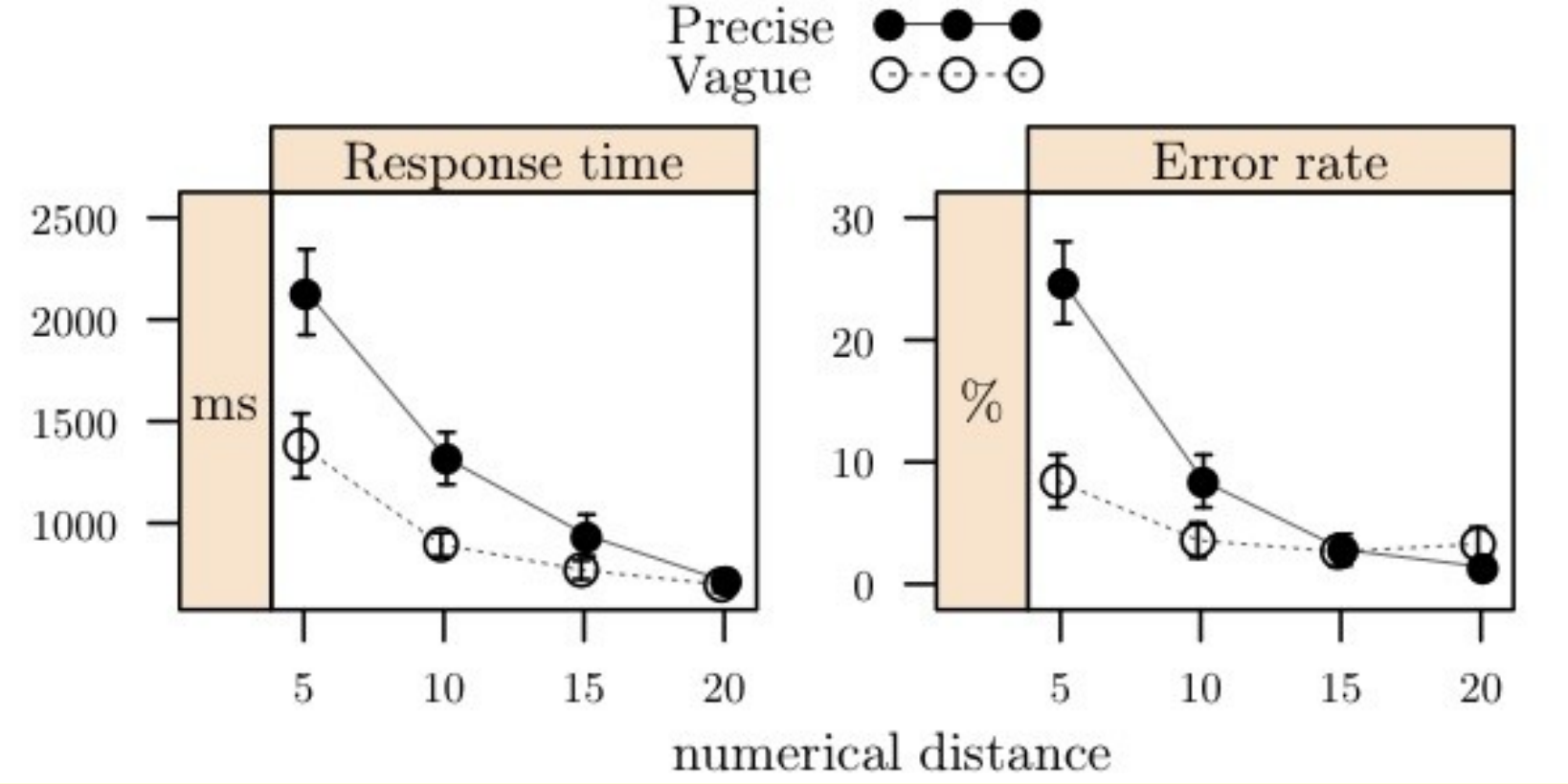
We varied

- the numbers of dots in the boxes: 5,10,15,20,(25),30,35,40,45
- The distance between the two numbers 5,10,15,20

All boxes were compared with

a box that contains **25** dots

G&vD: Experiment 2



G&vD: Experiment 2

- Vagueness helps for larger numbers
- Diminishing advantage for vagueness as gap size grows large
  - Subjects are able to pick *"the square with **45** dots"* without counting
  - (Interesting in its own right: an expression may be precise, yet interpreted almost as if it were vague!)

