

TIDES: Truncation Induced DEpendency among Summary statistics

Ian Hussey

Digitalisation of Psychology

Motivation

- Basic questions:
 - What values of SD are possible?
 - How do I develop an intuition for impossible or unlikely SDs?
- Test your knowledge:
 - Is $SD = 1.21$ small or large for a 1-7 Likert scale?
 - What about $SD = 3.72$?
 - What about $SD = 0.45$?

Information that is
repeated or overlooked

Granular & Truncated Data

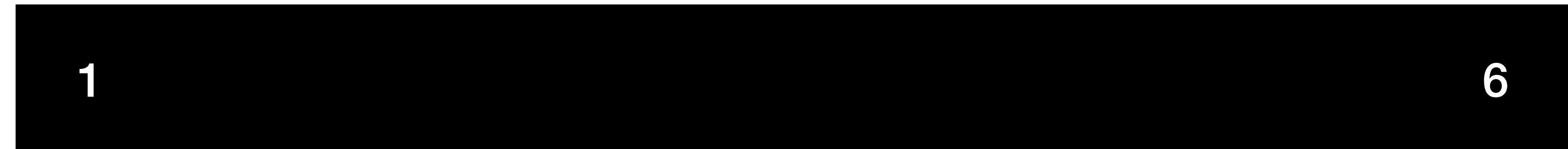
Granular but not (very) Truncated

E.g. Age in years; number of time you open TikTok per day



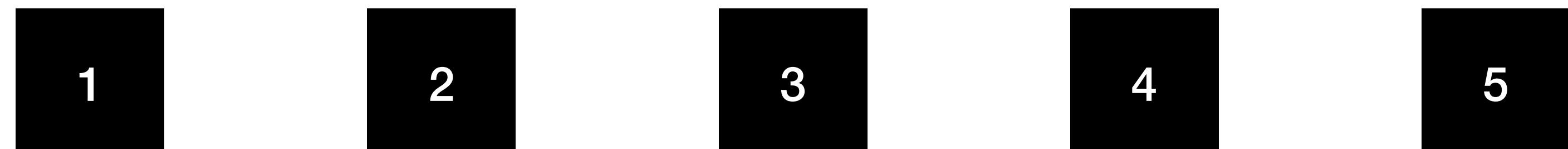
Truncated but not (very) Granular

E.g. University grades (before rounding)



