The utility of vagueness

an empirical test of the cost reduction hypothesis in a reference task

Matt Green & Kees van Deemter

NLG group
University of Aberdeen

7th June 2011

Outline

1. Rationale (what, why, how?)

Outline

1. Rationale (what, why, how?)
2. Experiment one (tests the water)

Outline

1. Rationale (what, why, how?)
2. Experiment one (tests the water)
3. Experiment two (extends the findings)

Outline

1. Rationale (what, why, how?)
2. Experiment one (tests the water)
3. Experiment two (extends the findings)
4. Interim summary (where do we stand now?)

Outline

1. Rationale (what, why, how?)
2. Experiment one (tests the water)
3. Experiment two (extends the findings)
4. Interim summary (where do we stand now?)
5. Experiment three (addresses problems with E1 & E2)

Outline

1. Rationale (what, why, how?)
2. Experiment one (tests the water)
3. Experiment two (extends the findings)
4. Interim summary (where do we stand now?)
5. Experiment three (addresses problems with E1 & E2)
6. Interpretation (what does it all mean?)

A definition of vagueness

Vagueness can be defined as the existence of borderline cases

- e.g., *tall*

Why study the effects of vagueness?

2 main reasons

1. Game theory needs an update
   - Game-theory predicts that precision always 'wins'
   - But we use vague language a lot
   - There must be some utility of vagueness that the models don't account for

Why study the effects of vagueness?

2 main reasons

1. Game theory needs an update
   - Game-theory predicts that precision always 'wins'
   - But we use vague language a lot
   - There must be some utility of vagueness that the models don't account for
2. NLG systems need help
   - Practical NLG systems must decide between generating vague and precise terms (e.g., Portet et al., 2009)
   - These decisions are (almost) guesswork currently
   - Data from human comprehenders would provide a better basis for these decisions
   - There isn't much data like this at the moment

Example NLG decision

Which of these is better?

1. 'Winds southwest 15 to 25 knots **becoming southeast 15** later this evening'
2. 'Winds southwest 15 to 25 knots **diminishing to light** later this evening'

Example NLG decision

Which of these is better?

1. 'Winds southwest 15 to 25 knots **becoming southeast 15** later this evening'
2. 'Winds southwest 15 to 25 knots **diminishing to light** later this evening'

- The answer might depend on the audience: offshore oil rig, mountain resuce, flying doctor, air traffic control ...

Example NLG decision

Which of these is better?

1. 'Winds southwest 15 to 25 knots **becoming southeast 15** later this evening'
2. 'Winds southwest 15 to 25 knots **diminishing to light** later this evening'

- The answer might depend on the audience: offshore oil rig, mountain resuce, flying doctor, air traffic control ...
- It depends quite a lot what 'better' means: speed, accuracy, recall, comprehension, effect on affect ...

Vagueness in reference

- Vagueness in reference is highly context-dependent
- Potential for vagueness is not necessarily realised

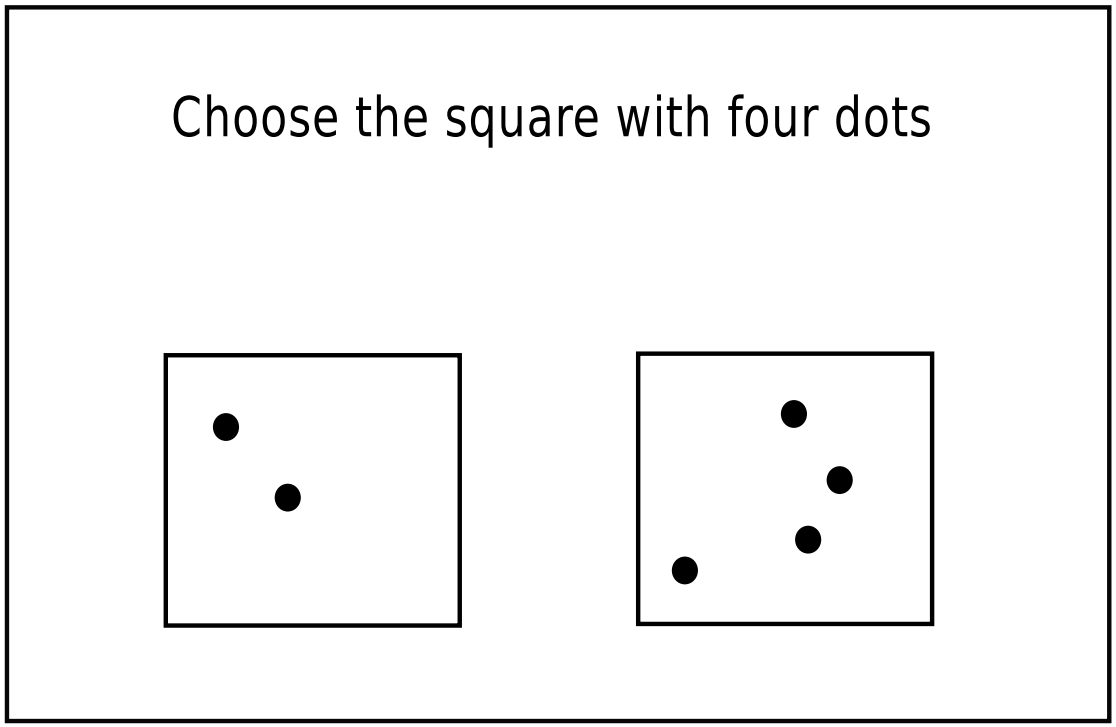
Pinning utility down

- Utility in co-operative situations is our focus
- Non-co-operative situations where vagueness has utility
  - Sales, marketing, advertising ✗
  - Politics, campaigning ✗
  - Persuasion, behaviour modification ✗
- Co-operative situations where vagueness may have utility
  - Local vagueness ✗
  - Information reduction ✗
  - Reducing cognitive load on the comprehender ✓

The cost reduction hypothesis

"For the listener, information which is too specific may require more effort to analyze" (Lipman, 2009)