G&vD: Experiment 2

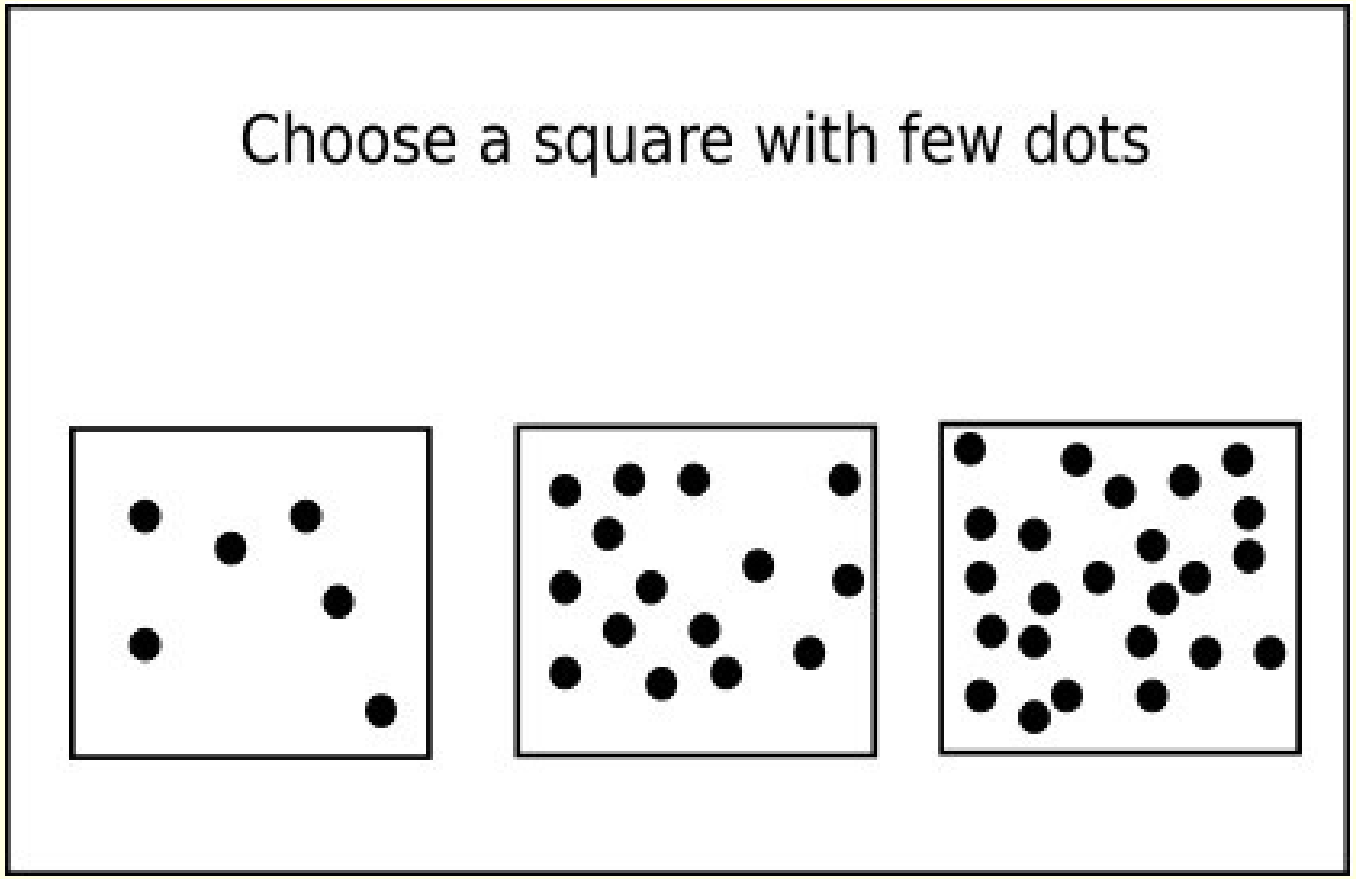
- Vagueness helps when subitisable numbers are excluded
- Diminishing advantage for vagueness as gap size grows large

- These were encouraging results, but …

Problems with Experiment 2

- Potential for vagueness not realised?
  - Two squares → no borderline case
  - Definite NPs (*"the square with .."*) identify the target uniquely
- Our solution for the next experiment:
  - use > 2 squares
  - use indefinite articles (*"a square with .."*)

G&vD: Experiment 3



Other problems with Experiment 2

- Vagueness confounded with absence of numbers?