Granular and Truncated

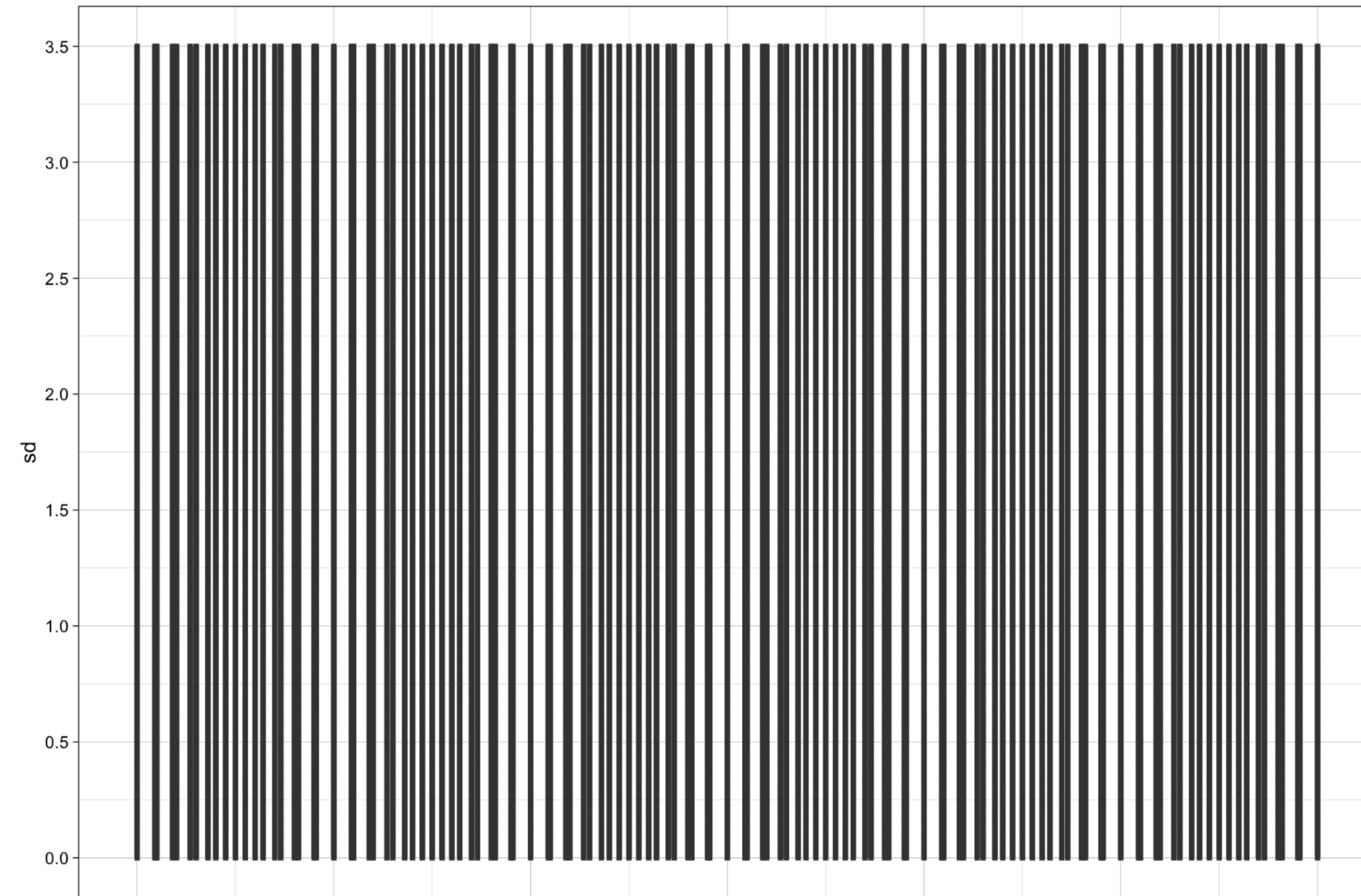
E.g. Likert scales



Granularity

GRIM

Only some means are possible

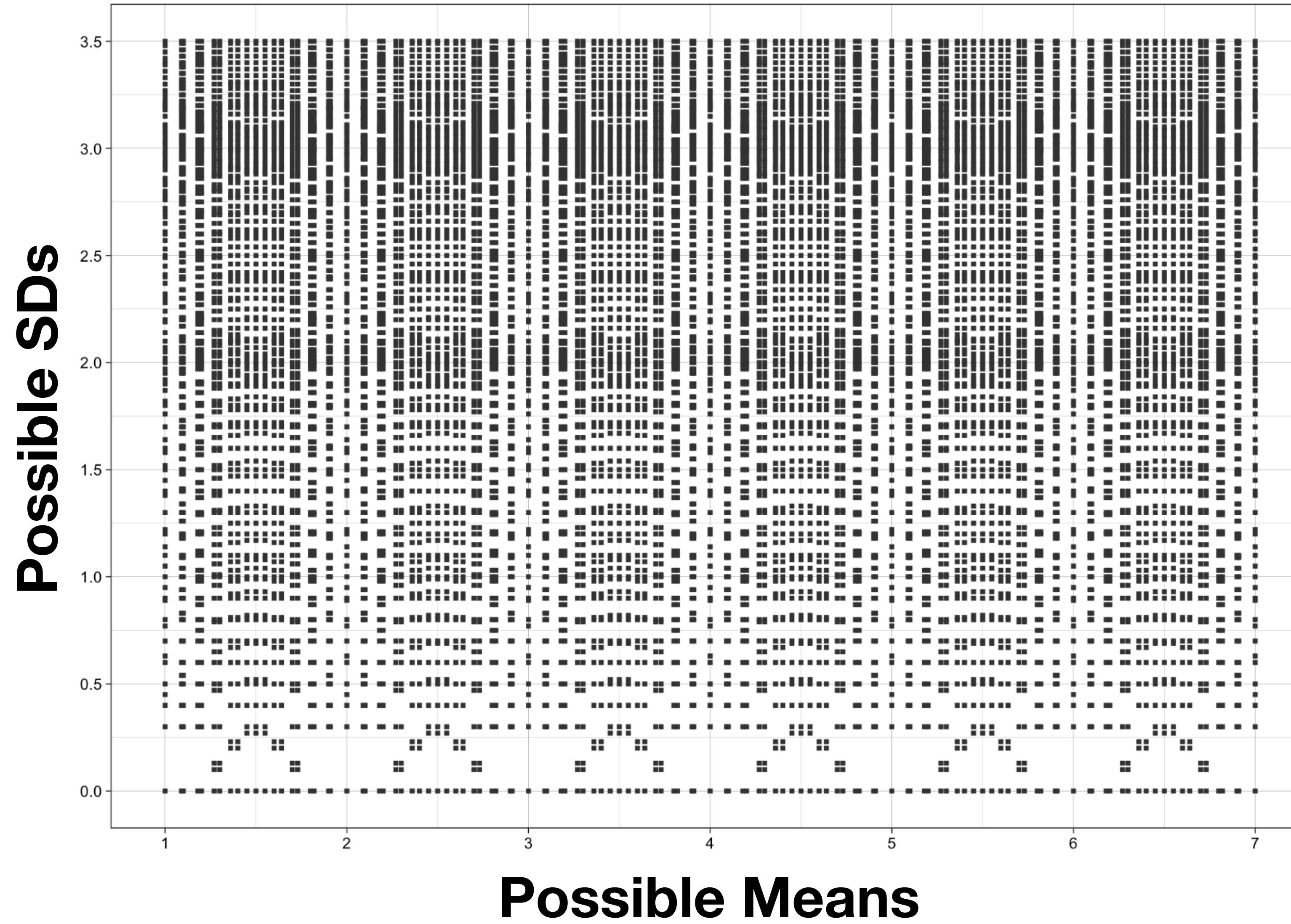


*when $N = 14 +$
1-7 Likert scale*

Granularity

GRIM + GRIMMER

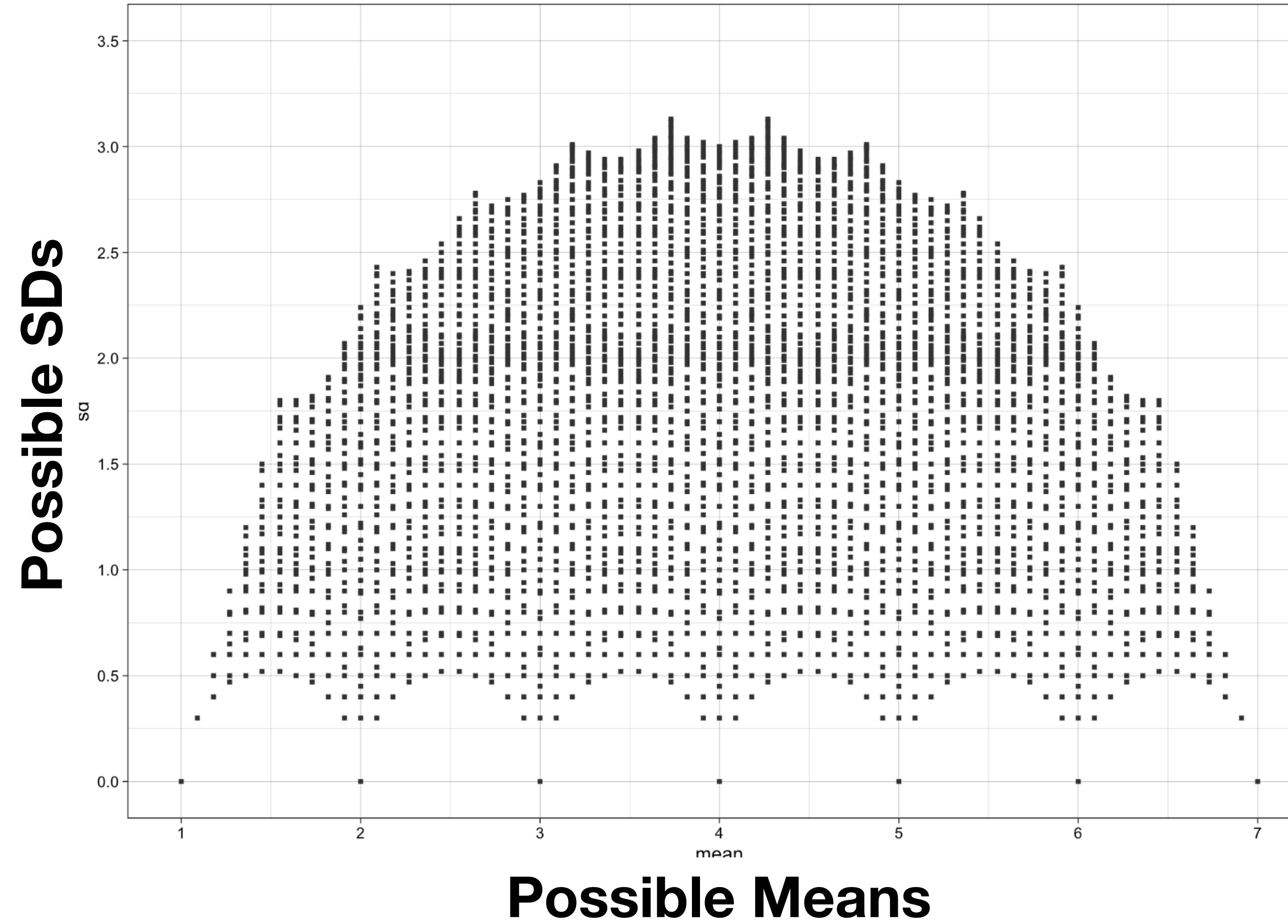
Only some means, some SDs are possible



Granularity
+
truncation

GRIM + GRIMMER + TIDES

Only some means + SDs are possible



“Umbrella plot”

Heathers (2018)



Truncation-induced limits on *Means*

- Truncation has simple implications for means
 - Means cannot be smaller than the lower truncation limit
 - Or larger than the upper truncation limit
- E.g., mean scores on a 1-7 Likert scale cannot $\neq 8.1$
- This is obvious ... but people rarely check it

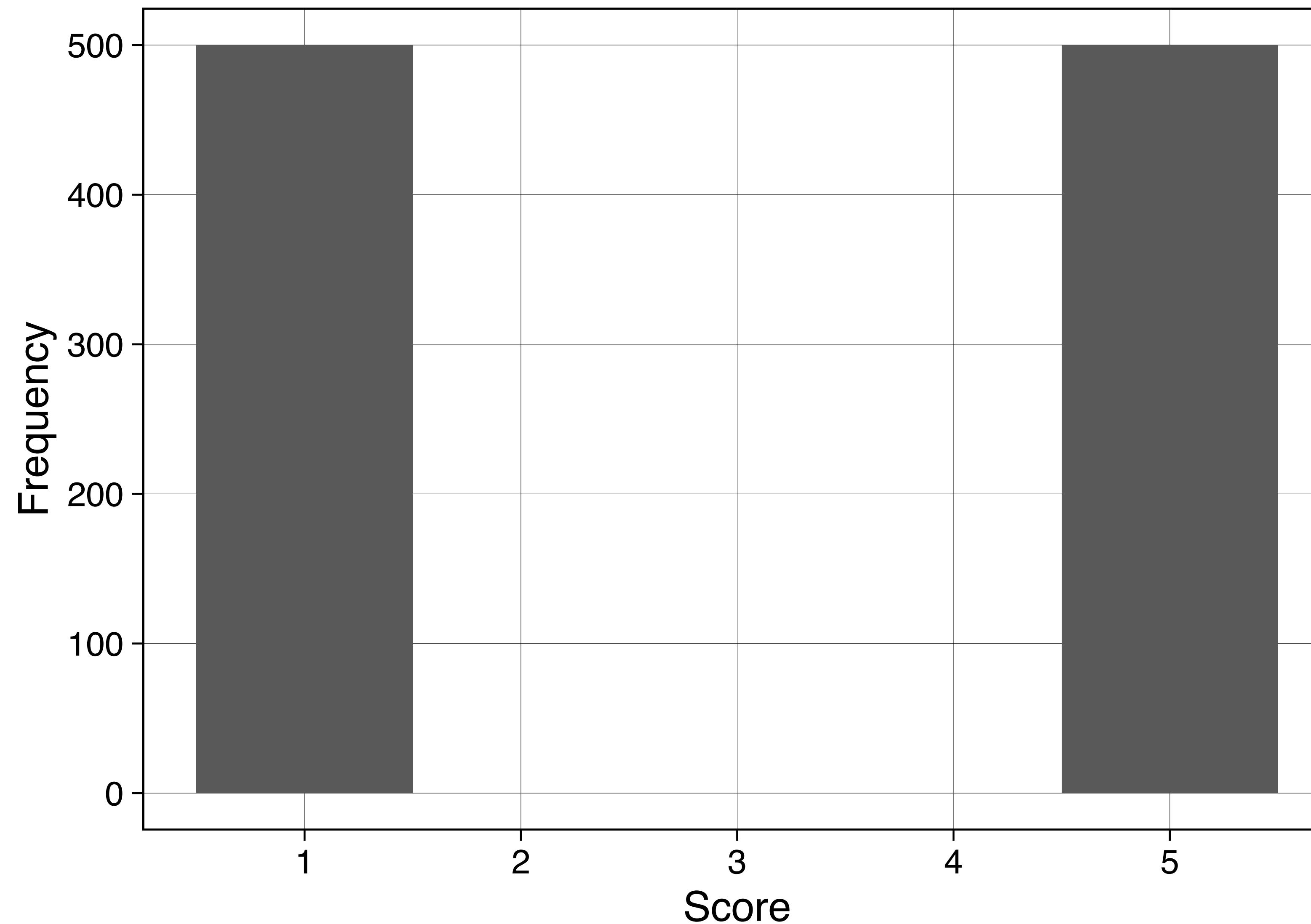
Truncation-induced limits on *SDs*

- Truncation has more complex implications for Standard Deviations (SDs)
 - Max SD
 - Min SD

[reminder to self: draw the rectangle-to-umbrella on the board for following slides]

