What is the best way to present likelihood ratios?

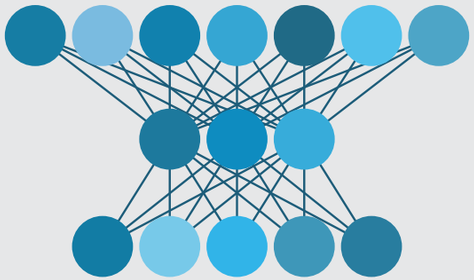
A review of past research and recommendations for future research

Geoffrey Stewart Morrison

Forensic Data Science Laboratory, Aston University

Forensic Evaluation Ltd

Aston University
BIRMINGHAM UK



$$\frac{p(E|H_p)}{p(E|H_d)}$$

Disclaimer

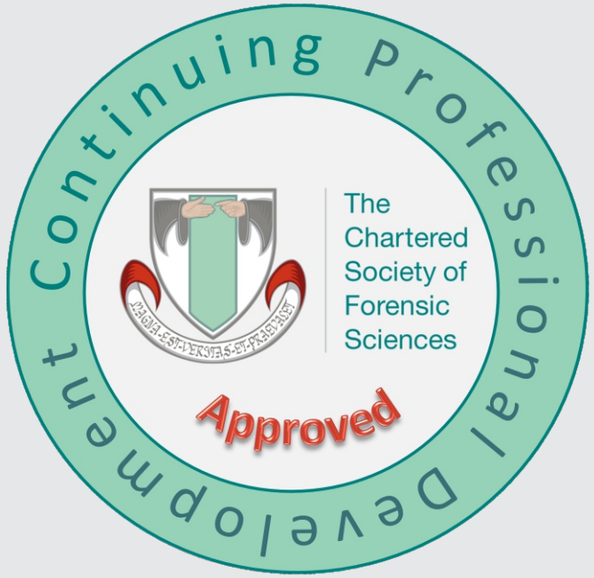
- All opinions expressed are those of the presenter and, unless explicitly stated otherwise, should not be construed as representing the policies or positions of any organizations with which the presenter is associated.

Funding

- The work of Morrison, Bali, and Martire was supported in part by Research England's Expanding Excellence in England Fund as part of funding for the Aston Institute for Forensic Linguistics 2019–2024.

Education

- **Concepts of forensic inference and statistics**
  - Master's level continuing professional development course
  - Online delivery – can be taken from anywhere in the world
  - Delivered in 22 weeks spread over 6 months
  - ~1 day per week workload
  - Active learning, flipped classroom, didactic testing and feedback
  - Weekly interactive sessions
  - Competency assessment



Slides

- https://forensic-data-science.net/communication/#AFDAA2026

Paper

- Morrison G.S., Bali A.S., Martire K.A., Grady R.H., Thompson W.C. (2025). **What is the best way to present likelihood ratios? A review of past research and recommendations for future research**. *Science & Justice*, 65, 101342. https://doi.org/10.1016/j.scijus.2025.101342

Content

- Question
- Step 1 (conduct a review)
- Formats (definitions)
- Inclusion criteria
- Papers included
- Formats (results)
- Values
- Evidence types
- Participants
- Experiment design
- Indicia of understanding
  - Sensitivity
  - Orthodoxy
  - Coherence
- Conclusion

Question

- What is the best way for forensic practitioners to present likelihood ratios so as to maximize their understandability for legal-decision makers?

Step 1

- Conduct a review of existing literature related to the question
    - what is already known
    - what still needs to be researched
    - methodology for future research

Formats

- likelihood ratio

$$\frac{p(E|H_1)}{p(E|H_2)}$$

  - H_1: same source
  - H_2: different source

Formats

- **numerical likelihood ratios**
    - the observations are 1,000 times more probable if H_1 were true than if H_2 were true
- **numerical random-match probabilities**
    - the observations made on the questioned-source item and the known-source item match, the probability of observations made on an item randomly selected from the relevant population matching the observations from the questioned-source item is 1 in 1,000

Formats

- **verbal likelihood ratios**
    - the observations are much more probable if H_1 were true than if H_2 were true
- **verbal strength-of-support statement**
    - the observations provide strong support for H_1 relative to H_2
    - the observations provide strong support for H_1