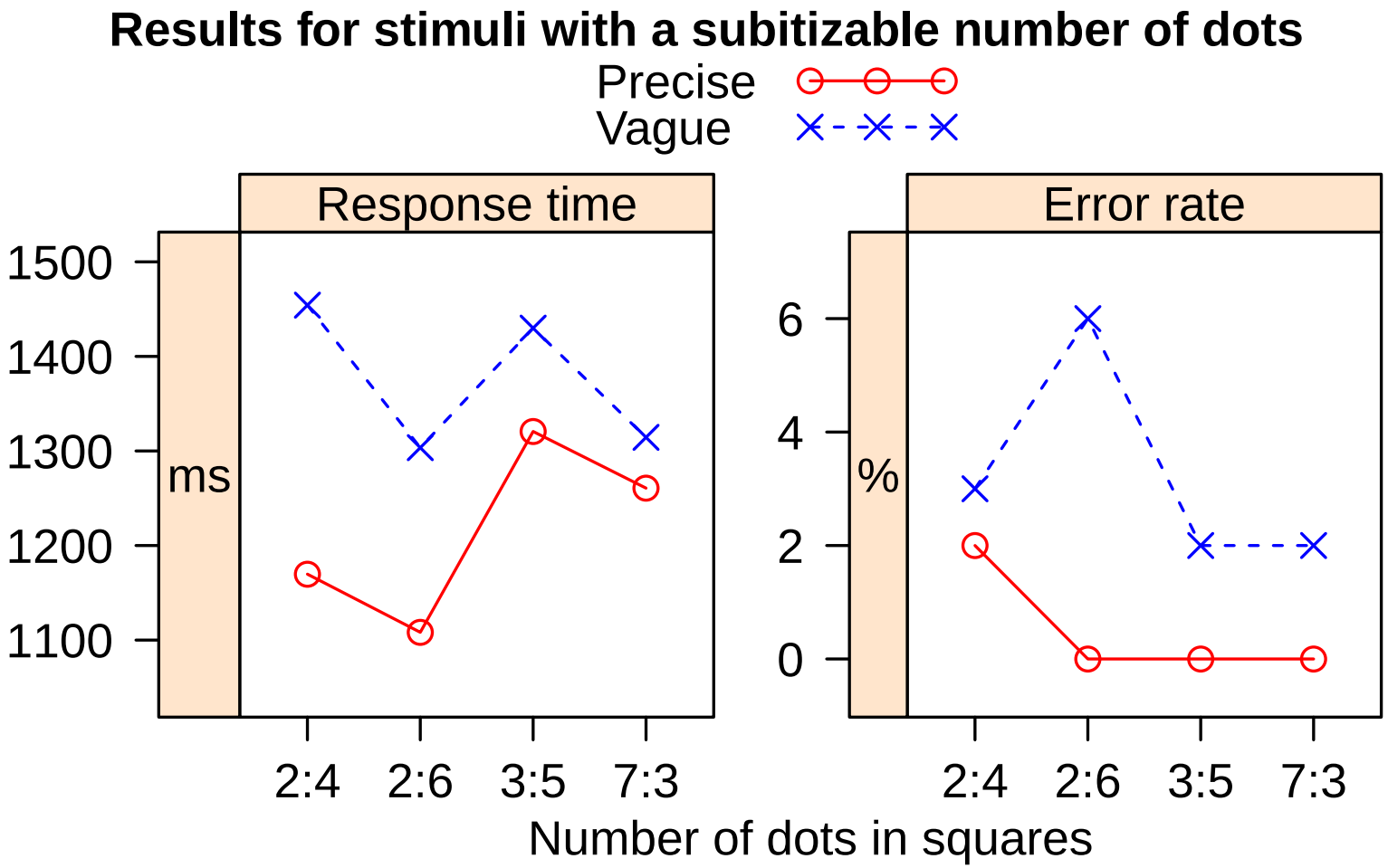
Experiment one



Experiment one

- Pilot experiment
- H_0: *No difference between vague and precise*
- H_1: *Some difference between vague and precise*
  . . . V better than P?
- Method: forced choice
- DVs: RT, errors
- IVs: Vagueness (Vague, Precise)
- Number combinations: $\{2, 4\}$, $\{2, 6\}$, $\{3, 5\}$, $\{5, 9\}$, $\{6, 8\}$, $\{7, 3\}$, $\{7, 9\}$, $\{8, 4\}$