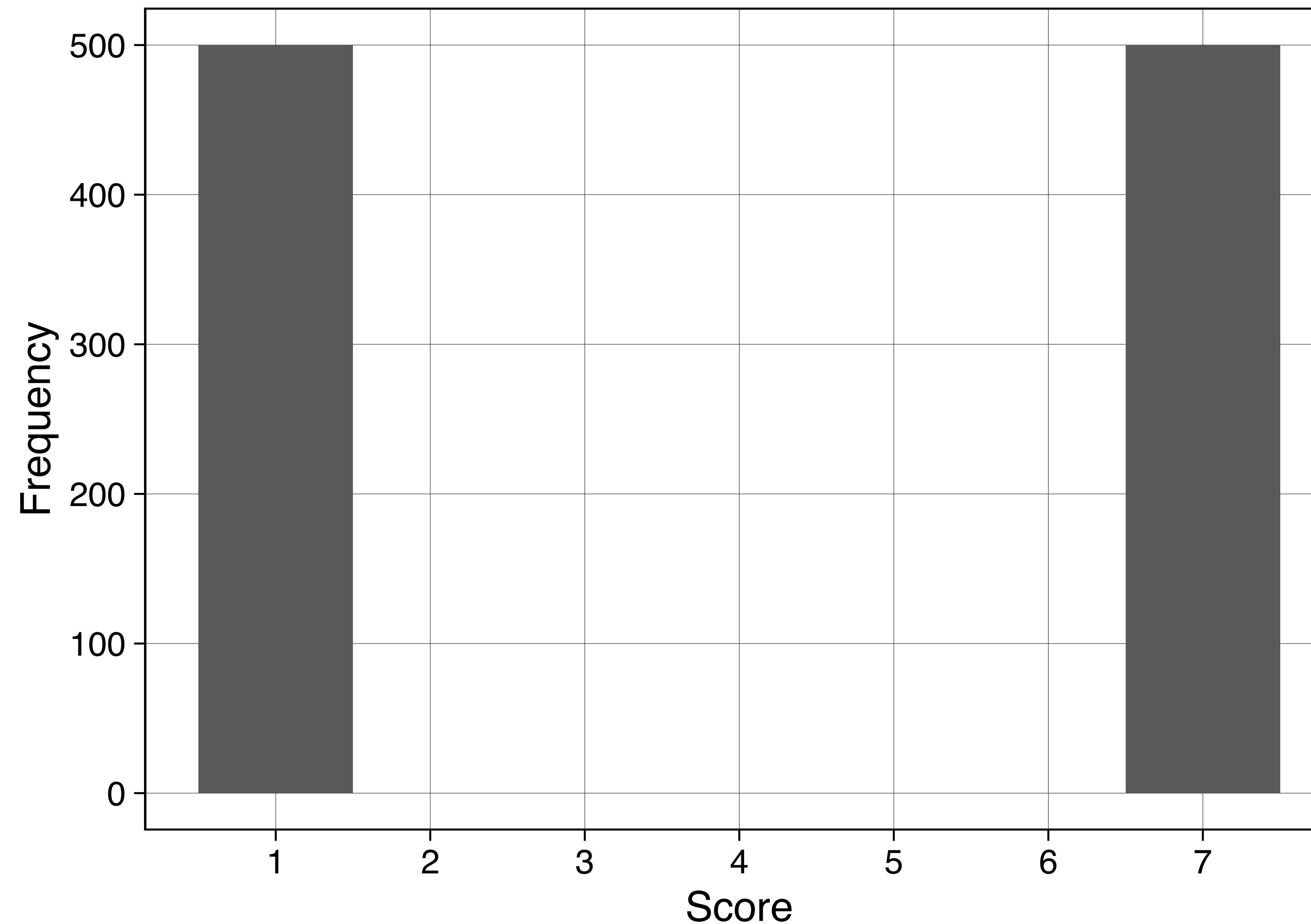
Max SD for whole scale

Max SD of 5-point scale:
 $M = 3.00$, $SD = 2.00$, $n = 1000$



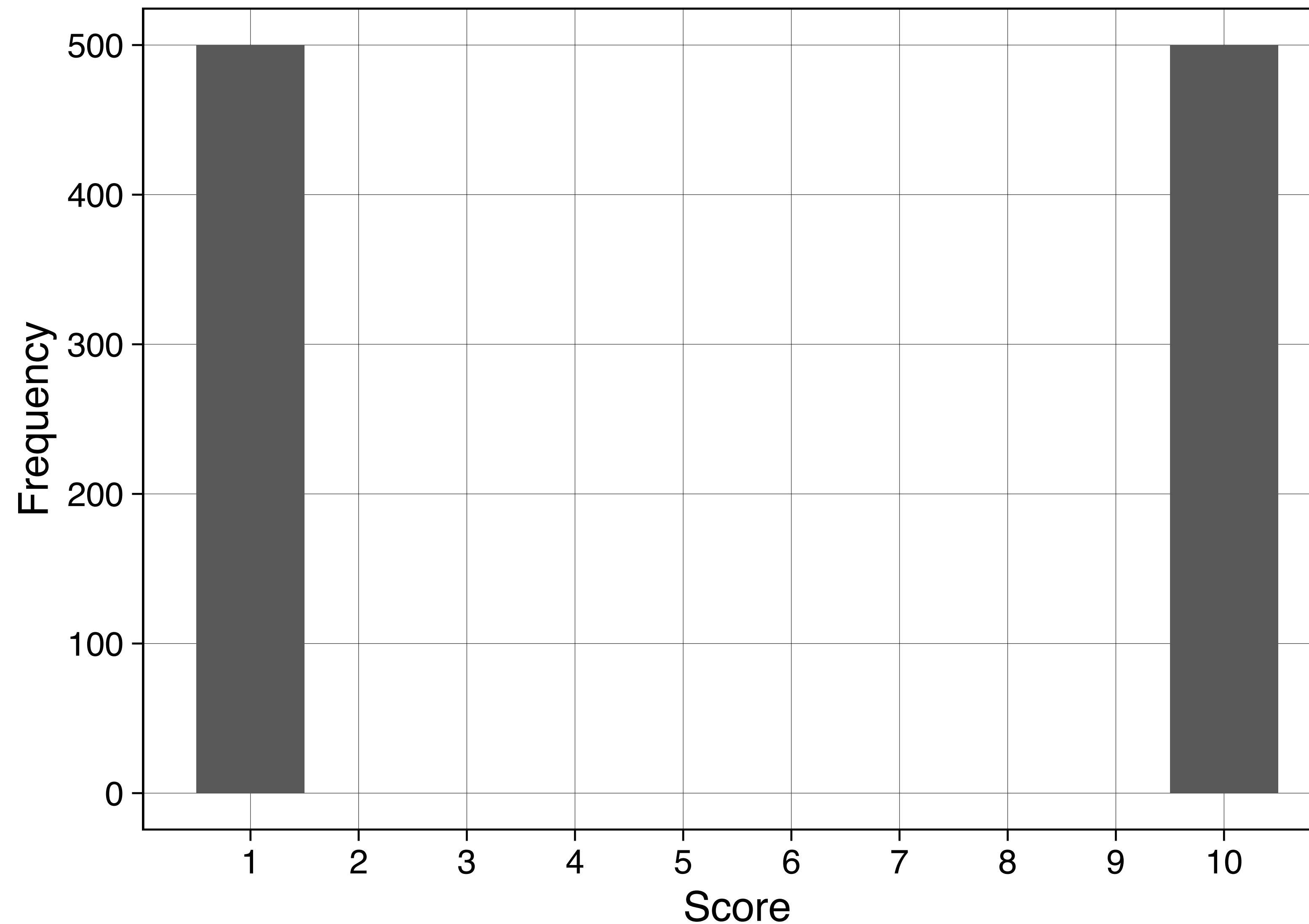
Max SD for whole scale

Max SD of 7-point scale:
 $M = 4.00$, $SD = 3.00$, $n = 1000$



Max SD for whole scale

Max SD of 10-point scale:
 $M = 5.50$, $SD = 4.50$, $n = 1000$



Max SD for whole scale

Response options	Max SD (in large N)
5	2
7	3
10	4.5
General Rule	$(\text{response_options} - 1) / 2$

Max SD for whole scale

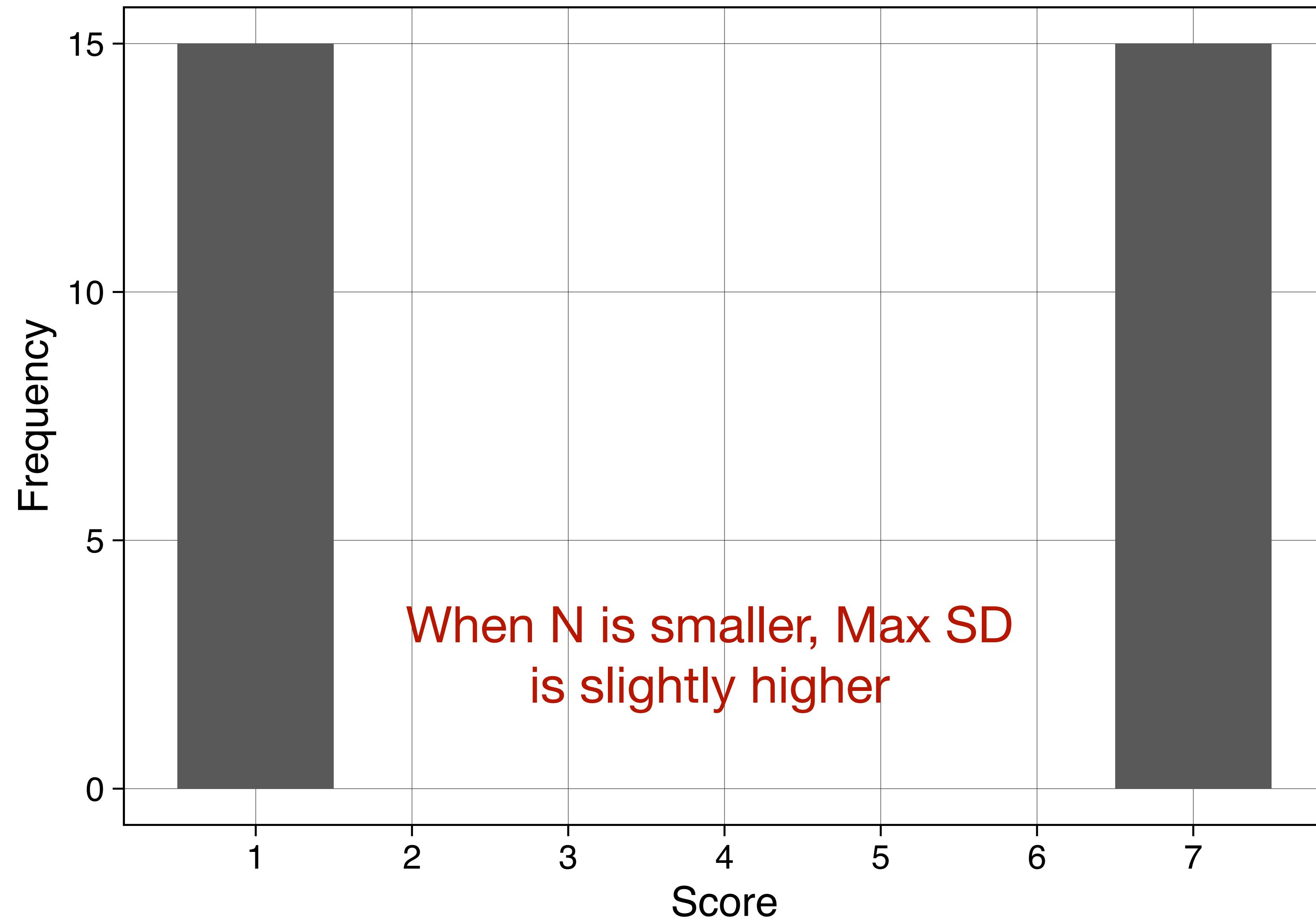
Example	Scale	[min, max]	SD
1	BDI	[0, 63]	13.49
2	BDI	[0, 63]	31.51
3	CES-D	[0, 60]	44.63
4	PHQ-9	[0, 27]	95.00
5	4-item Likert scale 1-7 response options mean scored	[1, 7]	3.11

Max SD for whole scale

Example	Scale	[min, max]	SD	
1	BDI	[0, 63]	13.49	Zemestani et al. 2020
2	BDI	[0, 63]	31.51	Ede et al. 2020
3	CES-D	[0, 60]	44.63	
4	PHQ-9	[0, 27]	95.00	Wright et al. 2022
5	4-item Likert scale 1-7 response options mean scored	[1, 7]	3.11	