Formats

Numerical range	Verbal likelihood ratios	Verbal strength of support
$0.5 < LR < 2$	The observations are ***approximately equally probable*** irrespective of whether H1 or H2 were true.	The observations provide ***no support*** for either H1 or H2.
$2 \leq LR < 10$	The observations are ***slightly more probable*** if H1 were true than if H2 were true.	The observations provide ***weak support*** for H1 relative to H2.
$10 \leq LR < 100$	The observations are ***more probable*** if H1 were true than if H2 were true.	The observations provide ***moderate support*** for H1 relative to H2.
$100 \leq LR < 1{,}000$	The observations are ***appreciably more probable*** if H1 were true than if H2 were true.	The observations provide ***moderately strong support*** for H1 relative to H2.
$1{,}000 \leq LR < 10{,}000$	The observations are ***much more probable*** if H1 were true than if H2 were true.	The observations provide ***strong support*** for H1 relative to H2.
$10{,}000 \leq LR < 1{,}000{,}000$	The observations are ***far more probable*** if H1 were true than if H2 were true.	The observations provide ***very strong support*** for H1 relative to H2.
$1{,}000{,}000 \leq LR$	The observations are ***exceedingly more probable*** if H1 were true than if H2 were true.	The observations provide ***extremely strong support*** for H1 relative to H2.

Inclusion criteria

- primary report of empirical research
- testing layperson understanding of likelihood ratios
- **numerical likelihood ratios** were presented to participants
    - could also include presentation of:
        - numerical random-match probabilities
        - verbal likelihood ratios
        - strength-of-support statements

Included papers

17 total:

- Koehler (1996)
- Taroni & Aitken (1998)
- Nance & Morris (2002)
- Nance & Morris (2005)
- Langenburg et al. (2013)
- Martire et al. (2013)
- Martire et al. (2014)
- Thompson & Newman (2015)
- Bayer et al. (2016)
- Thompson et al. (2018)
- Ribeiro et al. (2020)
- van Straalen et al. (2020)
- Bali et al. (2021)
- Ribeiro et al. (2023)
- van Straalen et al. (2023)
- Bali & Martire (2025)
- Thompson et al. (2025)

Formats

Format	Number of studies
numerical likelihood ratio	22
numerical random-match probability	12
verbal likelihood ratio	0
support statement (1 hypothesis)	9
support statement (2 hypotheses)	3
location on line	1

Values

- values presented $p(E|H_1)/p(E|H_2)$:

1/495,000

1/1,000

1/4.5

4.5

5

5.5

25

30

50

55

100

450

550

1,000

3,000

5,500

40,000

100,000

250,000

495,000

550,000

1,000,000

5,000,000

5,500,000

10,000,000

Evidence types

Evidence type	Number of studies
DNA	12
fingerprints	5
footwear	5
voice recordings	1

Participants

Participants	Number of studies
university students	6
general community	7
jury-eligible community	8
former jurors / jury-pool members	3
criminal-justice professionals	4

Experiment design

Response type	Number of studies
odds	5
probability (%)	8
multilevel scale	8
binary	2

Experiment design

Within / Between	Number of studies
within participant	4
between participants	16

Experiment design

Prior and posterior elicited from	Number of studies
same participant	11
different participants	2

Experiment design

Presentation format	Number of studies
written	20
video / live	2

Experiment design

Participants responded as	Number of studies
individuals (judges)	22
collaborating groups (juries)	0

Indicia of understanding

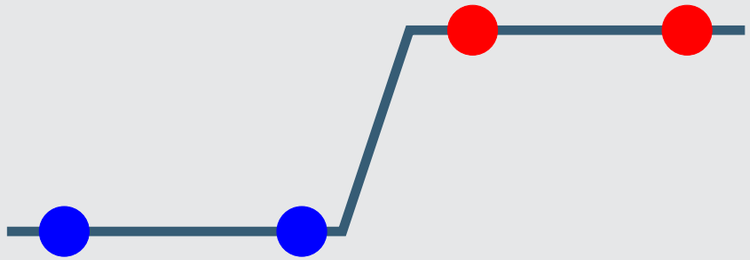
- **Sensitivity**
    - Participants' responses are *sensitive* if they reflect relative differences between different presented likelihood-ratio values.

Indicia of understanding

- **Sensitivity**
  - 15 studies used *sensitivity* as an indicium of understanding.
  - With the exception of 2 studies, and some conditions in 3 other studies, all studies found that participants were *sensitive* to differences in likelihood-ratio values.
  - This was true across all
    - formats
    - response types
    - evidence types
    - demographic groups

Indicia of understanding

- **Sensitivity**
  - Meta analysis, including exceptions and studies with more than 2 likelihood-ratio values, leads to hypothesis:
    - threshold somewhere between 100 and 450
    - sensitive to differences that cross the threshold
    - not sensitive to differences below the threshold or differences above the threshold



Indicia of understanding

- **Orthodoxy**
    - Participants' responses are *orthodox* if they reflect use of the values of presented likelihood ratios to update priors to posteriors as per correct application of Bayes' theorem.