Results, experiment one



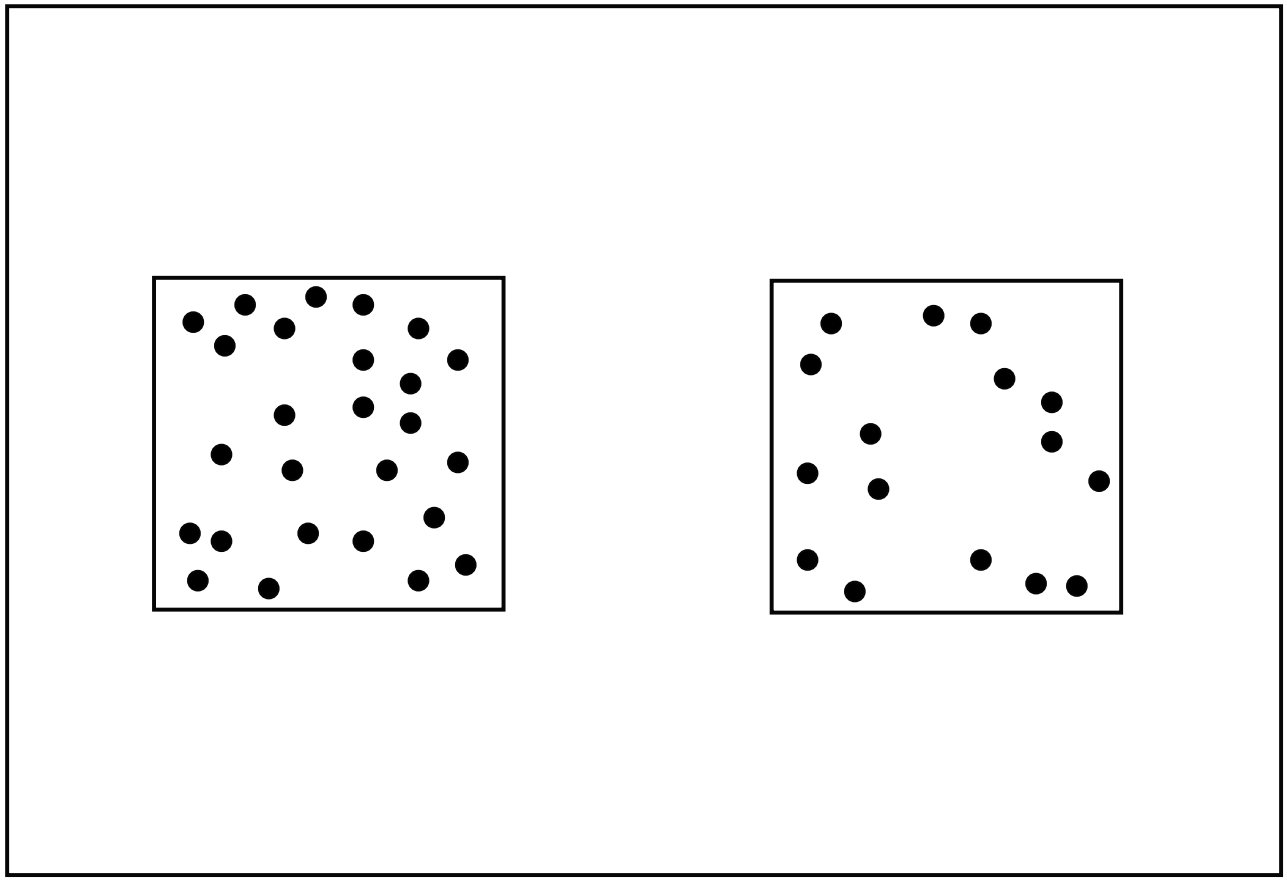
Experiment two

- Try larger numbers (not subitizable)
- H_0: *No difference between vague and precise*
- H_1: *Vagueness helps when the numbers are bigger, and when the difference between the numbers is greater*
- Number combinations: $\{5, 25\}$ $\{10, 25\}$ $\{15, 25\}$ $\{20, 25\}$ $\{30, 25\}$ $\{35, 25\}$ $\{40, 25\}$

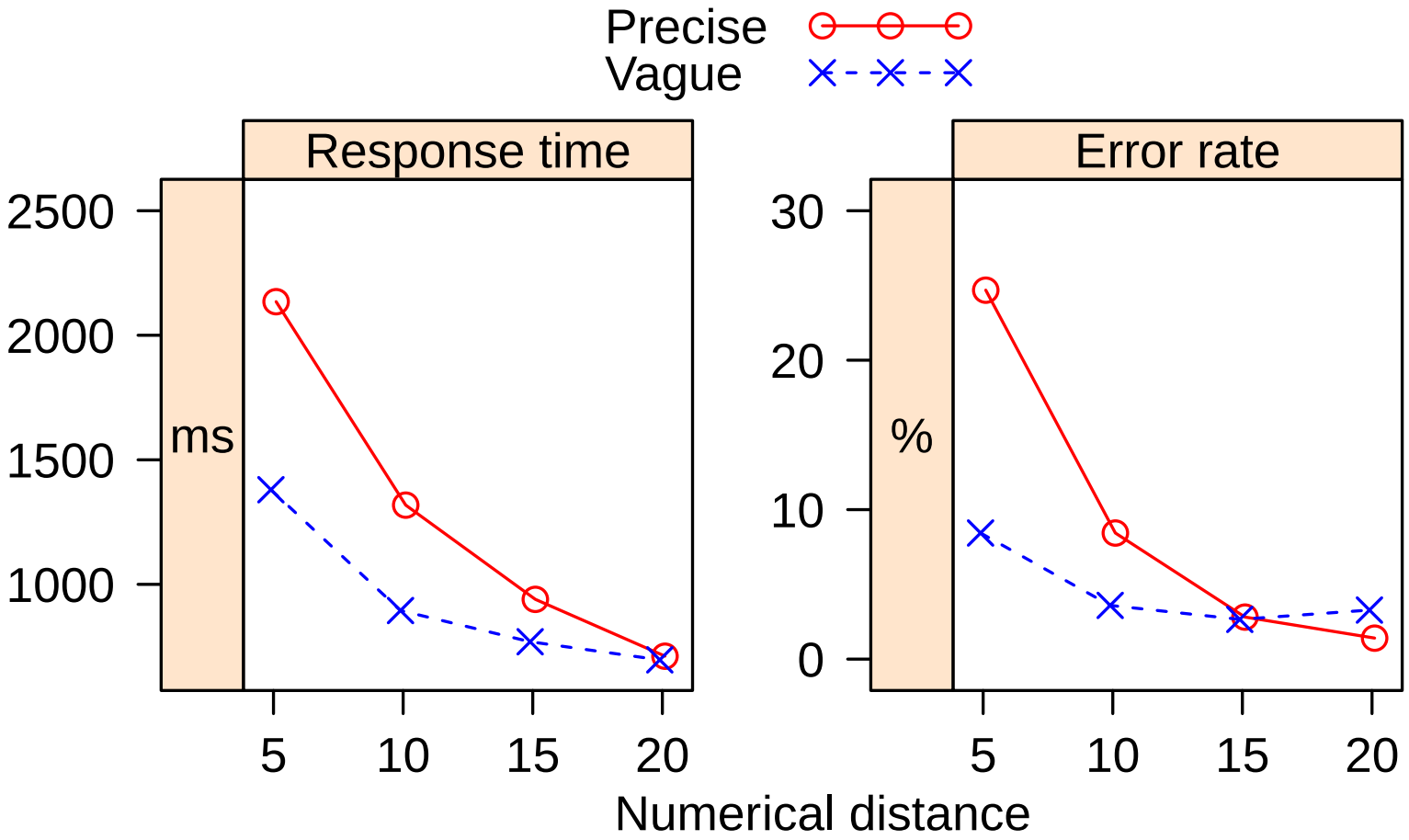
Instruction first

Choose the square with many dots

Stimulus next



Results, experiment two



Interim summary (where do we stand?)

- Vagueness is worse for very small numbers (expt 1)

Interim summary (where do we stand?)

- Vagueness is worse for very small numbers (expt 1)
- Vagueness is better for larger numbers (expt 2)

Interim summary (where do we stand?)

- Vagueness is worse for very small numbers (expt 1)
- Vagueness is better for larger numbers (expt 2)
- Diminishing returns for vagueness as gap size grows very large (expt 2)

Issues to address

- Potential for vagueness not realised? (expts 1 & 2)

Issues to address

- Potential for vagueness not realised? (expts 1 & 2)
  - Definite articles uniquely identify

Issues to address

- Potential for vagueness not realised? (expts 1 & 2)
  - Definite articles uniquely identify
  - Two squares means no borderline case

Issues to address

- Potential for vagueness not realised? (expts 1 & 2)
  - Definite articles uniquely identify
  - Two squares means no borderline case
  - Solution: use indefinite articles; use > 2 squares

Issues to address

- Potential for vagueness not realised? (expts 1 & 2)
  - Definite articles uniquely identify
  - Two squares means no borderline case
  - Solution: use indefinite articles; use > 2 squares
- Vagueness confounded with absence of a number in the instructions?