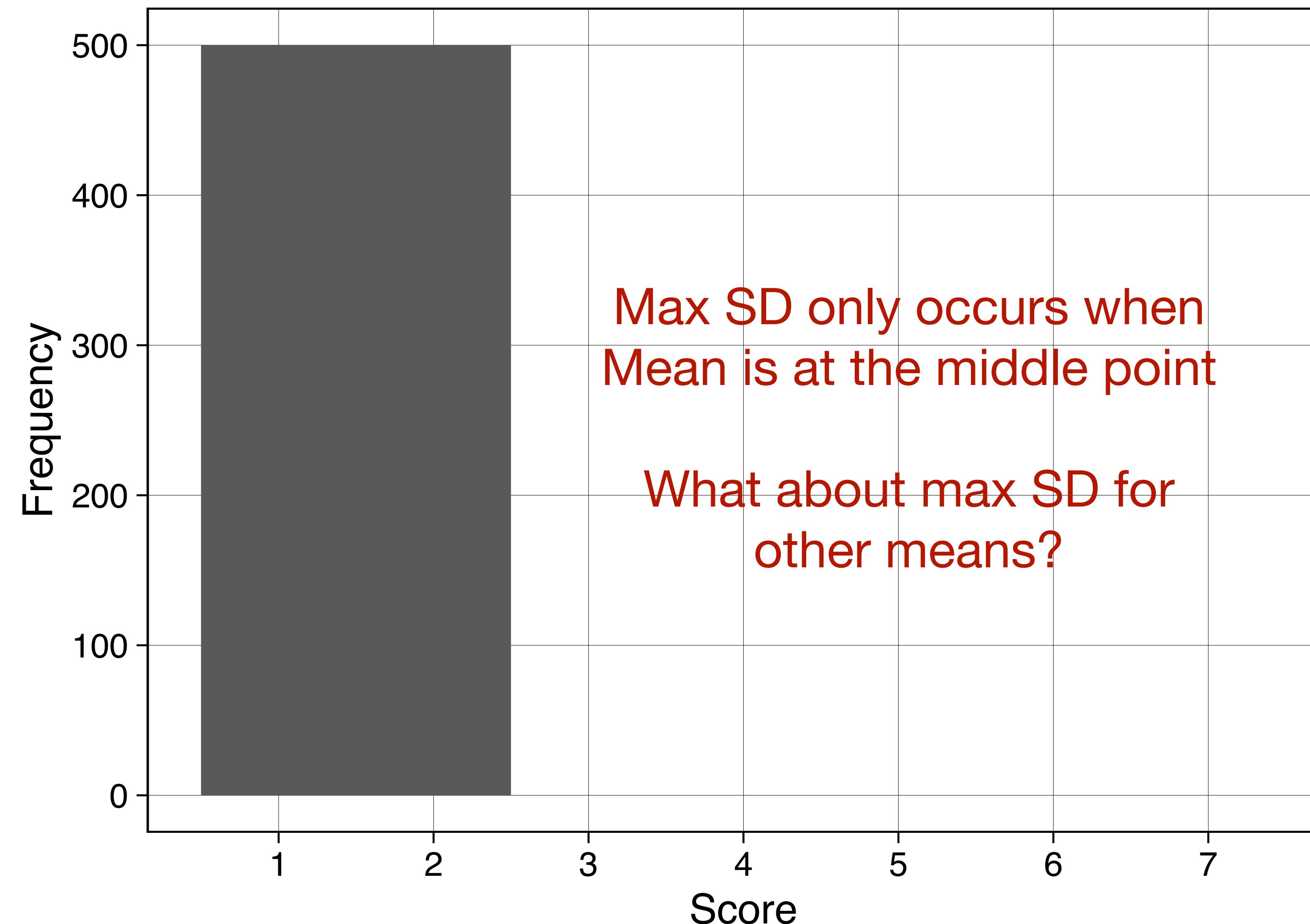
Max SD for whole scale

Max SD of 7-point scale:
 $M = 4.00$, $SD = 3.05$, $n = 30$



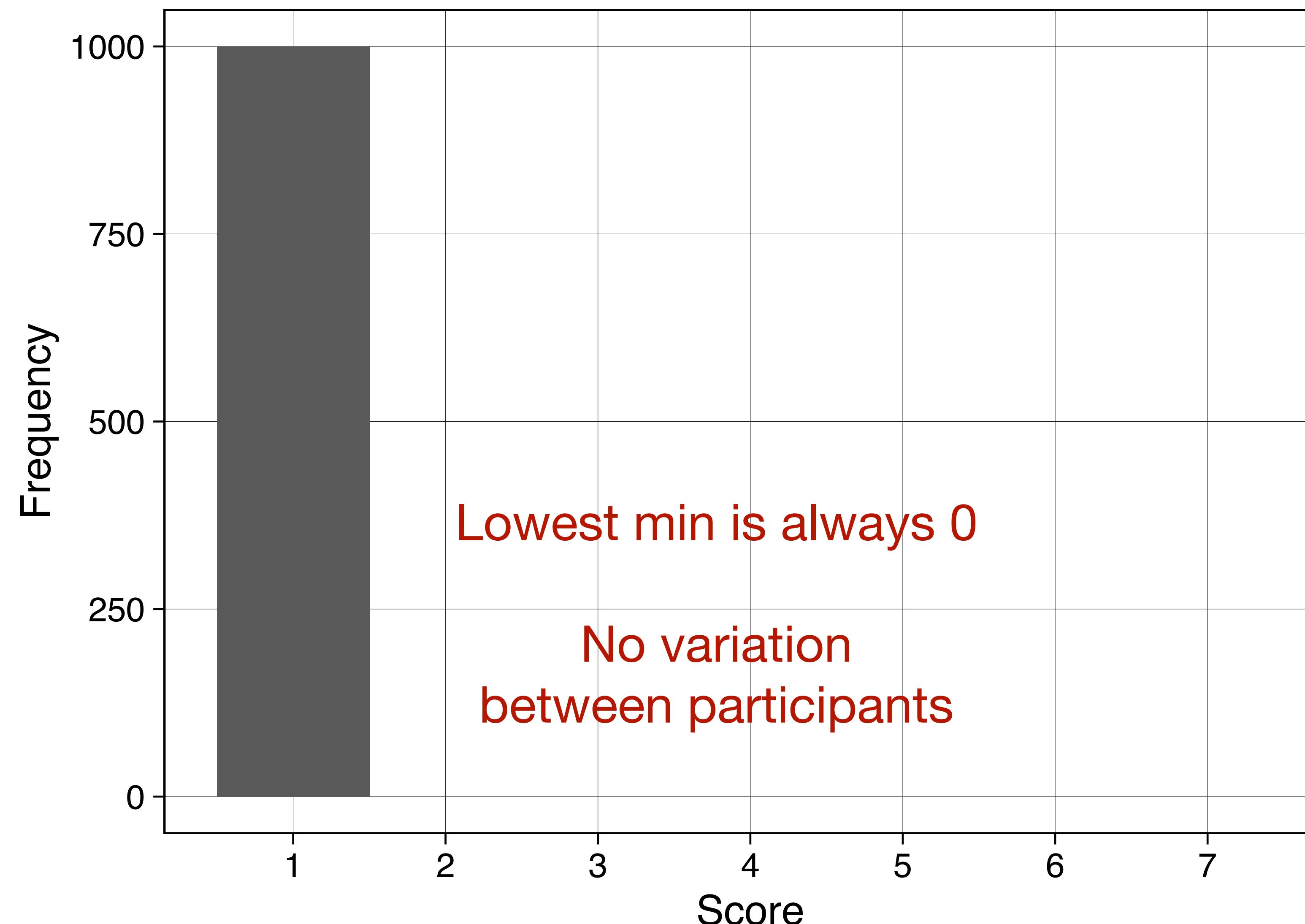
Max SD for a specific mean

Upper bound of min SD of 7-point scale:
 $M = 1.50$, $SD = 0.50$, $n = 1000$



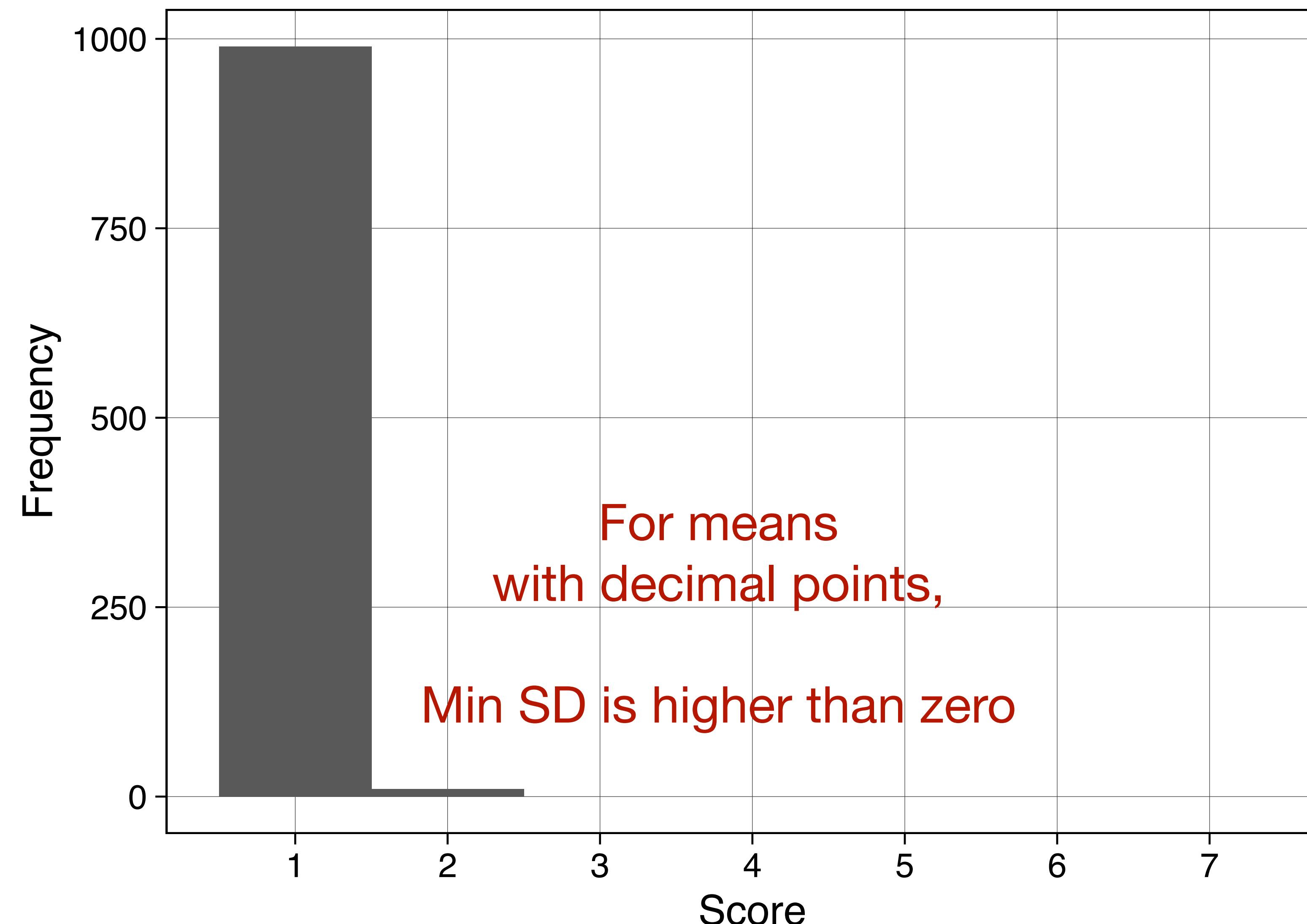
Min SD for whole scale

Lower bound of min SD of 7-point scale:
 $M = 1.00$, $SD = 0.00$, $n = 1000$



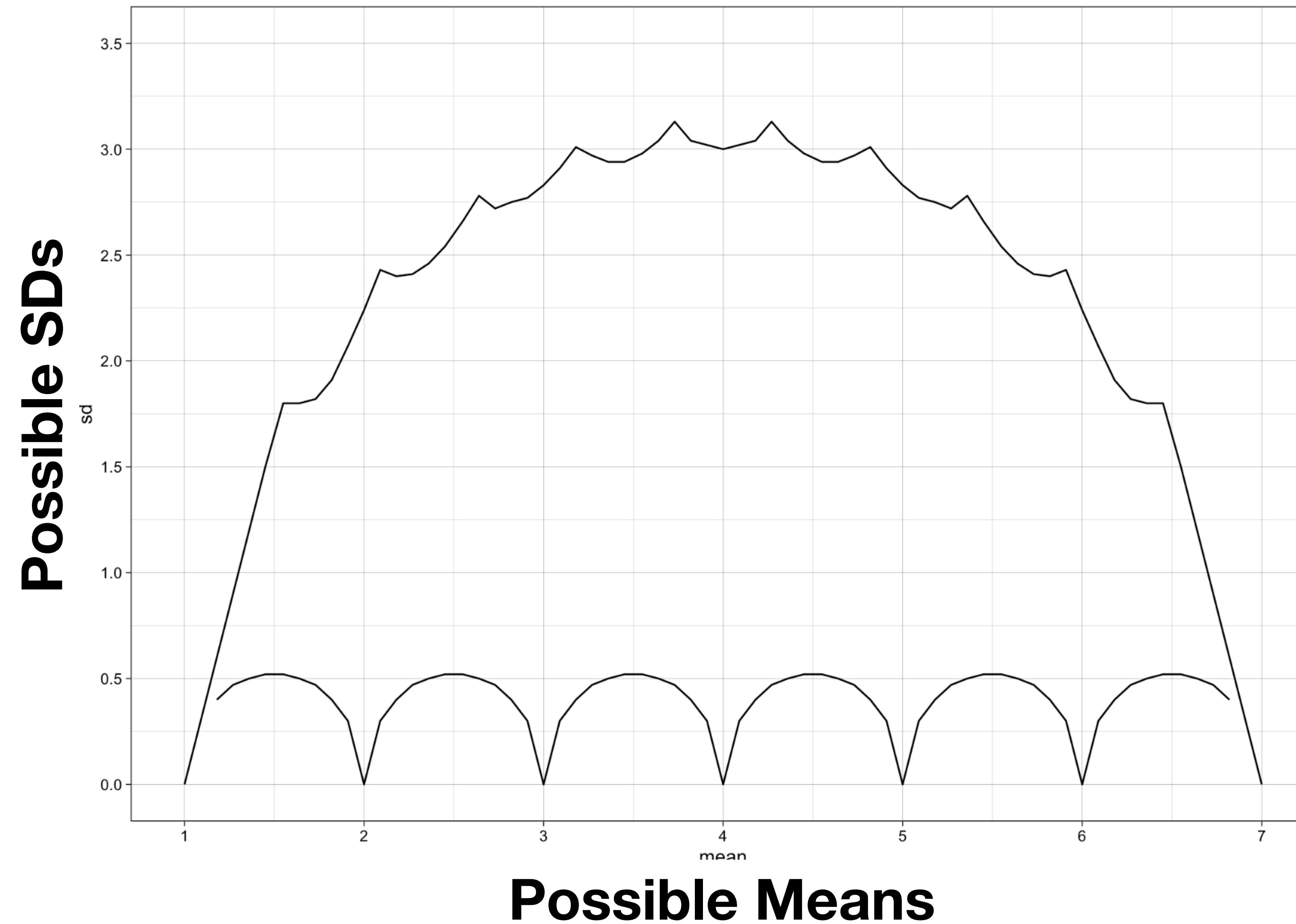
Min SD for a specific mean

Lower bound of min SD of 7-point scale:
 $M = 1.01$, $SD = 0.10$, $n = 1000$



TIDES

Only some means + SDs are possible



Calculate the
min and max SD
for every mean

I call this TIDES

For **any** N +
1-7 Likert scale