Indicia of understanding

- **Orthodoxy**
  - Bayes' theorem:

$$posterior\ odds = prior\ odds\ \times likelihood\ ratio$$

  - effective likelihood ratio:

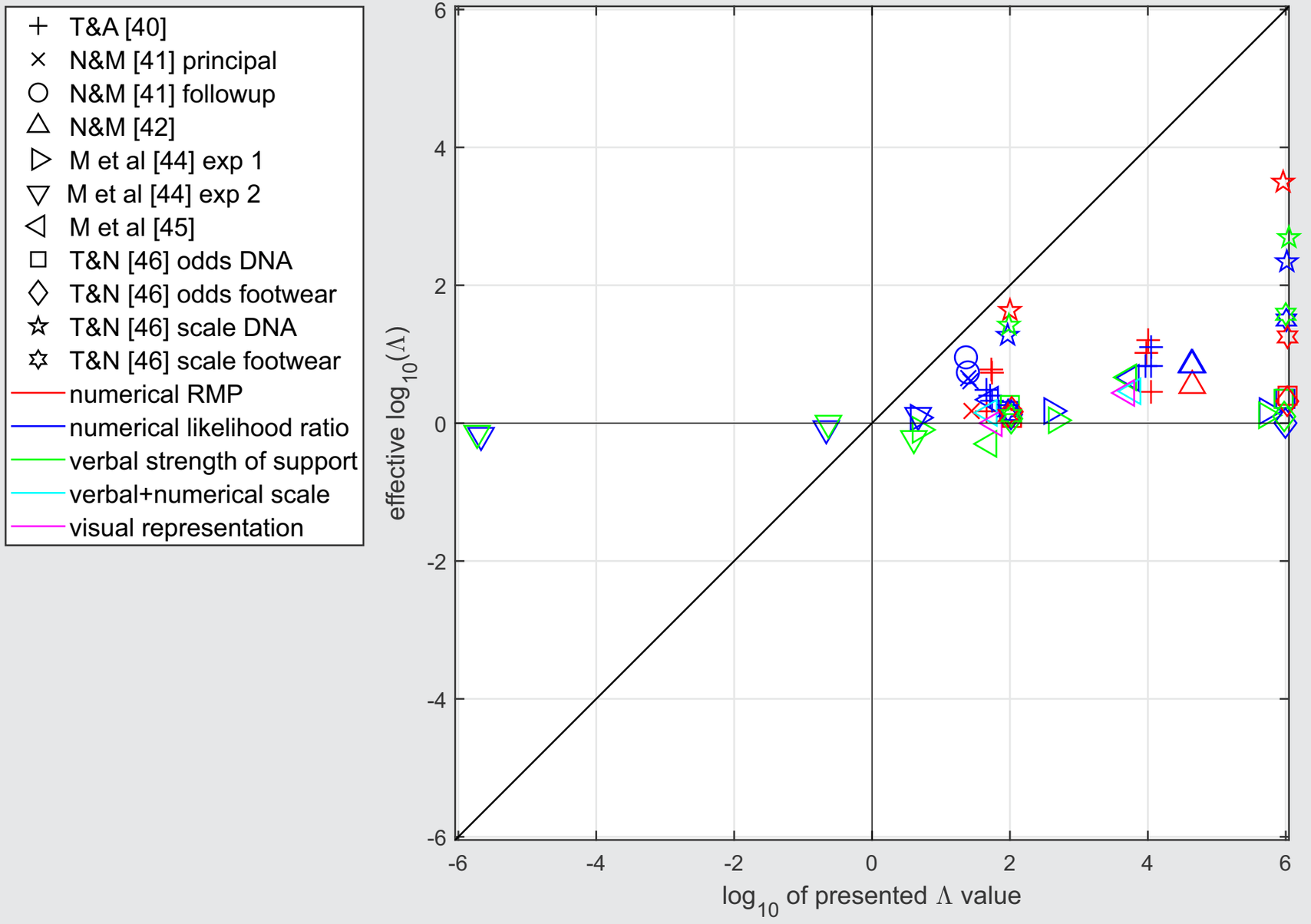
$$effective\ likelihood\ ratio = \frac{posterior\ odds}{prior\ odds}$$

Indicia of understanding

- **Orthodoxy**
  - 14 studies used *orthodoxy* as an indicium of understanding.
  - Average effective likelihood ratios were always weaker (closer to the neutral value of 1) than presented likelihood ratios, e.g.:
    - presented: 1 million
    - effective: less than 10
  - This was true across all formats.