  Vague: *few, many*
  Precise: *5, 25*

- Solution: factorial instruction format (2 x 2)

	Precise	**Vague**
Numerical	…	…
Verbal	…	…

Example: The triple (16,25,34)

	Precise	**Vague**
Numerical	*16 dots*	*about 20 dots*
Verbal	*the fewest dots*	*few dots*

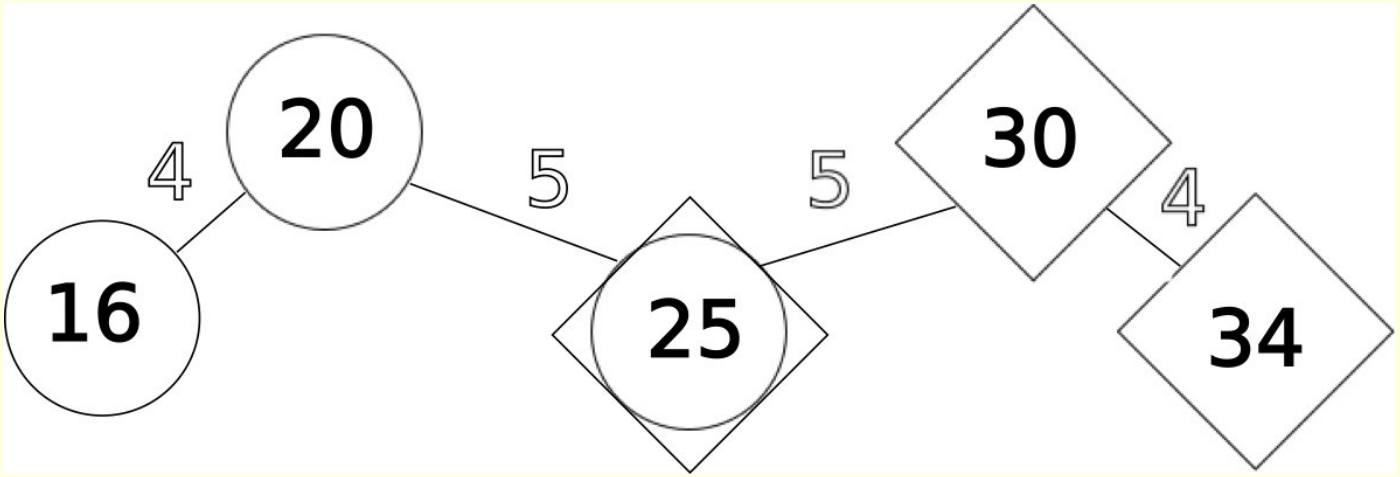
"a square with about 20 dots":