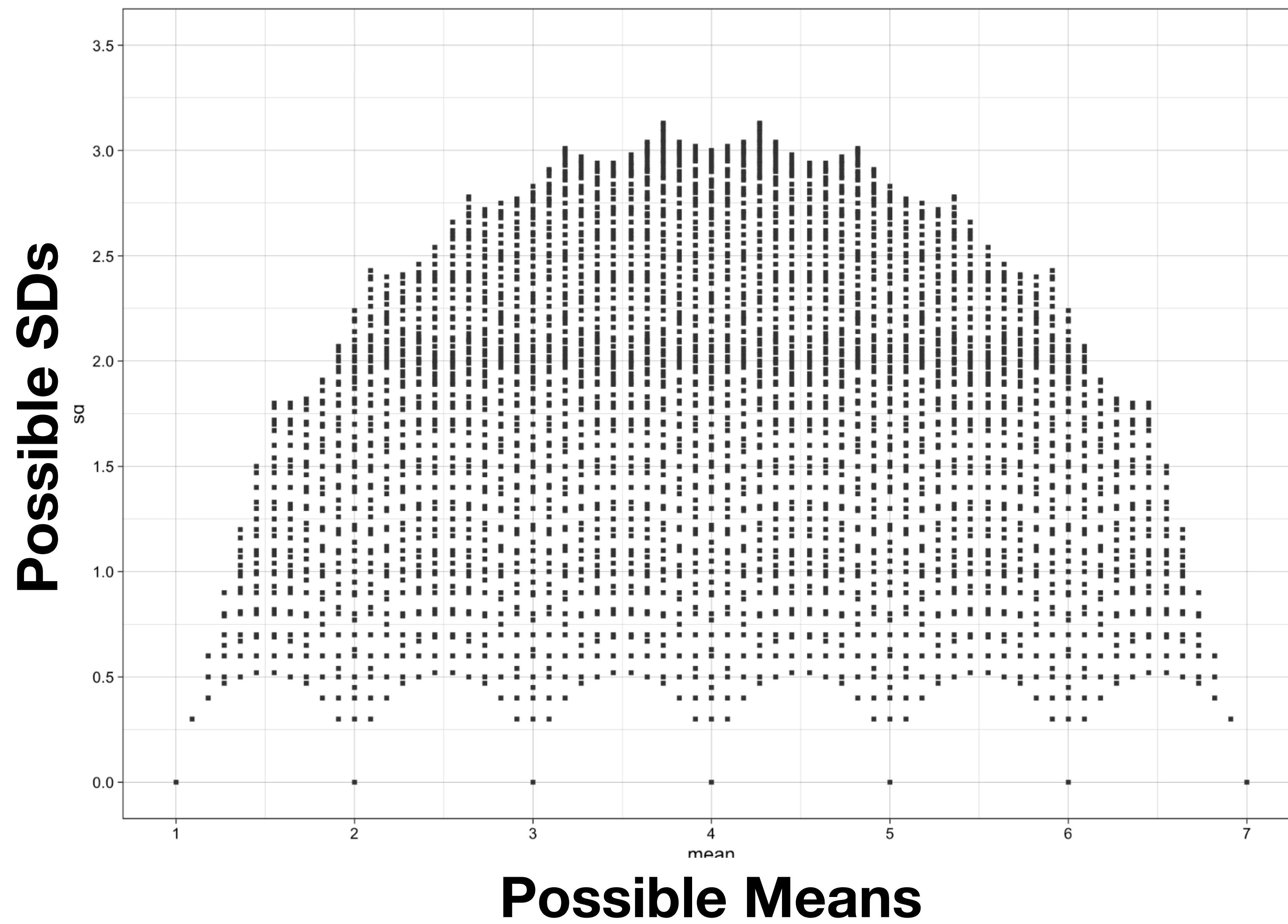
TIDES

Only some means + SDs are possible



GRIM + GRIMMER + TIDES

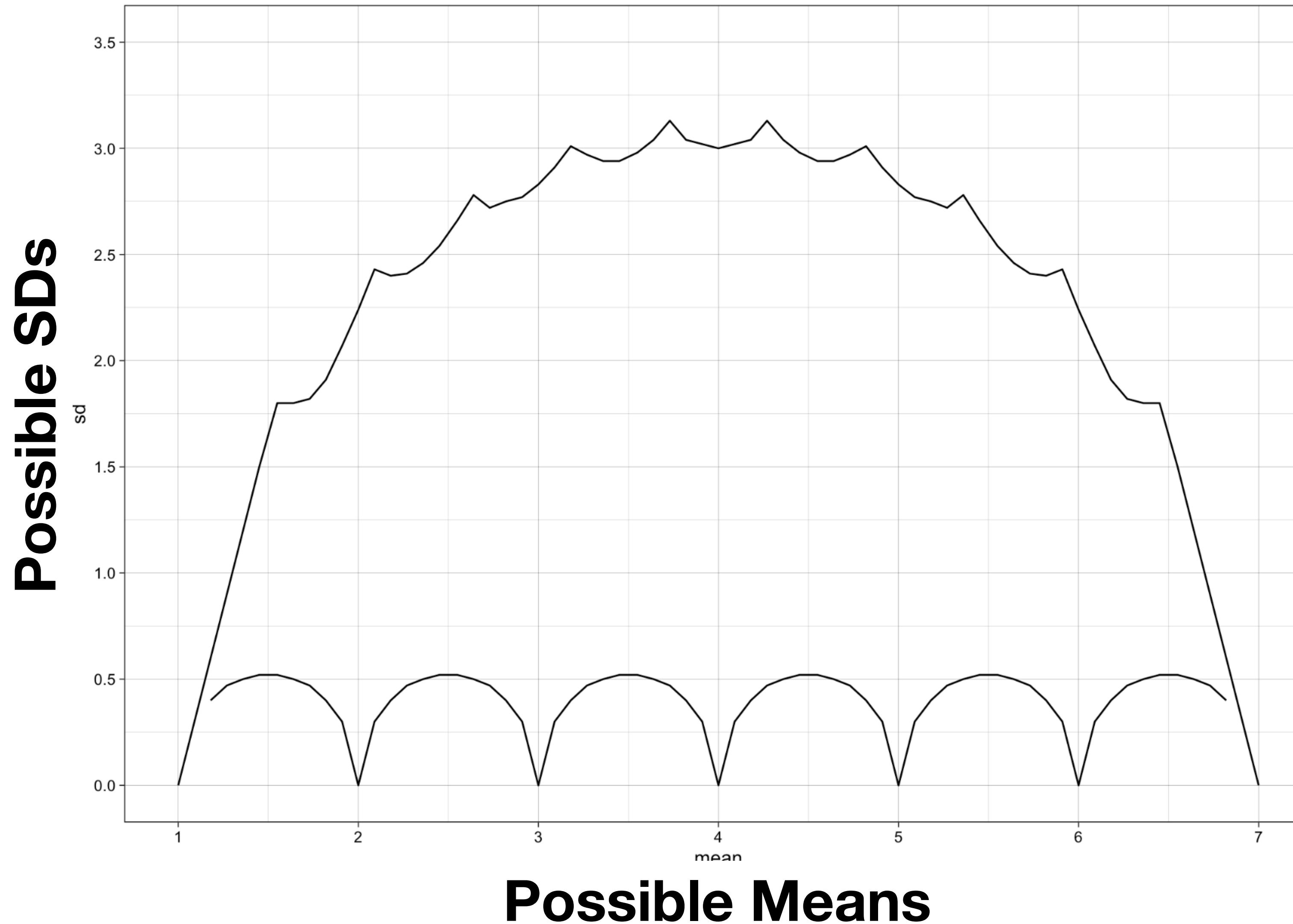
Only some means + SDs are possible



TIDES

Only some means + SDs are possible

Even when $N > 100$ and GRIM/MER stop working



TIDES

errors.shinyapps.io/TIDES

- What happens when:
 - N is very small?
 - N is very large?
 - Scale range is very small?
 - Scale range is very large?

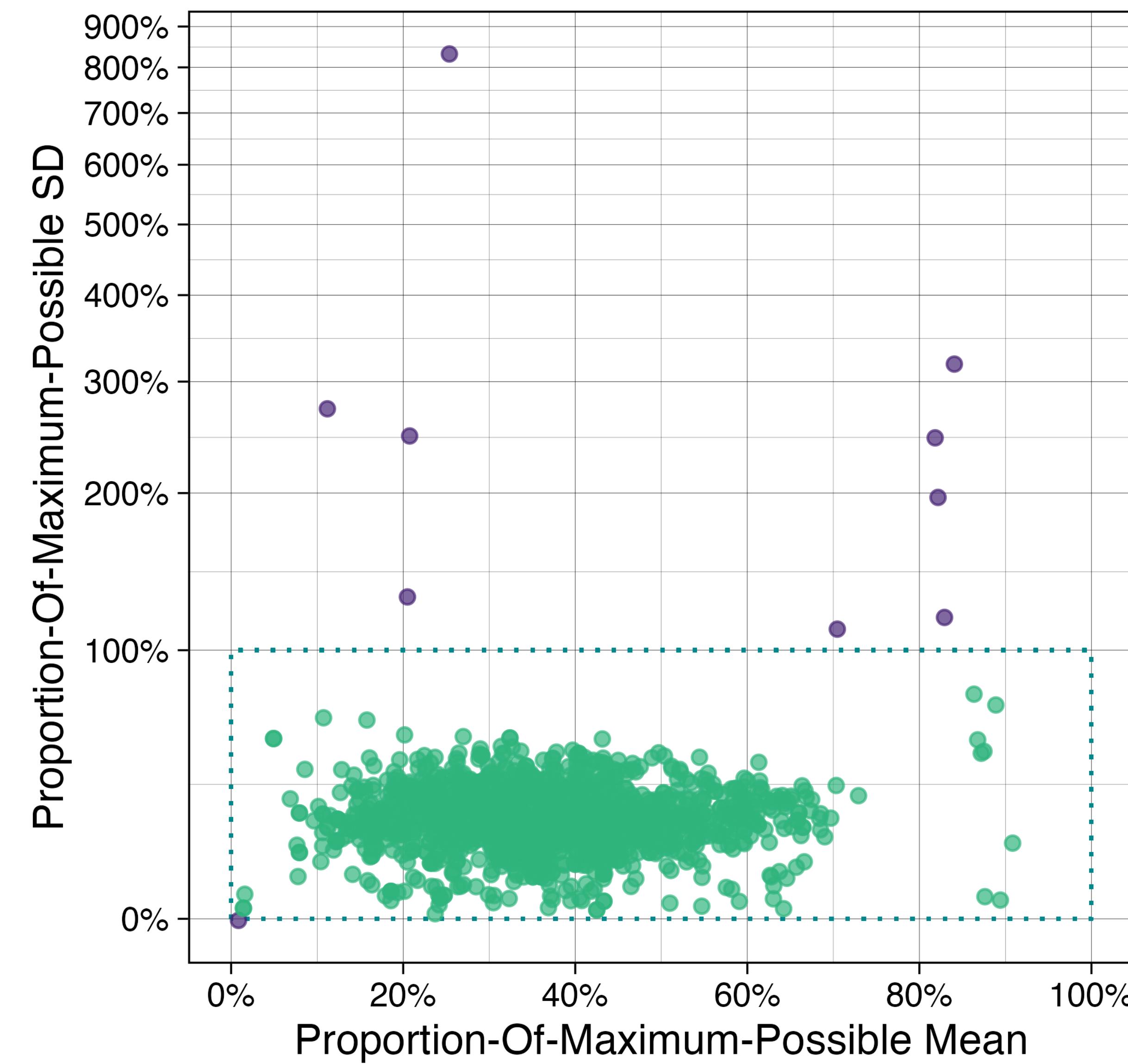
TIDES

errors.shinyapps.io/TIDES

- How can we improve our general intuitions for possible and plausible SDs?
 - Stretch the Umbrella plot to be a rectangle
 - Instead of units of Mean and SD, use units of Percent-Of-Maximum-Possible
 - Range 0% to 100%