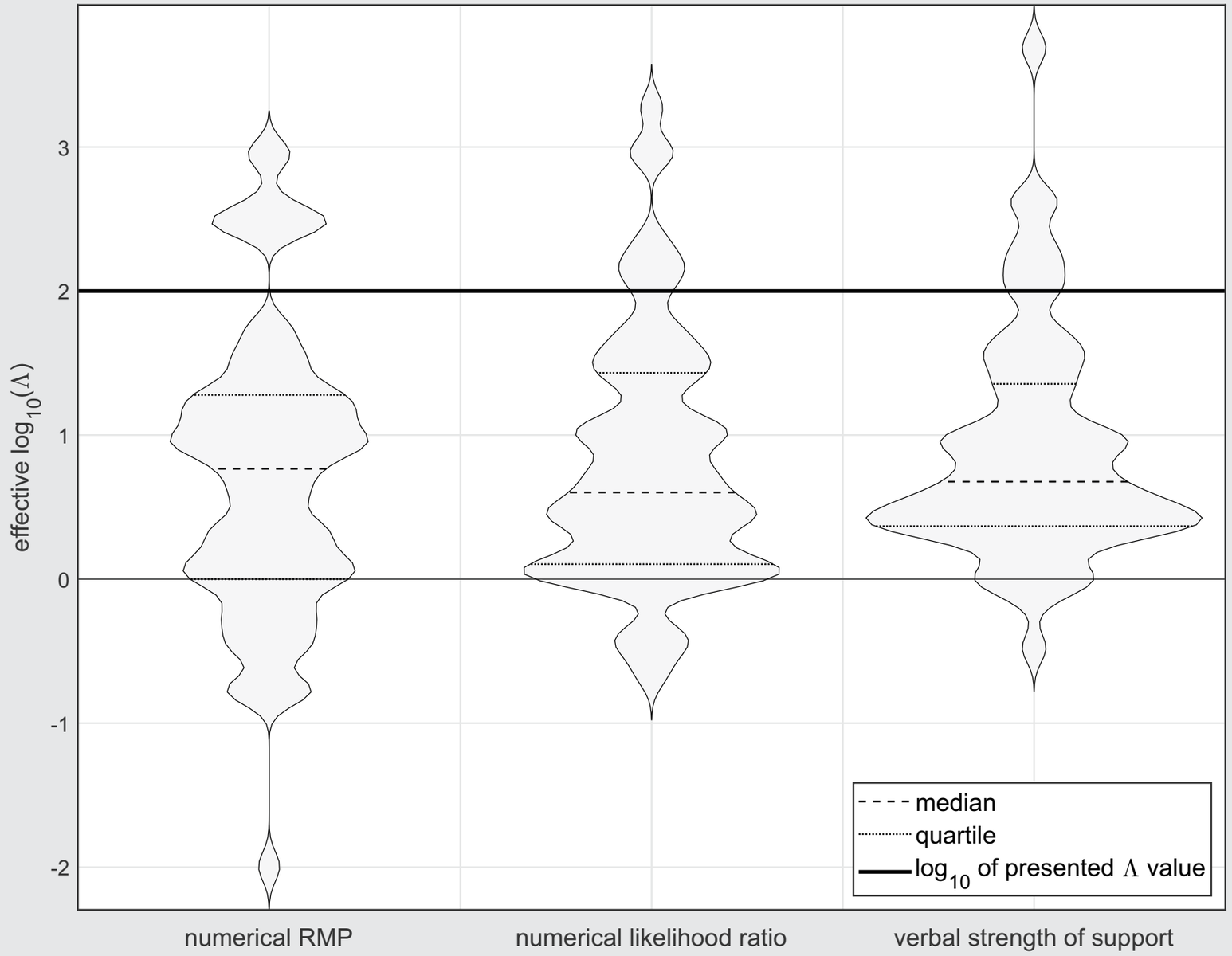
Indicia of understanding

- **Orthodoxy**



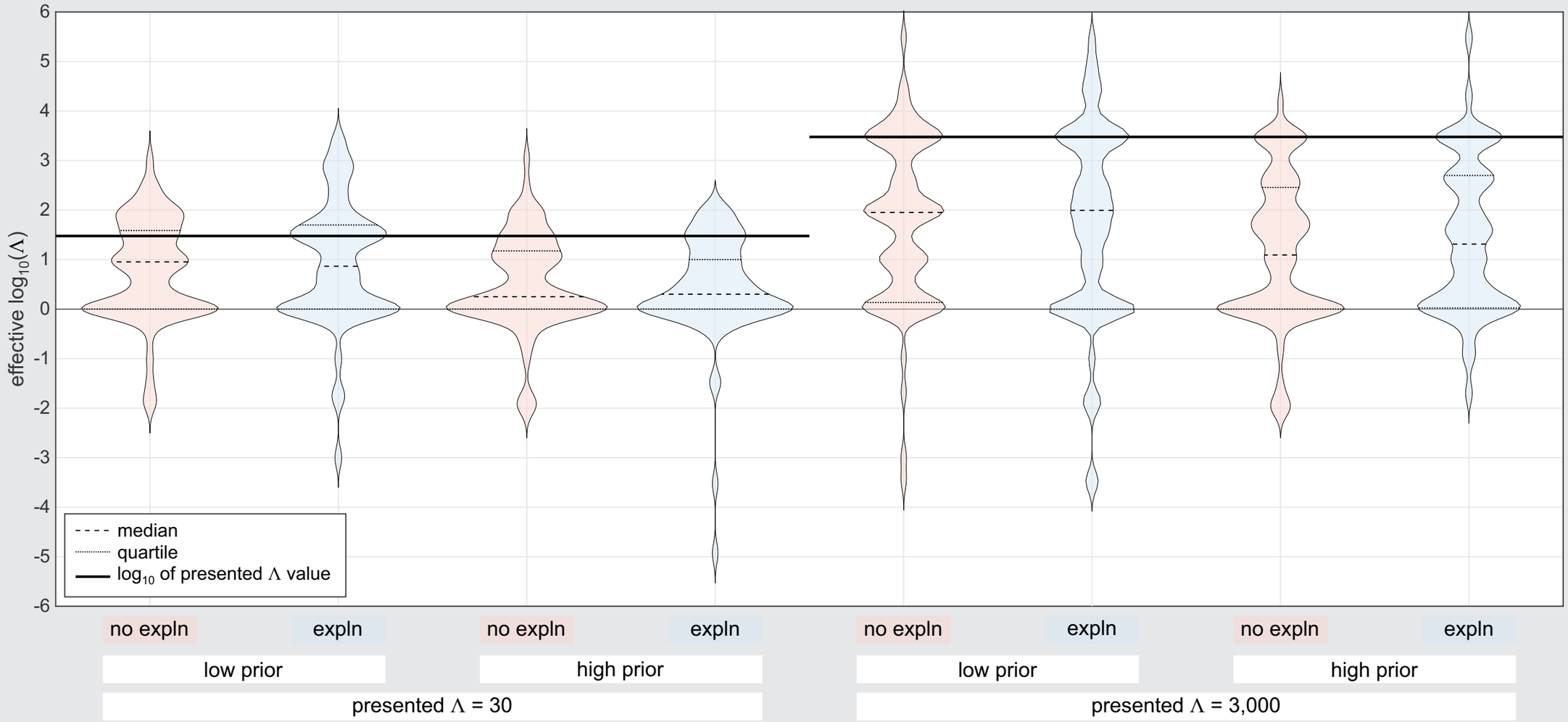
Indicia of understanding

- **Orthodoxy**
  - Bali et al. (2021)



Indicia of understanding

- **Orthodoxy**
  - Thompson et al. (2025)
  - numerical likelihood ratios



Indicia of understanding

- **Orthodoxy**
  - Thompson et al. (2025)
    - excluding participants whose *prior odds* = 1
    - num participants whose *effective likelihood ratio = presented likelihood ratio*
      - given explanation: 7 / 232 (3.0%)
      - not given explanation: 2 / 272 (0.74%)

Indicia of understanding

- **Orthodoxy**
    - Some studies elicited offence-level rather then source-level prior and posteriors
    - Many studies included substantial extraneous case information
    - Participants may have weighted the likelihood ratios using
        - a priori beliefs about validity of branch of forensic science (DNA vs footwear)
        - perception of quality of testimony presented

Indicia of understanding

- **Coherence**
    - Participants' responses are *coherent* if they reflect logically correct interpretation of likelihood ratios, i.e., if they indicate that participants have avoided reasoning errors and logical fallacies.