Issues to address

- Potential for vagueness not realised? (expts 1 & 2)
  - Definite articles uniquely identify
  - Two squares means no borderline case
  - Solution: use indefinite articles; use > 2 squares
- Vagueness confounded with absence of a number in the instructions?
  - V: {few, many}; P: {5, twenty}

Issues to address

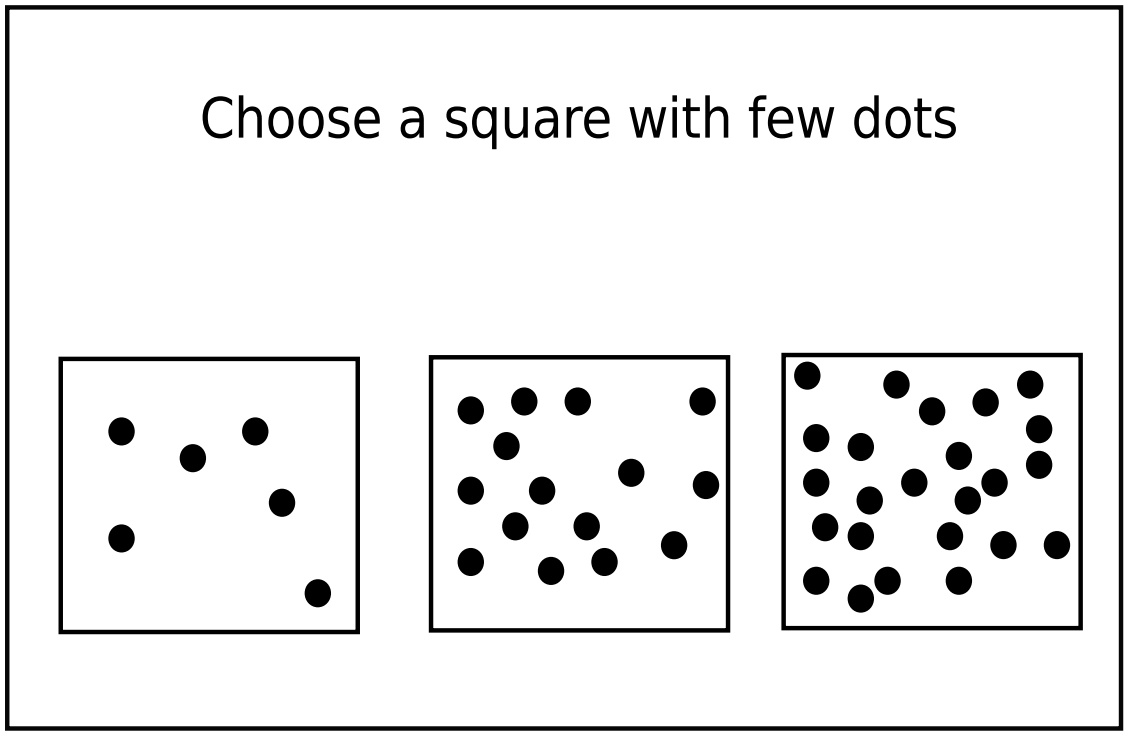
- Potential for vagueness not realised? (expts 1 & 2)
  - Definite articles uniquely identify
  - Two squares means no borderline case
  - Solution: use indefinite articles; use > 2 squares
- Vagueness confounded with absence of a number in the instructions?
  - V: {few, many}; P: {5, twenty}
  - Solution: factorially manipulate instruction format 2 x 2

Issues to address

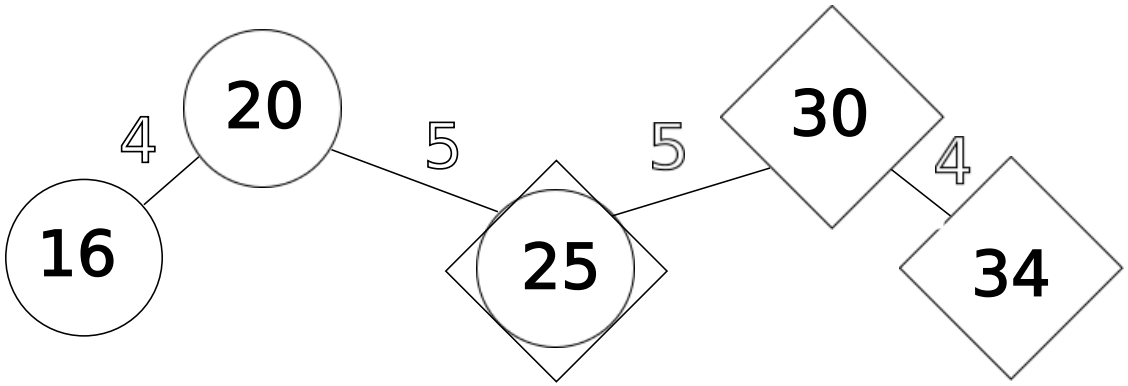
- Potential for vagueness not realised? (expts 1 & 2)
  - Definite articles uniquely identify
  - Two squares means no borderline case
  - Solution: use indefinite articles; use > 2 squares
- Vagueness confounded with absence of a number in the instructions?
  - V: {few, many}; P: {5, twenty}
  - Solution: factorially manipulate instruction format 2 x 2

	Vague	Precise
Numerical	{about 20, about 30}	{16, 34}
Verbal	{few, many}	{fewest, most}