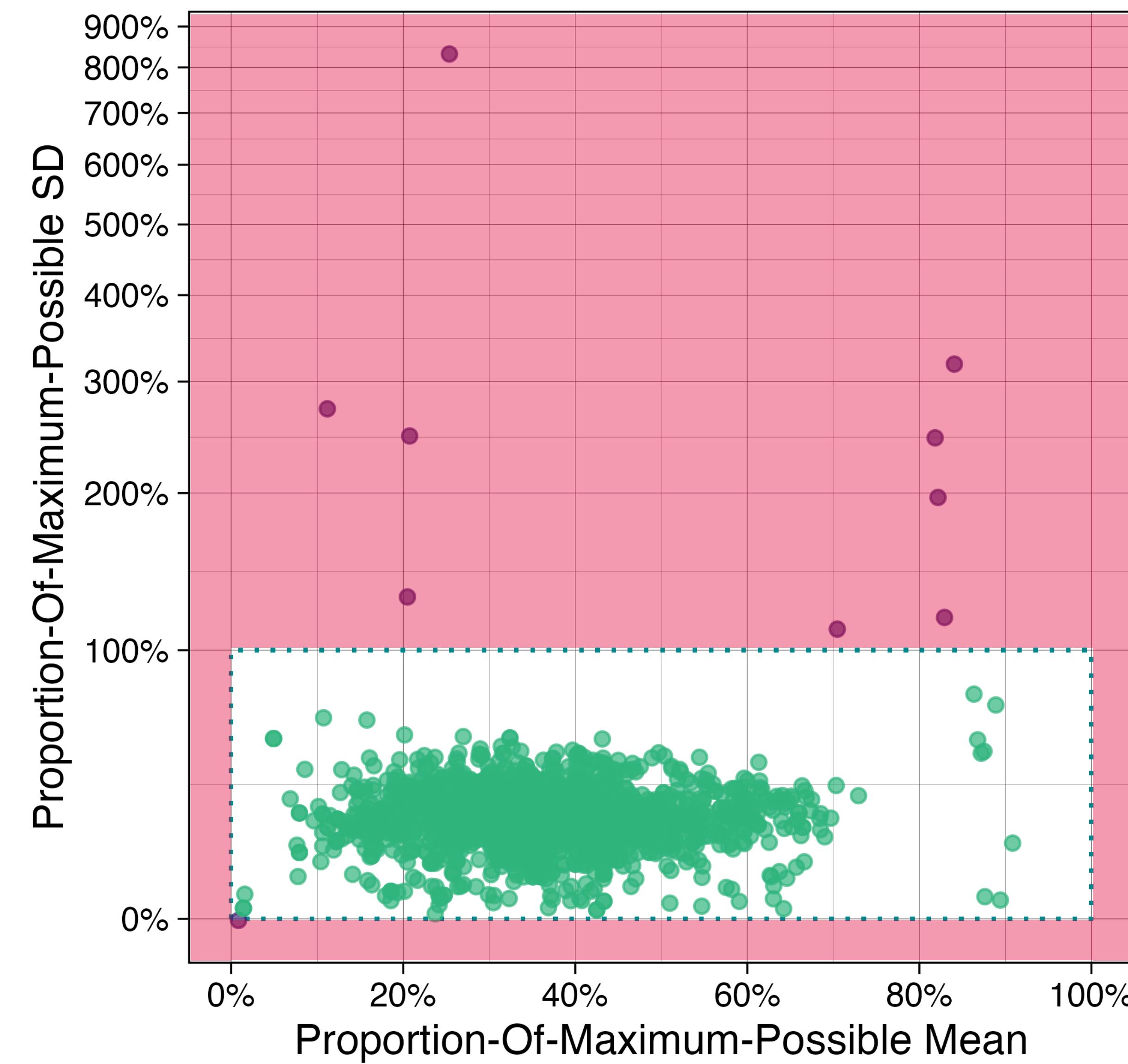
TIDES

Data points RCTs from
a large meta-analysis
database on
psychotherapy for
depression

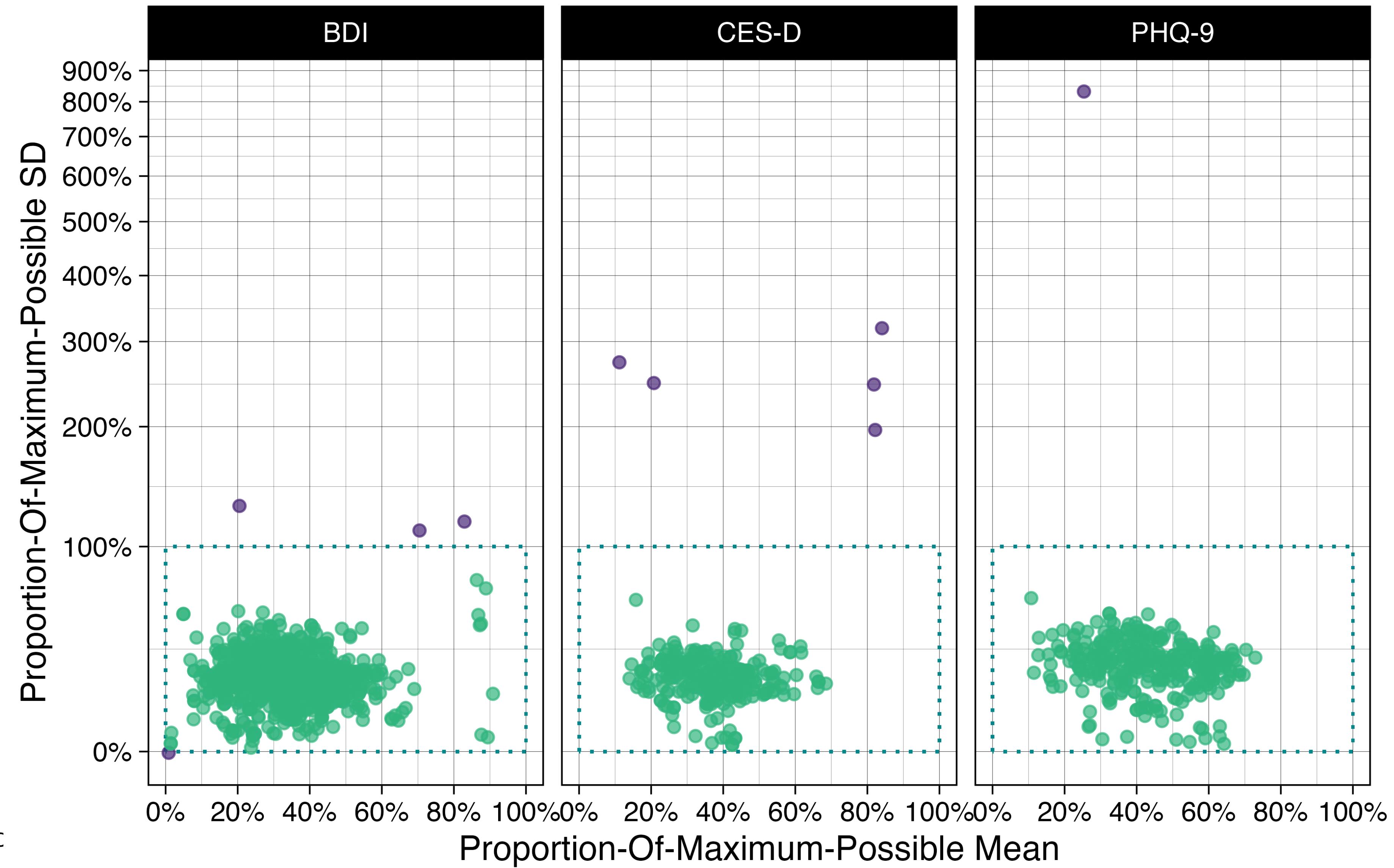


TIDES

Data points RCTs from
a large meta-analysis
database on
psychotherapy for
depression



TIDES



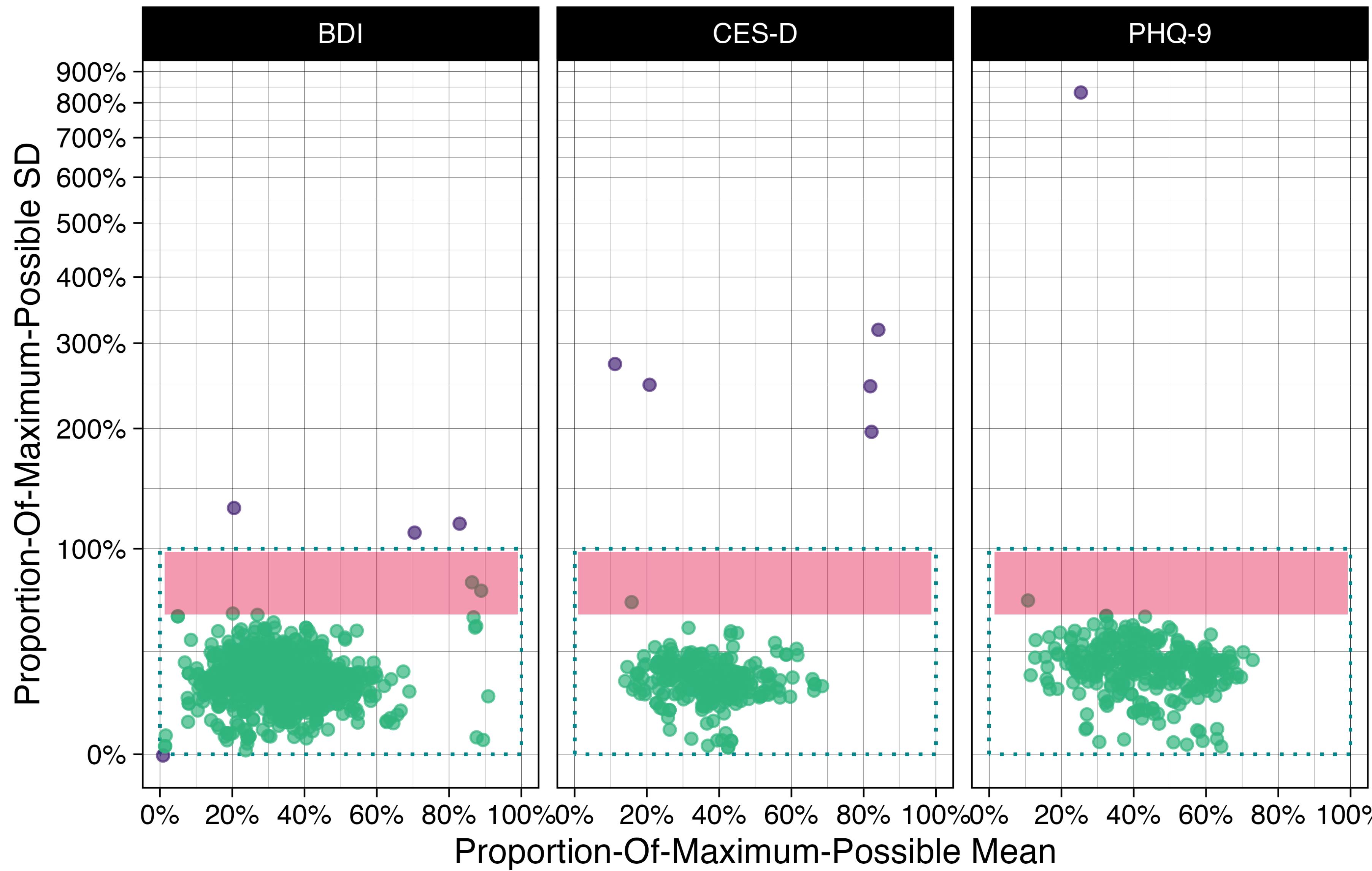
Inside possible ranges

- FALSE
- TRUE

TIDES

Few SDs close to max

Why?