Indicia of understanding

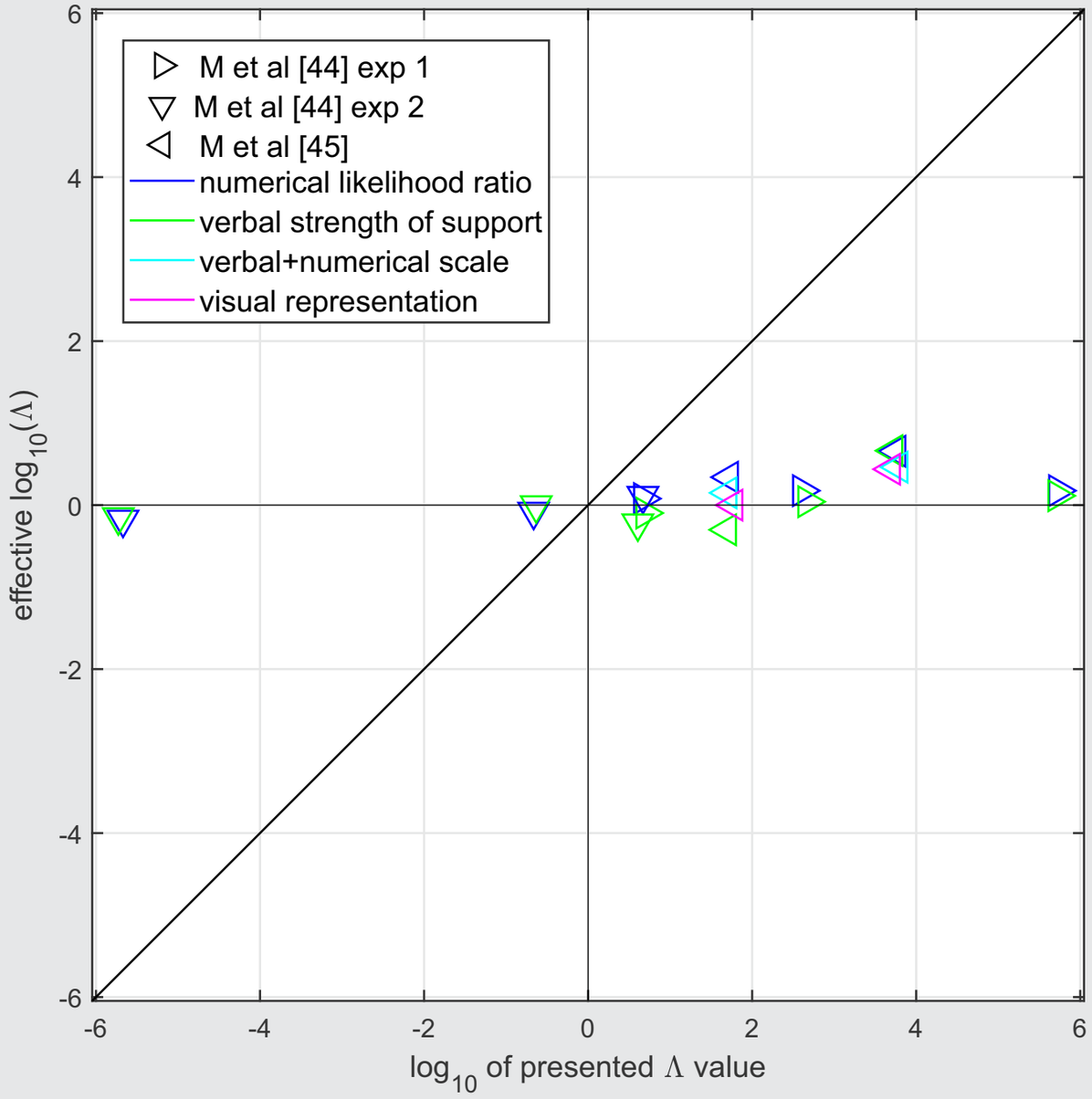
- **Coherence**
    - 14 studies used *coherence* as an indicium of understanding.
        - weak-evidence effect
        - prosecutor’s fallacy

Indicia of understanding

- **Coherence**
  - **weak-evidence effect**
    - "weak support" for H_1 is interpreted as support for H_2
    - a numerical likelihood ratio that is a little larger than 1 is interpreted as if it were a numerical likelihood ratio that is less than 1

Indicia of understanding

- **Coherence**
  - **weak-evidence effect**
    - common for verbal support statements (64%)
    - reduced by providing whole verbal scale (32%)
    - not common for $p(E|H_1)/p(E|H_2) < 1$