a clear case: 16 dots

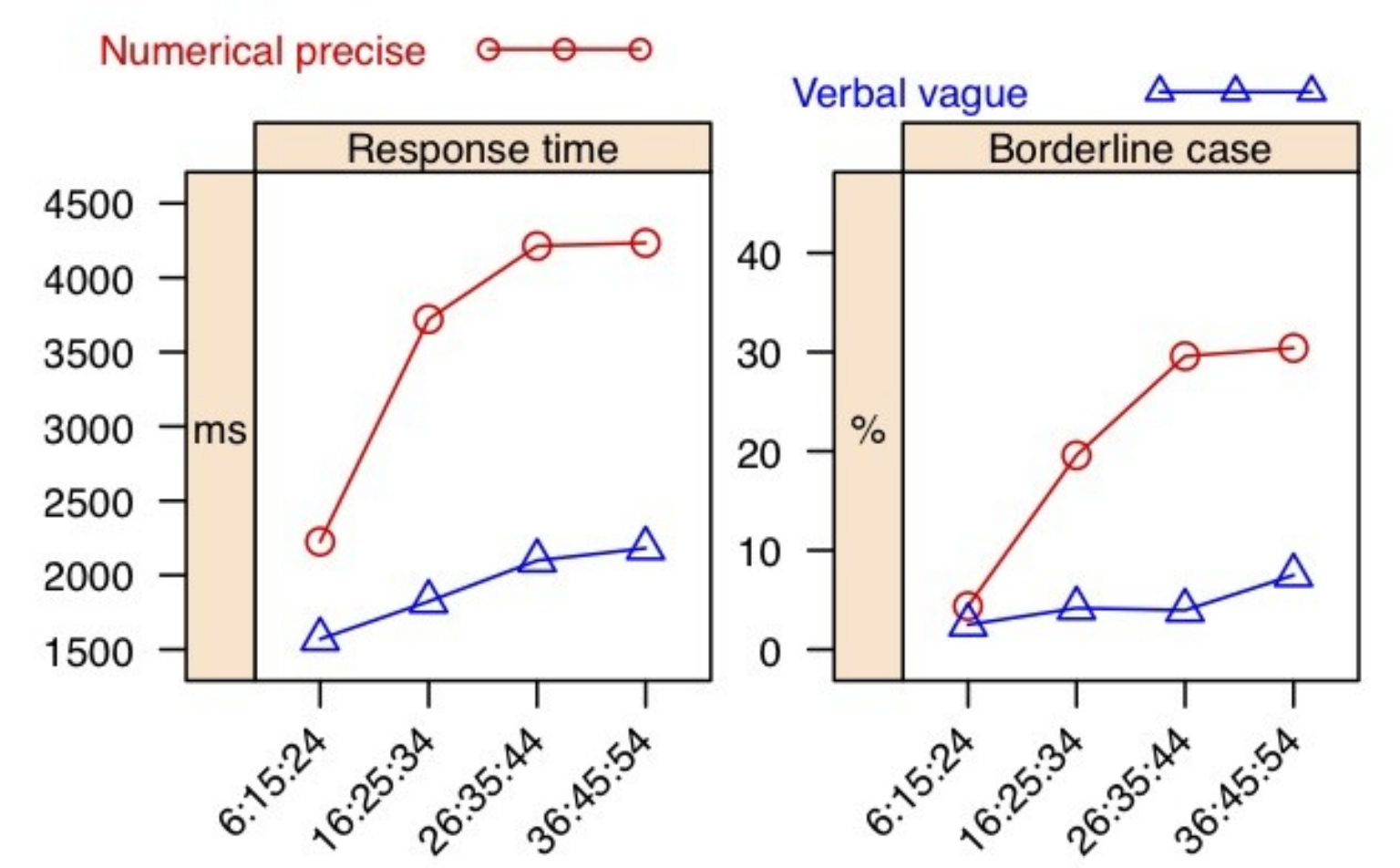
borderline case: 25 dots

Example: the triple (16,25,34)