TIDES

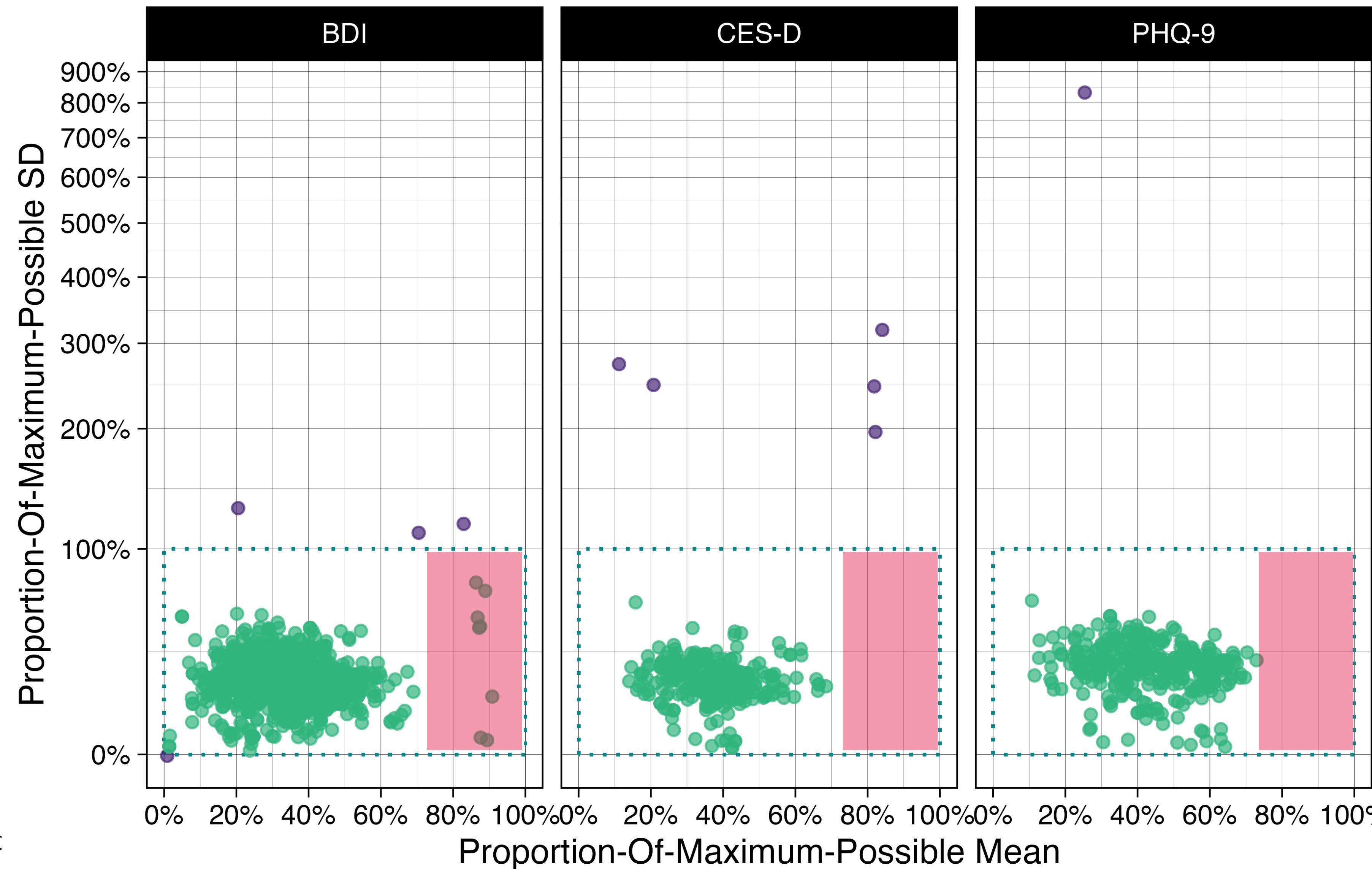
Few means close to right-hand max

Why?

Remember these are
depression scales

Inside possible ranges

- FALSE
- TRUE



Summary of interpretation of GRIM/MER & TIDES

Result fails GRIM/ER/TIDES

Meaning

Reported result is not possible

Causes

Typos

Errors

... *There was no dataset*

Result passes GRIM/ER/TIDES

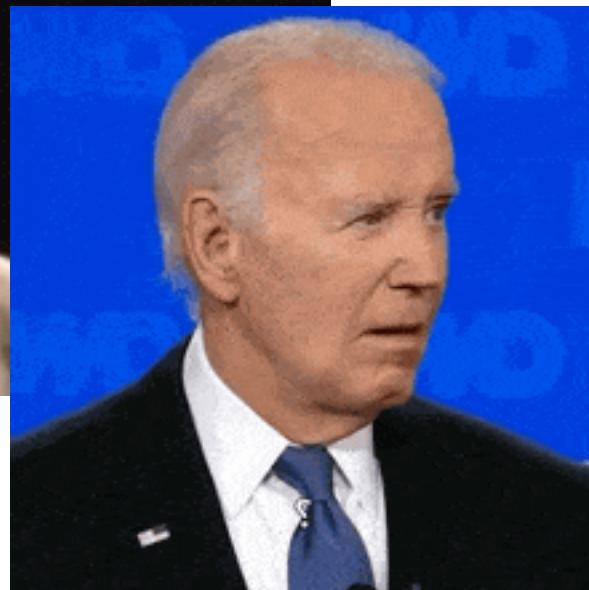
Meaning

Reported result is possible

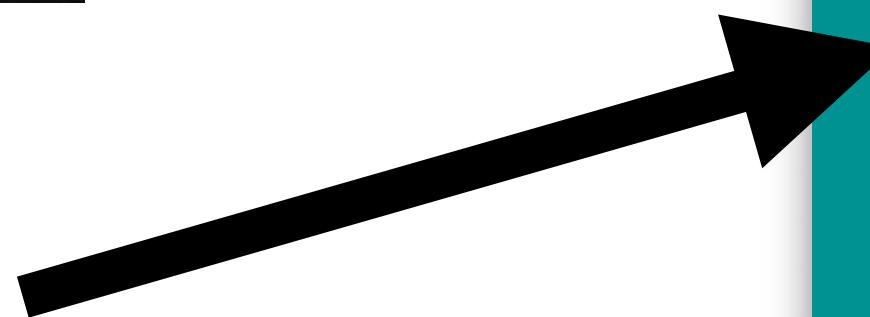
Causes

There was a dataset and summary stats were reported accurately

Summary of interpretation of GRIM/MER & TIDES



**Could this be bad?
How?**



Result passes GRIM/ER/TIDES

Meaning

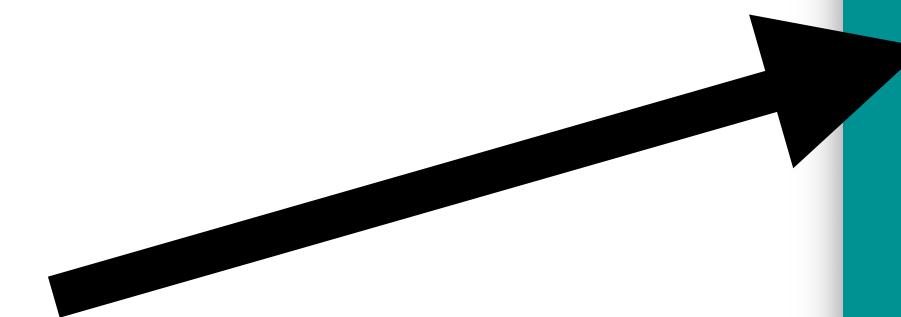
Reported result is possible

Causes

There was a dataset and summary
stats were reported accurately

Summary of interpretation of GRIM/MER & TIDES

**The underlying data
could be fake
or implausible**



Result passes GRIM/ER/TIDES

Meaning

Reported result is possible

Causes

There was a dataset and summary stats were reported accurately

SPRITE

Sample Parameter Reconstruction via Iterative TEchniques

When N is small, datasets can sometimes be reconstructed from summary stats

Maybe there are very few datasets that can produce the summary stats

Eg only one (1) dataset can generate

- M = 2.20
- SD = 2.53
- N = 10

```
## # A tibble: 10 × 2
##       solution score
##       <int>   <dbl>
## 1          1      1
## 2          2      1
## 3          3      1
## 4          4      1
## 5          5      1
## 6          6      1
## 7          7      1
## 8          8      1
## 9          9      7
## 10        10      7
```

SPRITE

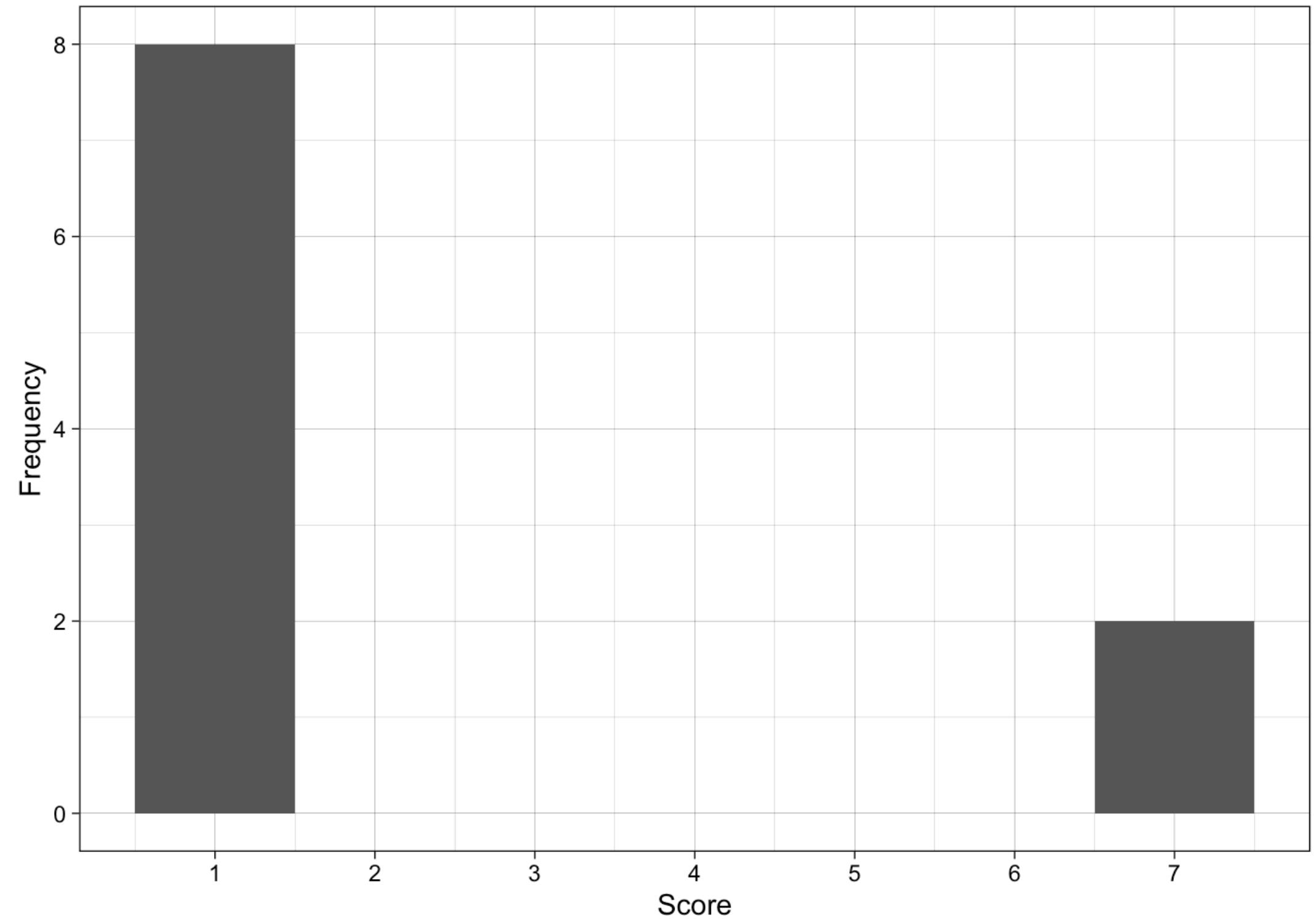
Sample Parameter Reconstruction via Iterative TEchniques