Stimulus, experiment three



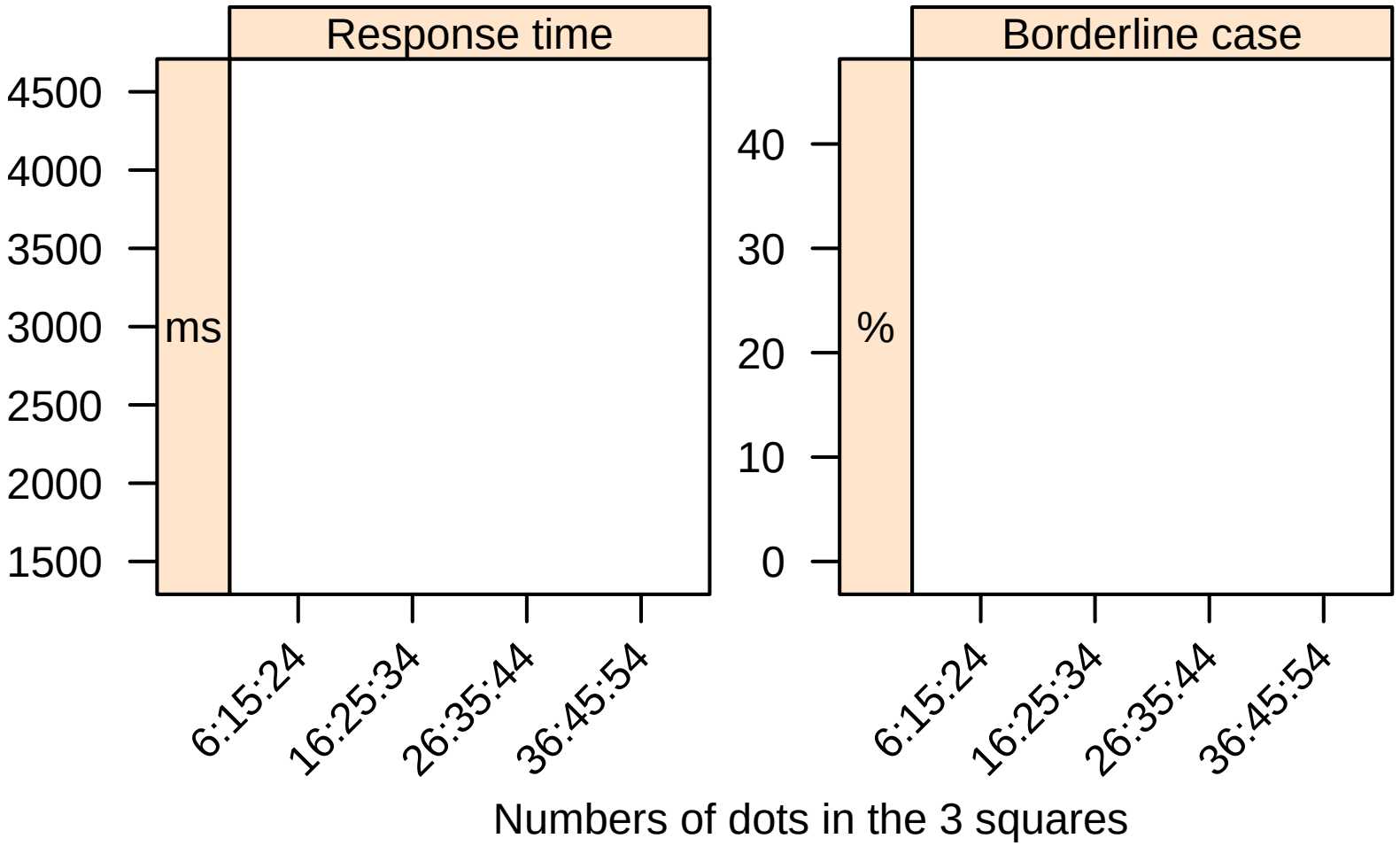
Operationalising borderline cases



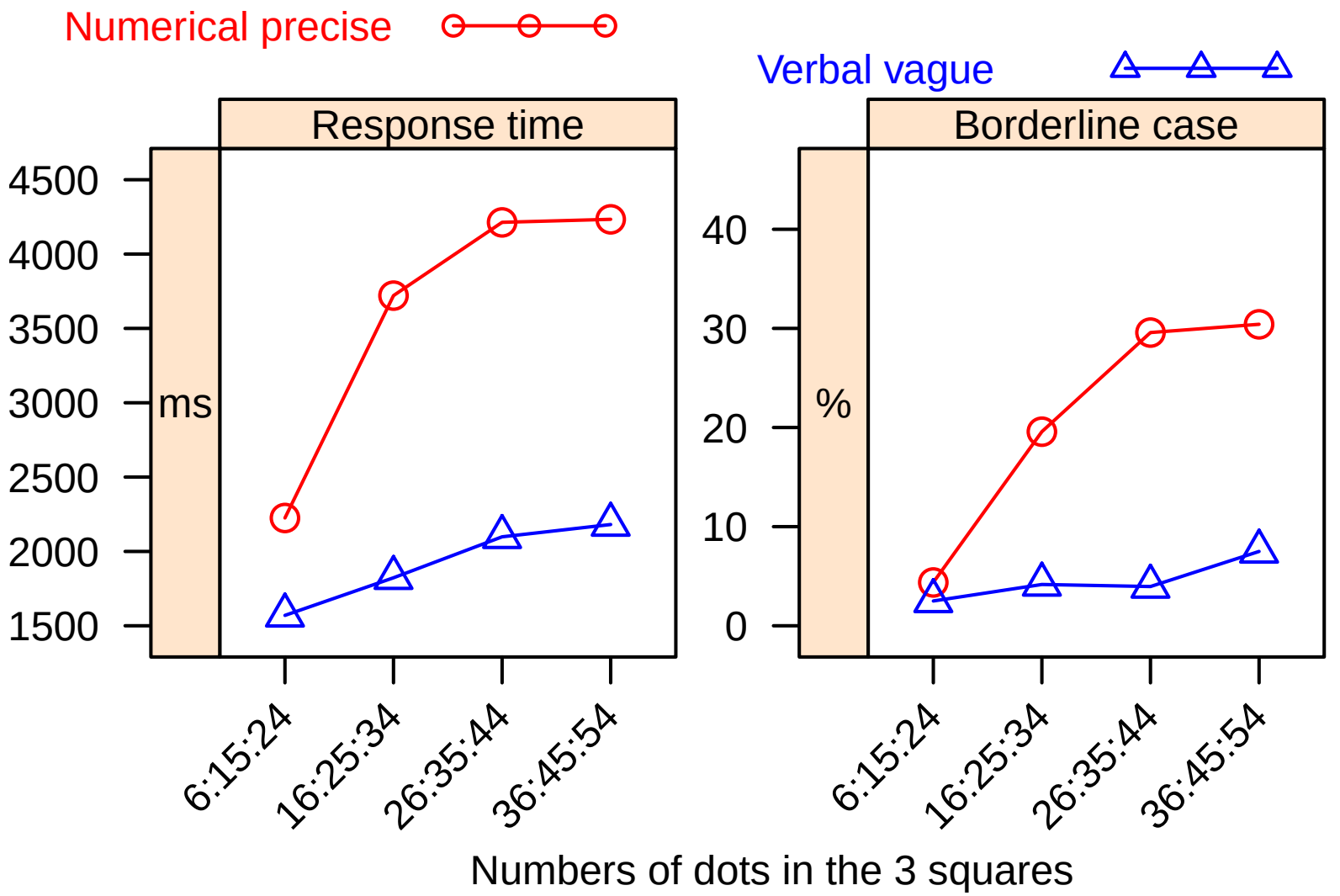
Number combinations: $\{6, 15, 24\}$, $\{16, 25, 34\}$, $\{26, 35, 44\}$, $\{36, 45, 54\}$