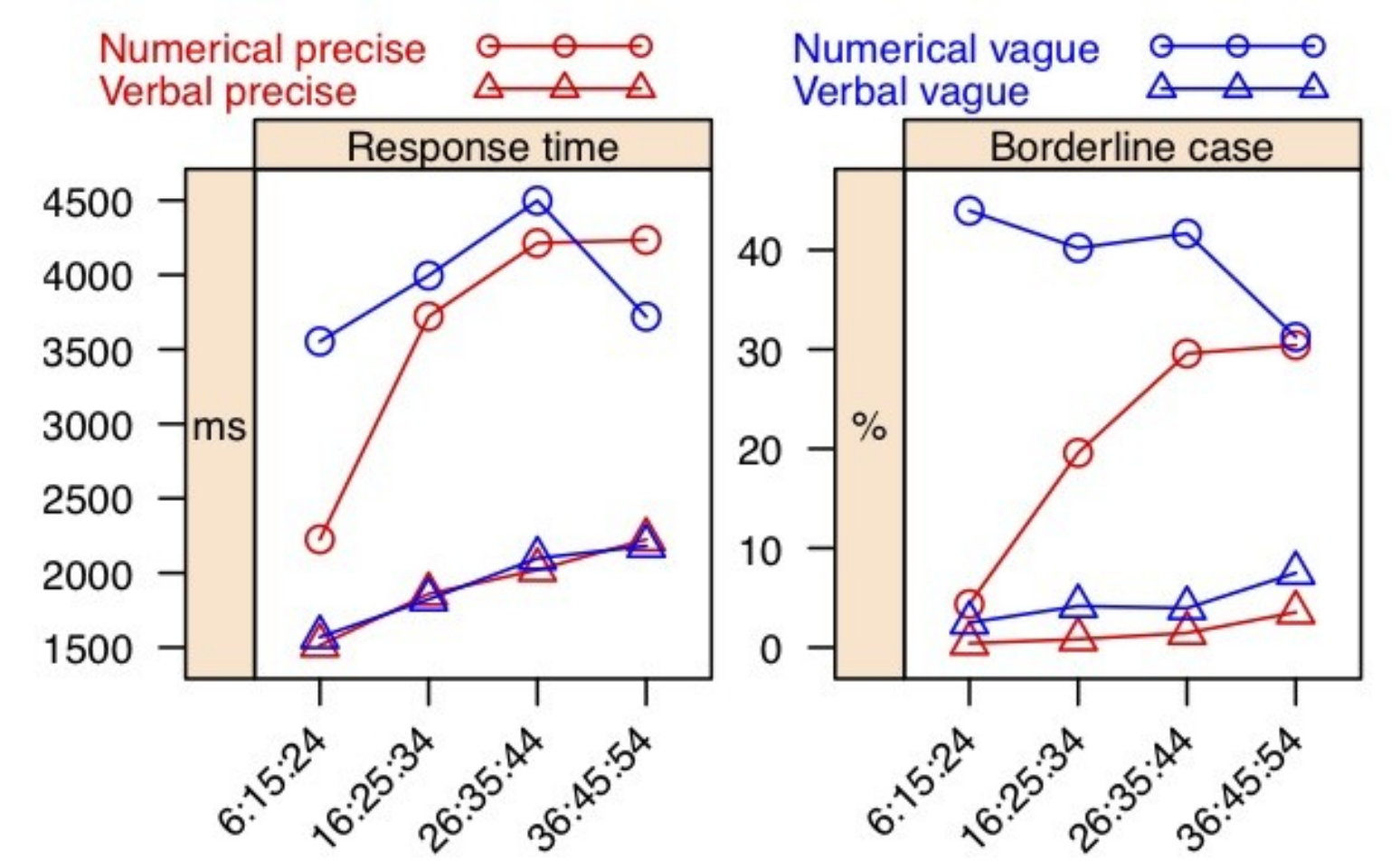
Vague & Numerical: *"about 20"* and "*about 30*"



G&vD: Experiment 3 (1)



G&vD: Experiment 3 (3)



G&vD: Experiment 3

- **Vague** expressions were not reliably faster than Non-vague ones (p=0.73)
- **Numerical** expressions were much slower than non-numerical ones
- No interaction between vagueness and numericity

What this suggests

- Perhaps the benefits of vague words are not about vagueness but about number avoidance
- Or maybe it's not about the presence of numbers *per se*, but about the existence of comparison strategies
- Look at an example from Experiment 3 again:

Example: The triple (16,25,34)

	Precise	**Vague**
Numerical	*16 dots*	*about 20 dots*
Verbal	*the fewest dots*	*few dots*

Both verbal items rely on a comparison task

Both numerical items rely on a matching task

Two final experiments

a four types of REs that contain numbers

b four types of REs that do not contain numbers

In both cases, a 2 x 2 design:

precision x comparison

	Precise	Vague
Comp		
Match		

New experiment **a** (focussing on **Numerical** expressions)

	Precise	**Vague**
Comp	*fewer than 30*	*far fewer than 30*
Match	*16*	*about 20*

- Even though the two **Comp** items contain a number, the task can be performed by finding the smallest number

New experiment **b** (focussing on **Non-numerical** expressions)