Indicia of understanding

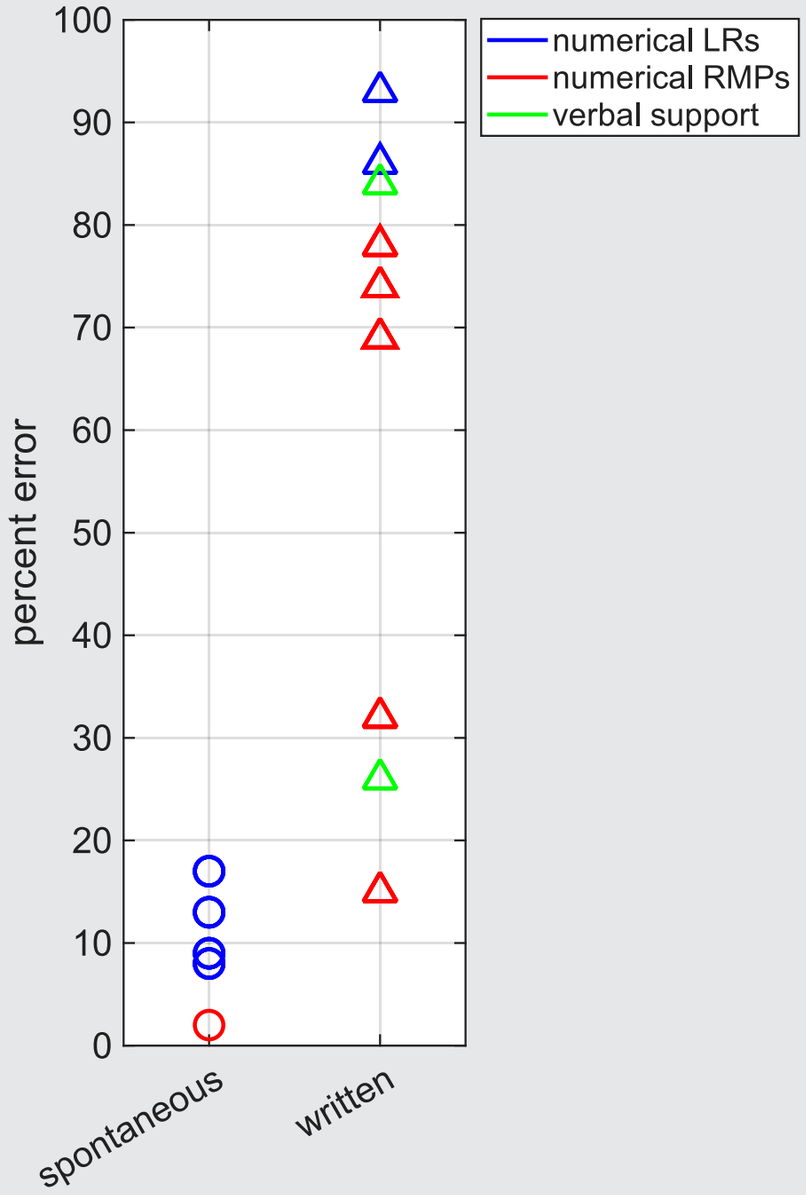
- **Coherence**
  - **prosecutor's fallacy**
    - the likelihood ratio $\frac{p(E|H_1)}{p(E|H_2)}$

      is interpreted as if it were

    - the posterior odds $\frac{p(H_1|E)}{p(H_2|E)}$

Indicia of understanding

- **Coherence**
  - **prosecutor's fallacy**
    - **spontaneous** occurence of prosecutor's fallacy when likelihood ratio presented and posterior odds elicited
    - failure to recognize that a **written** statement included the prosecutor's fallacy



Indicia of understanding

- **Coherence**
  - **prosecutor's fallacy**
  - Thompson et al. (2025)
    - excluding participants whose *prior odds* = 1
    - num participants whose *posterior odds = presented likelihood ratio*
      - given explanation: 31 / 232 (13%)
      - not given explanation: 47 / 272 (17%)

Conclusion

- Results from published studies suggest understanding of likelihood ratios is poor irrespective of
    - presentation format (numerical LR, numerical RMP, verbal support statement)
    - provision of explanation
    - provision of table/graph for converting from prior probabilities to posterior probabilities
    - provision of whole verbal scale