- Numerical-precise: Choose the square with 16 dots
- Numerical-vague: Choose a square with about 20 dots
- Verbal-precise: Choose the square with fewest dots
- Verbal-vague: Choose a square with few dots

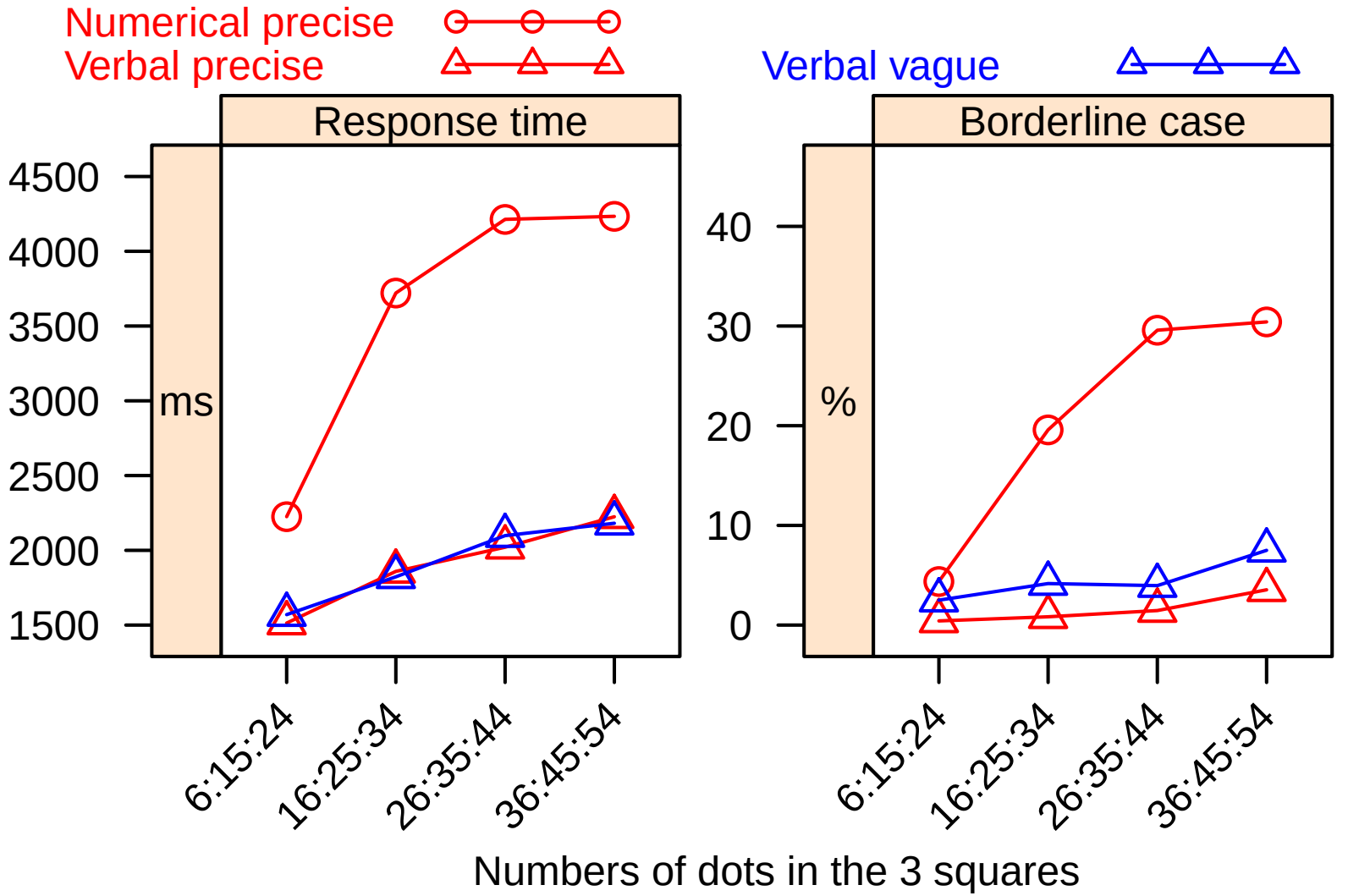
Experiment three (more squares & use indefinites)



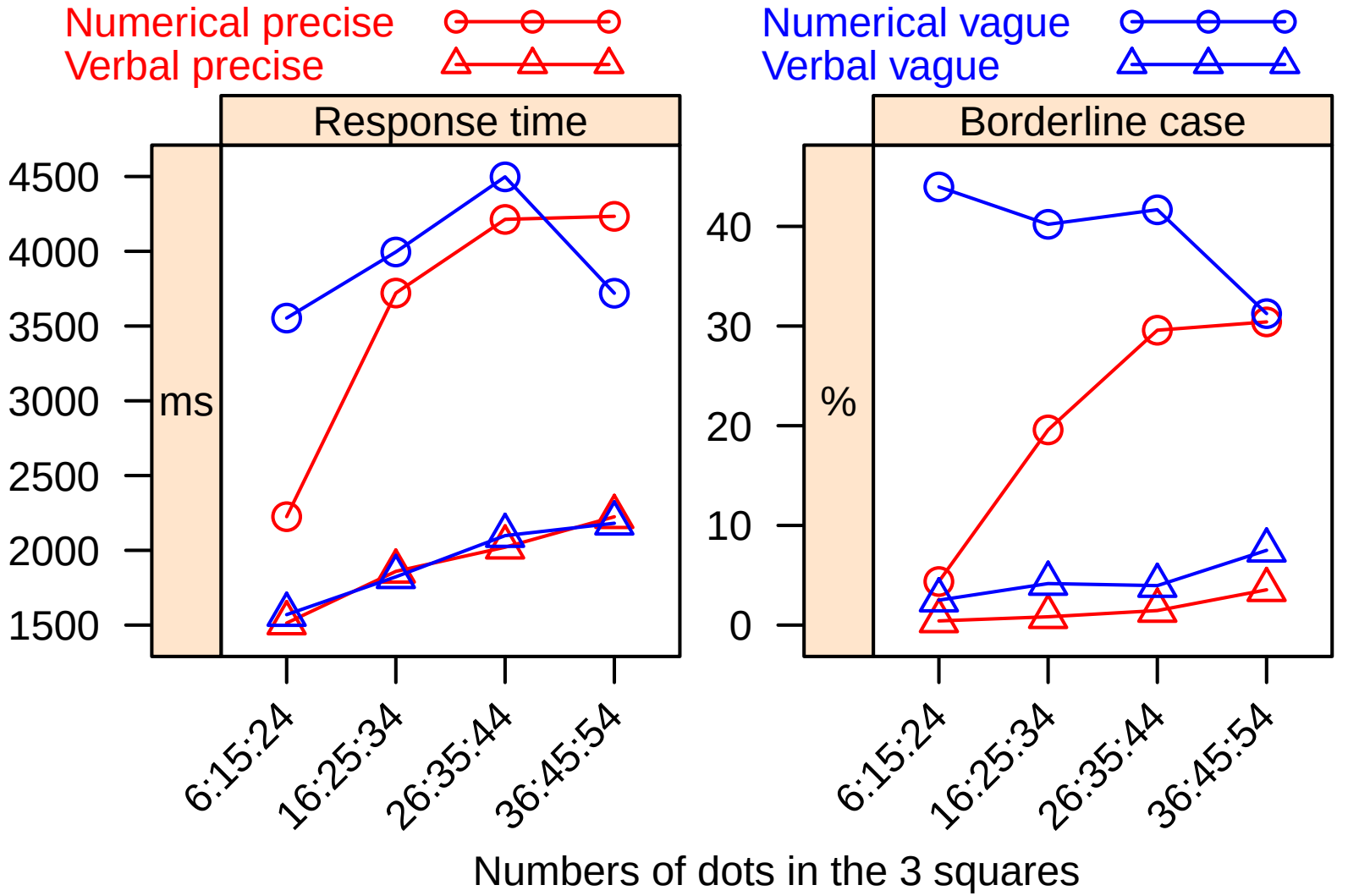
Experiment three (more squares & use indefinites)