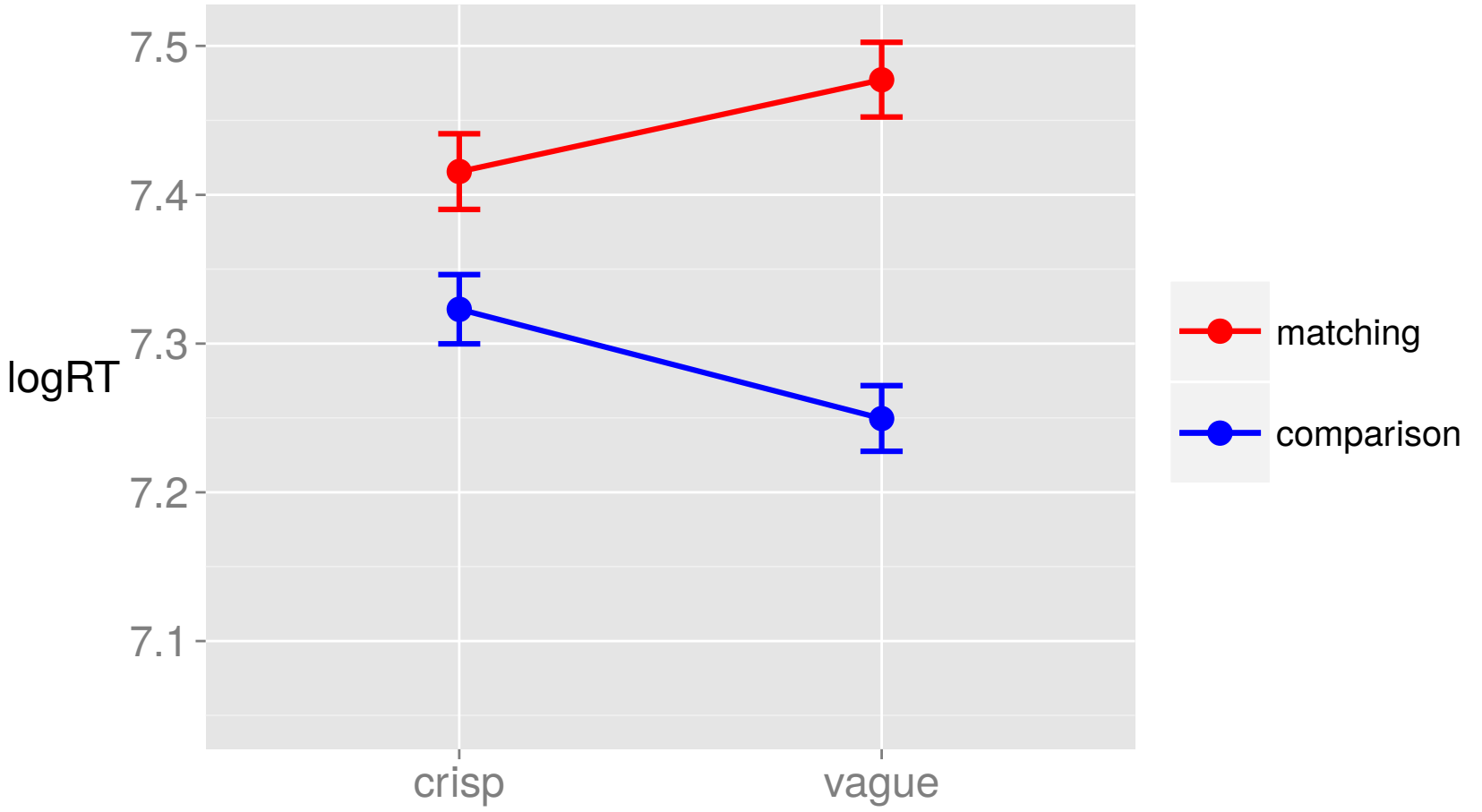
	Precise	**Vague**
Comp	*fewer than X*	*far fewer than X*
Match	*same number as X*	*approx same number as X*

- *X* has been shown a few seconds earlier

G&vD: Experiment 4

1. Instructions
   - *Choose a square with 6 dots*
2. Show squares
   - 3 squares with 24, 15, 6 dots
3. Collect response