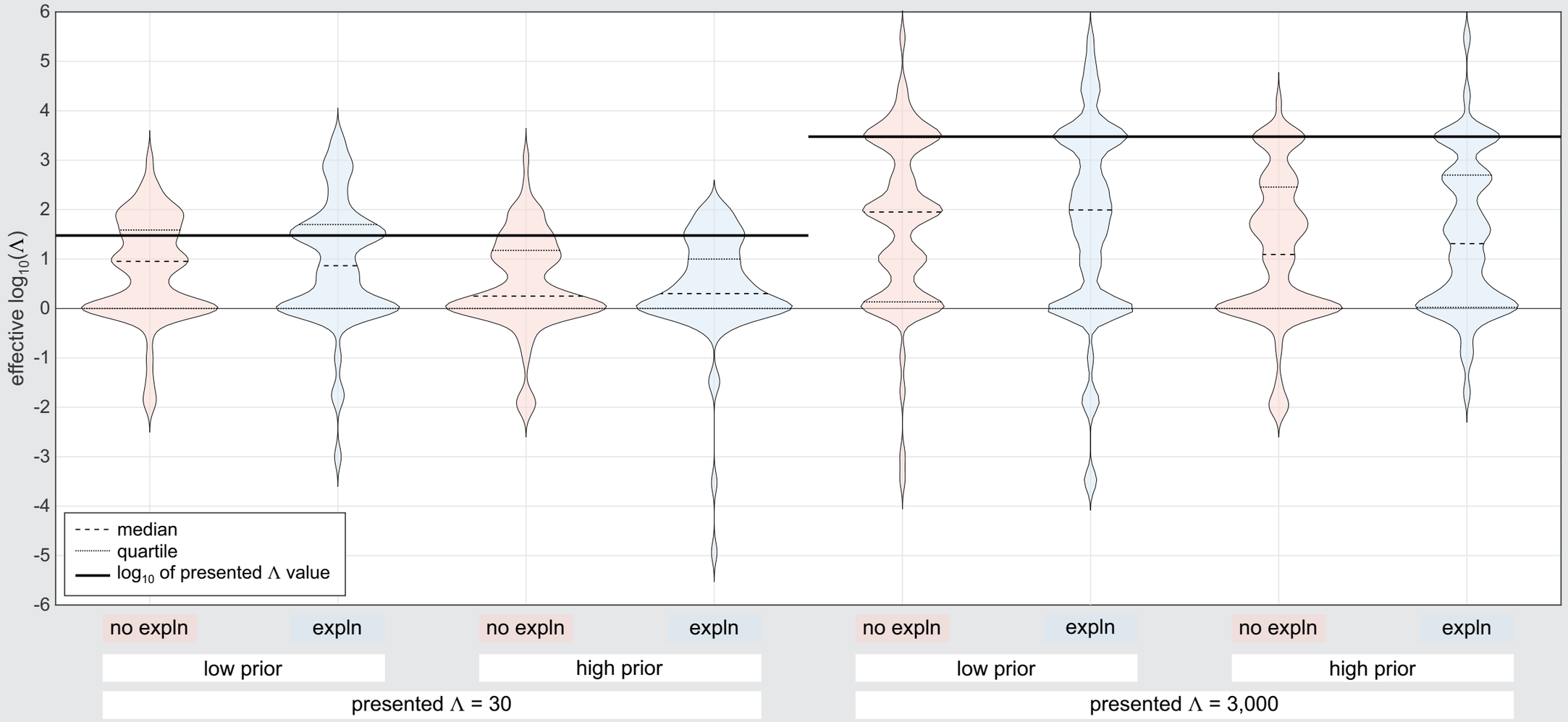
- Most published studies do not address our research question

- Most published studies have weaknesses in research design

Conclusion

- Future research focussed on our research question
  - systematic series of experiments
  - numerical likelihood ratios
  - multiple values below and above hypothesized threshold
  - $p(E|H_1)/p(E|H_2) > 1$ and $p(E|H_1)/p(E|H_2) < 1$
  - video
  - better(?) explanation
  - avoid explanations with prior odds = 1
  - describe method and validation results
  - minimal case information
  - avoid scenarios that elicit prior odds = 1
  - elicit prior odds and posterior odds
  - ask for posterior odds if Bayes' theorem had been applied
  - include proctored experiments
  - include legal-decision makers as participants
  - include groups of collaborating participants

Thank You


effective log10(Λ)
6
5
4
3
2
1
0
-1
-2
-3
-4
-5
-6
median
quartile
log10 of presented Λ value
no expln
expln
no expln
expln
no expln
expln
no expln
expln
low prior
high prior
low prior
high prior
presented Λ = 30
presented Λ = 3,000