Experiment three (more squares & use indefinites)



Experiment three (more squares & use indefinites)



Interpretation and Conclusions

Summary of findings

Interpretation and Conclusions

Summary of findings

- We did find some benefits for using some vague terms . . .

Interpretation and Conclusions

Summary of findings

- We did find some benefits for using some vague terms . . .
- . . . but verbal format is largely responsible

Interpretation and Conclusions

Summary of findings

- We did find some benefits for using some vague terms . . .
- . . . but verbal format is largely responsible

Summary of explanations

Interpretation and Conclusions

Summary of findings

- We did find some benefits for using some vague terms . . .
- . . . but verbal format is largely responsible

Summary of explanations

- Hierarchy of gist account

Interpretation and Conclusions

Summary of findings

- We did find some benefits for using some vague terms . . .
- . . . but verbal format is largely responsible

Summary of explanations

- Hierarchy of gist account
- Heuristic / Systematic account

Interpretation and Conclusions

Summary of findings

- We did find some benefits for using some vague terms . . .
- . . . but verbal format is largely responsible

Summary of explanations

- Hierarchy of gist account
- Heuristic / Systematic account

Lessons for NLG

Interpretation and Conclusions

Summary of findings

- We did find some benefits for using some vague terms . . .
- . . . but verbal format is largely responsible

Summary of explanations

- Hierarchy of gist account
- Heuristic / Systematic account

Lessons for NLG

- Data from human comprehenders can help NLG systems make better decisions

Interpretation and Conclusions

Summary of findings

- We did find some benefits for using some vague terms . . .
- . . . but verbal format is largely responsible

Summary of explanations

- Hierarchy of gist account
- Heuristic / Systematic account

Lessons for NLG

- Data from human comprehenders can help NLG systems make better decisions
- Vagueness is really hard to pin down experimentally

Interpretation and Conclusions

Summary of findings

- We did find some benefits for using some vague terms . . .
- . . . but verbal format is largely responsible

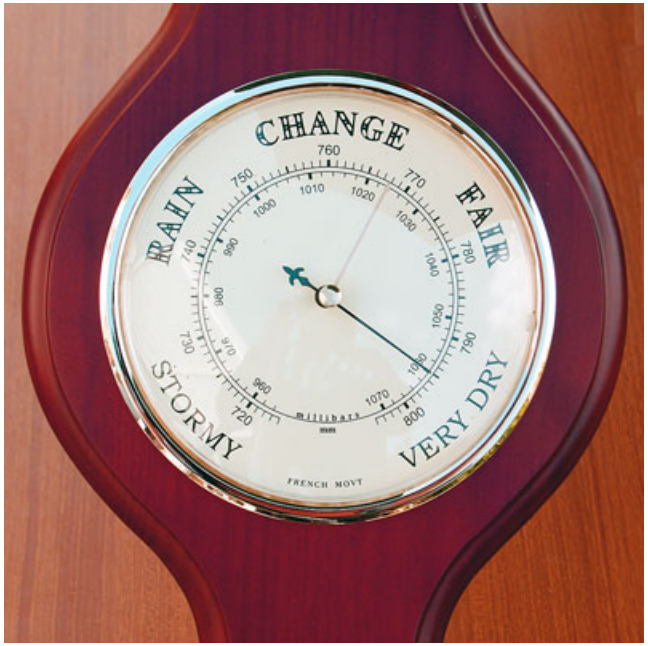
Summary of explanations

- Hierarchy of gist account
- Heuristic / Systematic account

Lessons for NLG

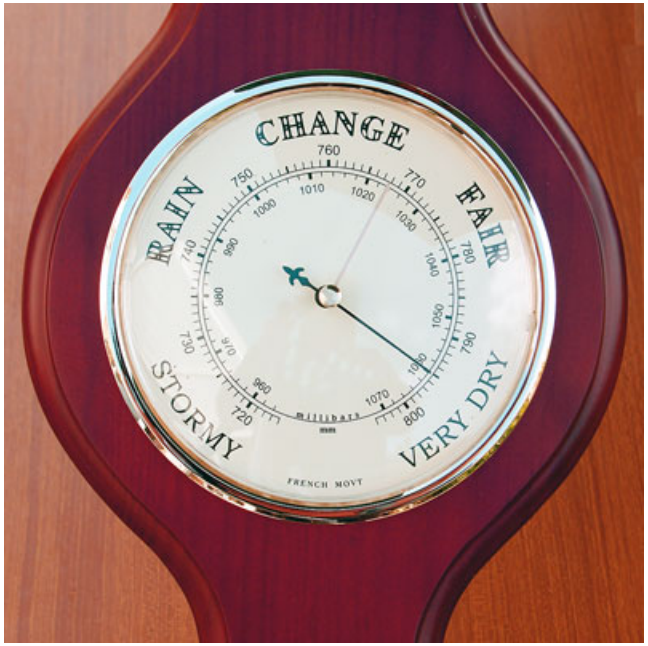
- Data from human comprehenders can help NLG systems make better decisions
- Vagueness is really hard to pin down experimentally
- Using vague terms can make a big difference for cognitive load on comprehenders