G&vD: Experiment 4



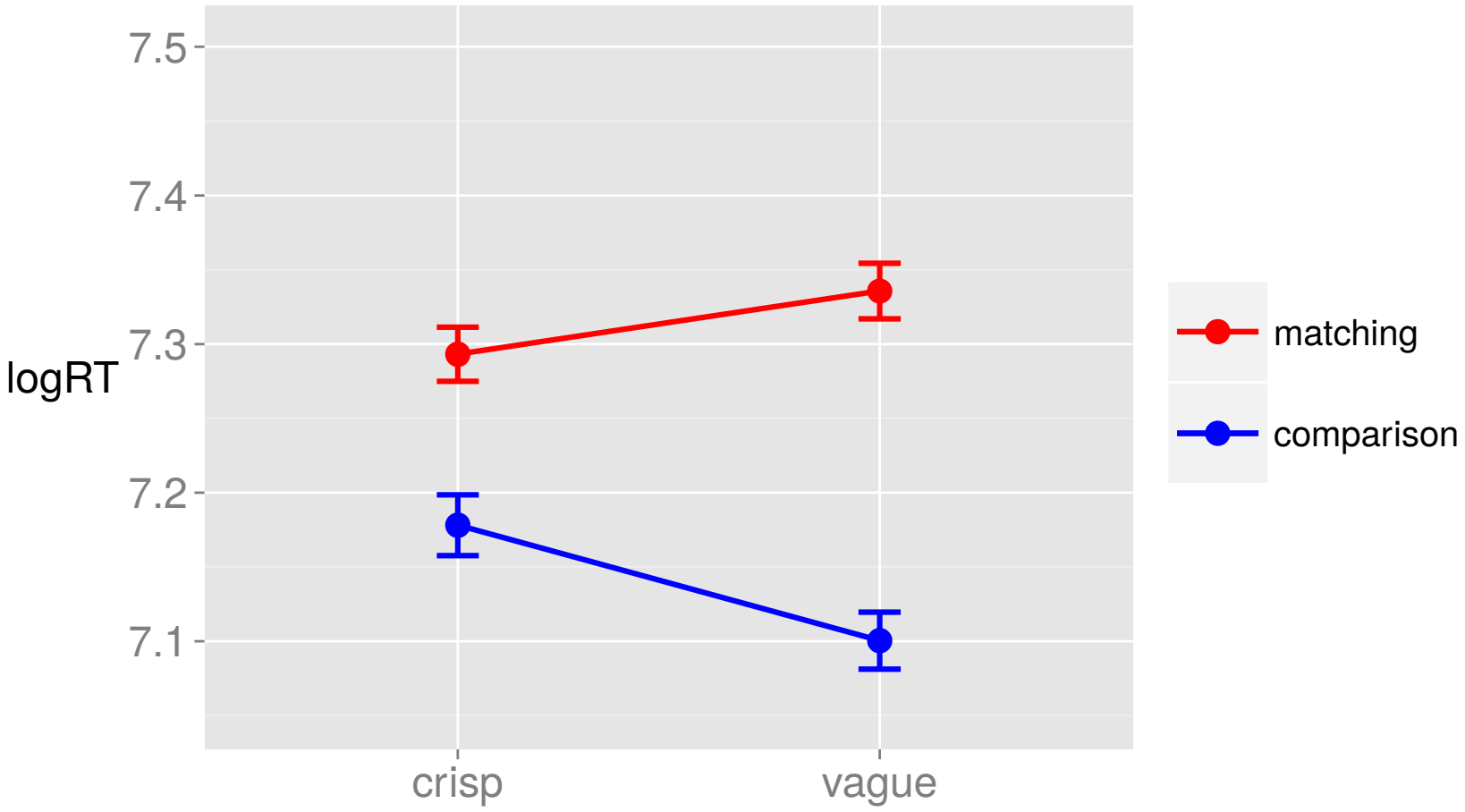
G&vD: Experiment 4

- Results show
  - A benefit for comparison
  - A benefit for vagueness in comparison, but a disadvantage of vagueness in matching

G&vD: Experiment 5

1. Instructions
   - *Choose a square with the same number of dots as the target*
2. Show target
   - a square with 6 dots
3. Show candidates
   - 3 squares with 24, 15, 6 dots
4. Collect response

G&vD: Experiment 5



G&vD: Experiment 5

- Results show
  - The task took slightly longer overall than the task in e4 (perhaps due to the effort of visual memory)
  - A benefit for comparison, as before
  - A benefit for vagueness in comparison, but a disadvantage of vagueness in matching, as before

A possible explanation: range reduction

- Compare is faster than Match because, for Compare, choosing an extreme value (in the right direction) always works. No counting is needed.
- Having more items to choose between makes choosing harder. This predicts the reversal:
  - *"Far fewer than thirty"* [Compare, Vague] is easier than *"Fewer than thirty"* [Compare, Crisp]
  - *"Sixteen"* [Match, Precise] is easier than *"About twenty"* [Match, Vague]