When N is small, datasets can sometimes be reconstructed from summary stats

Maybe there are very few datasets that can produce the summary stats

Eg only one (1) dataset can generate

- M = 2.20
- SD = 2.53
- N = 10



SPRITE

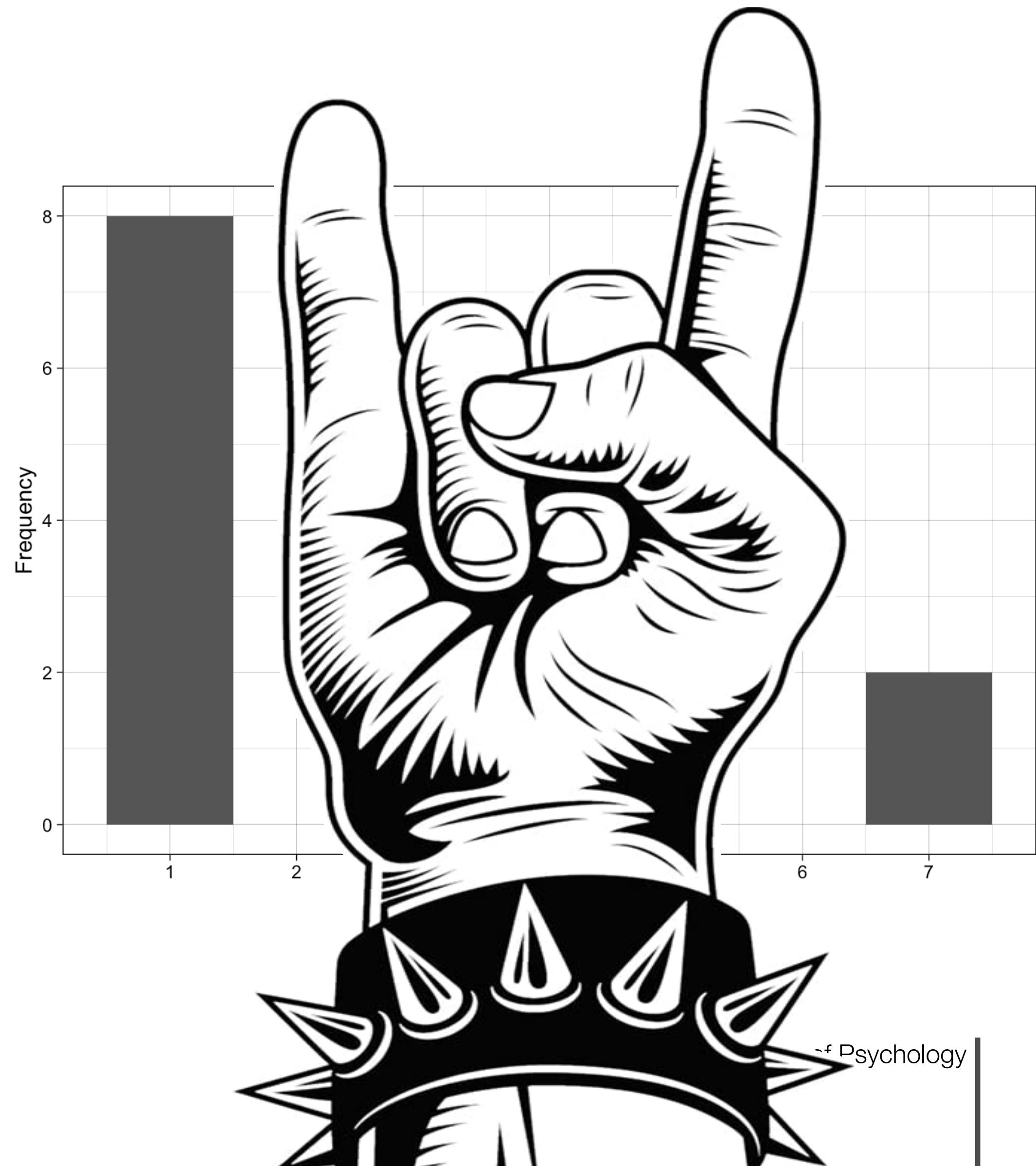
Sample Parameter Reconstruction via Iterative TEchniques

This data is very unlikely to have a legitimate and natural explanation!

‘horns of no confidence’ (Heathers, 2017)



Dr. Ian Hume



~f Psychology

SPRITE

Sample Parameter Reconstruction via Iterative TEchniques

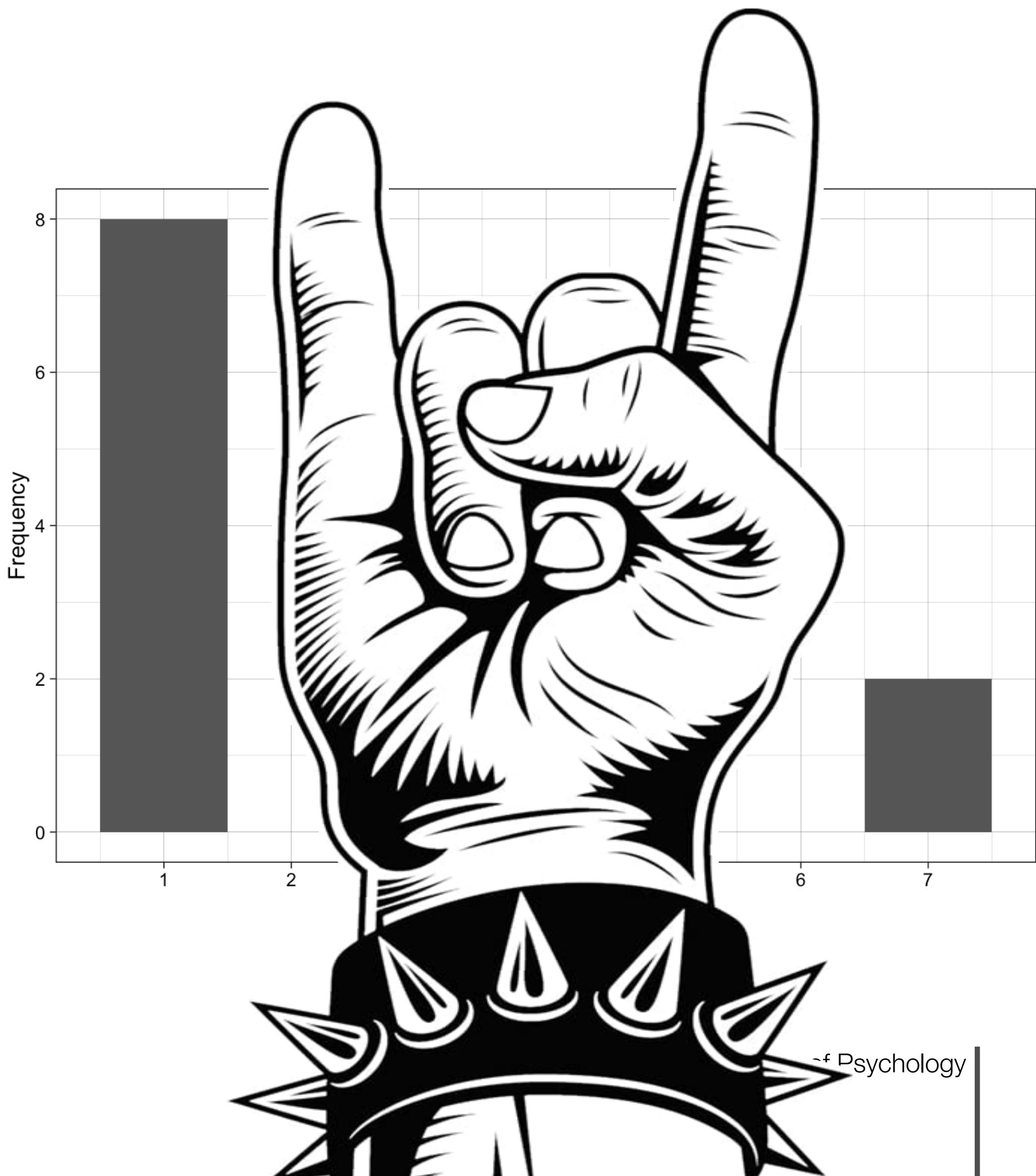
This data is very unlikely to have a legitimate and natural explanation!

‘horns of no confidence’ (Heathers, 2017)

LinkedIn edition



Di



~f Psychology

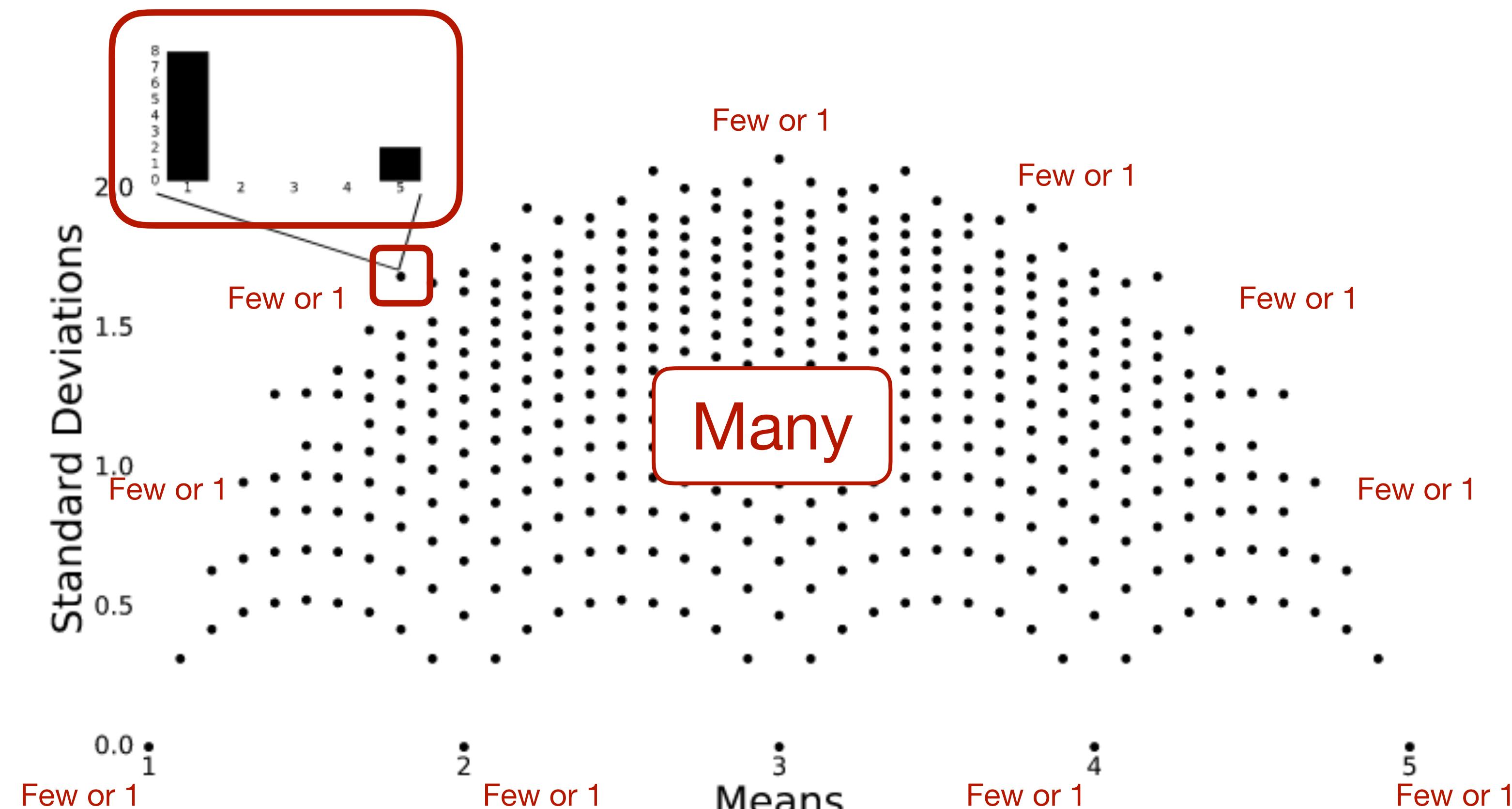
SPRITE

Sample Parameter Reconstruction via Iterative TEchniques

Umbrella plot tells you about **how many** datasets can generate a given set of summary statistics

- Near the edge: few or just 1
- Near the centre: many

*But because SPRITE is more complicated,
we won't use it in the assignment*





{rant}
Foody et al. 2018

Readings

- Relevant to TIDES
 - Heathers (2017) The Umbrella Graph — Connecting GRIM and SPRITE <https://jamesheathers.medium.com/sprite-interlude-the-umbrella-graph-connecting-grim-and-sprite-also-brunch-sucks-1266c629c974>
 - Heathers et al. (2018) Recovering data from summary statistics: Sample Parameter Reconstruction via Iterative TEchniques (SPRITE). <https://doi.org/10.7287/peerj.preprints.26968v1>

Assignment

No assignment this week

Questions?