Take home message



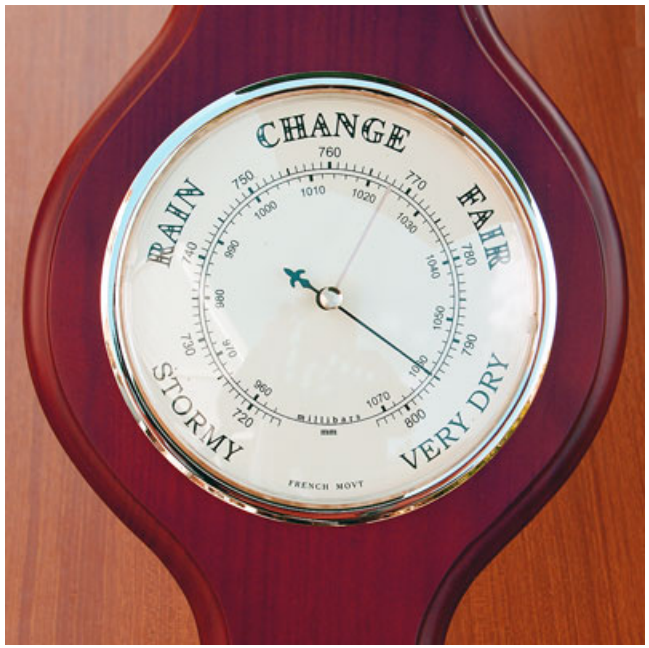
Take home message

Vagueness does make things easier in some situations, but really it is non-numerical format that does the heavy lifting.



Take home message

Vagueness does make things easier in some situations, but really it is non-numerical format that does the heavy lifting.



Thank you for listening. 5 mins for Q & A

Next

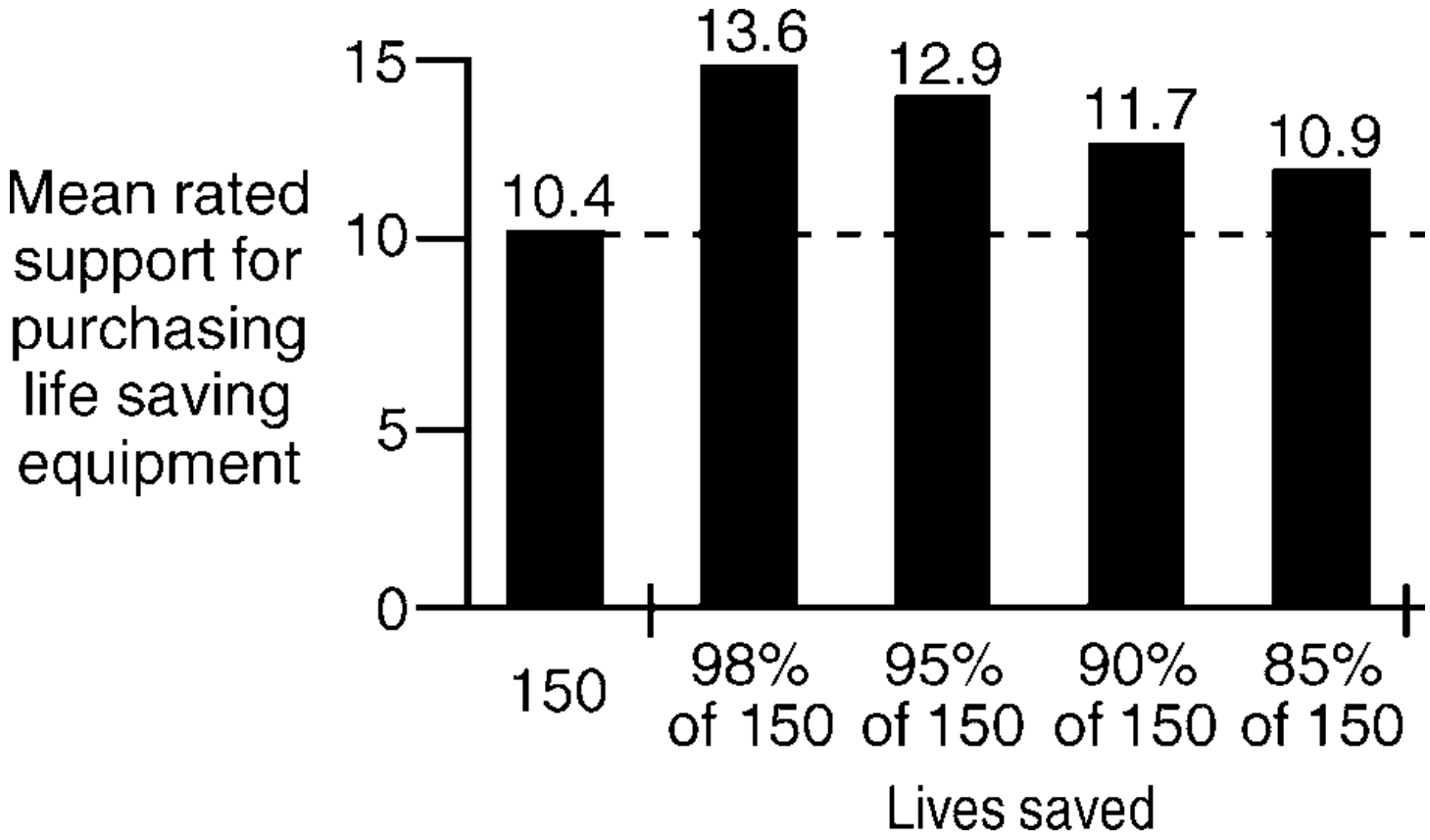


Fig. 4. Saving a percentage of 150 lives received higher support than saving 150 lives (Slovic *et al.*, 2002).

References

Lipman, B. L. (2009). *Why is Language Vague?* (unpublished draft manuscript retrieved 12 April 2011 from `http://people.bu.edu/blipman/Papers/vague5.pdf`)

Portet, F., Reiter, E., Gatt, A., Hunter, J., Sripada, S., Freer, Y., et al. (2009). Automatic generation of textual summaries from neonatal intensive care data. *Artificial Intelligence*, *173*(7-8), 789–816.