Maybe …

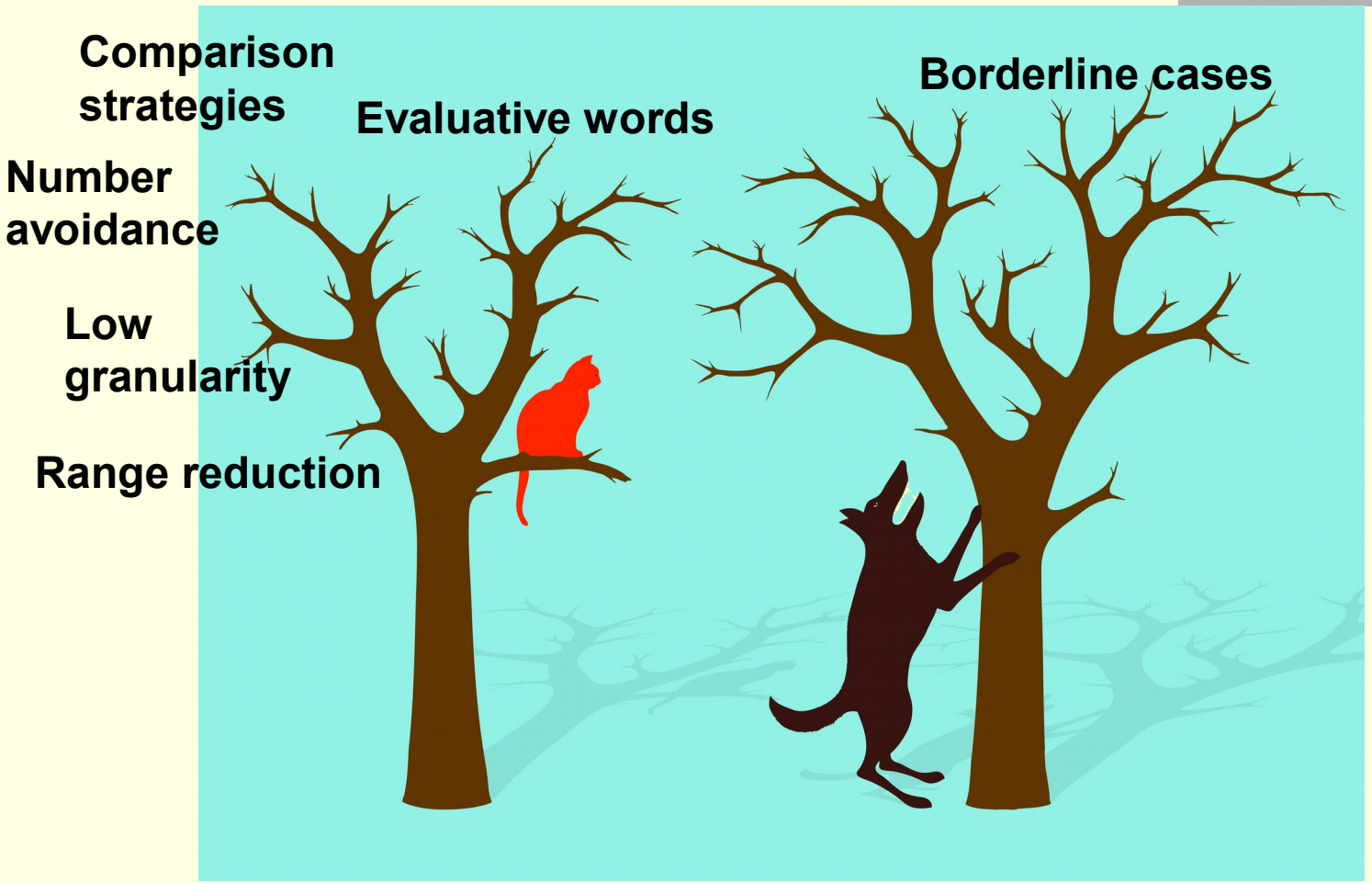
Vague expressions are only better than crisp ones because they **tend** to …

- express value judgments (Peters et al.)
- have low granularity (Mishra et al.)
- avoid numbers (Green & van Deemter)
- allow comparison strategies
- lead to range reduction for reference

Are we barking up the wrong tree?

Are we barking up the wrong tree?

Are we barking up the wrong tree?



Caveat

- Vagueness may not always work in the same way. Compare
  - Instructing a hearer to choose a box
  - Reporting on an experience
  - Asking a question
  - Etc.
- Another reason not to jump to conclusions!

Questions?

- M.Green & K.van Deemter (2011). Vagueness as Cost Reduction: an Empirical Test. In Proceedings of the workshop *Production of Referring Expressions*, CogSci 33.
- K.van Deemter (2012) *Not Exactly: In Praise Of Vagueness.* Oxford University Press.