

# IS UN-UNLIKE NOT, OR NOT?

Xander MacSwan



Clara McMahon

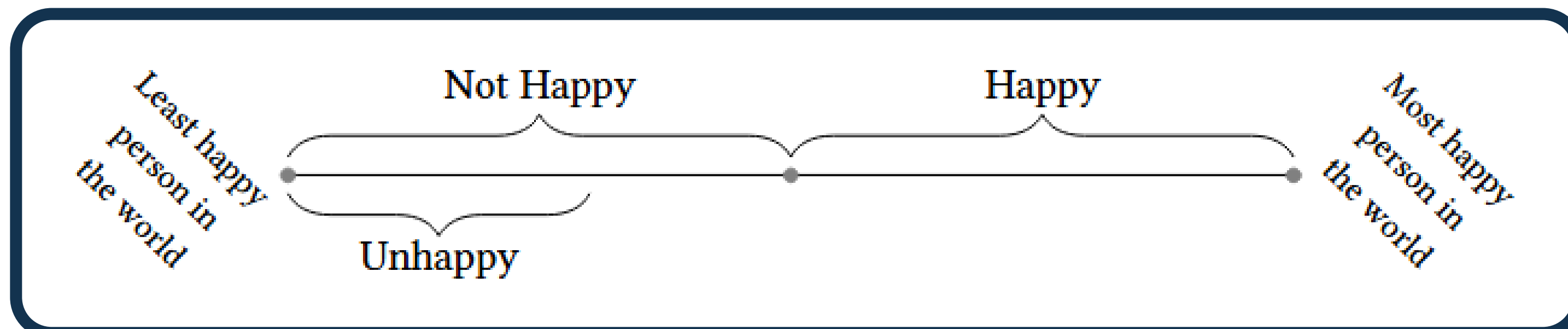


## THE STRENGTH OF UN-

Theorists like Horn (1989; 2020) and Krifka (2007) have suggested a difference in meaning between sentences like (1) and (2):

- (1) Sara is not happy.
- (2) Sara is unhappy.

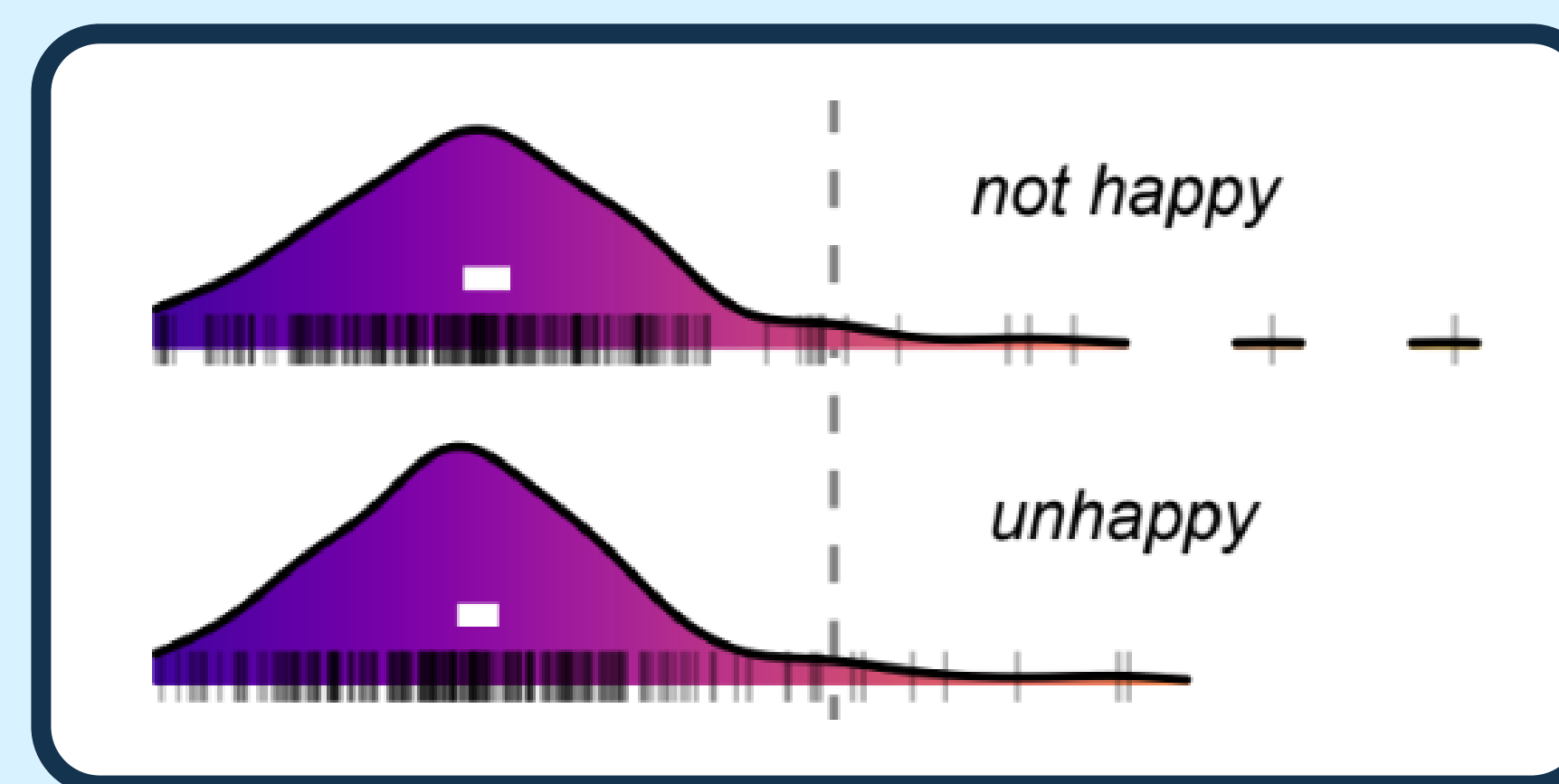
At least in comparison, (2) is thought to be *stronger* than (1) - it suggests a less happy person. Represented on scale, this might look like:



## NEGATION OUT-OF-THE-BLUE

Contra Horn, Tessler & Franke (2019) found that in out-of-the-blue uses, *un-* and *not* are interpreted identically. This suggests that any difference between *un-* and *not* is pragmatic, rather than semantic, in nature.

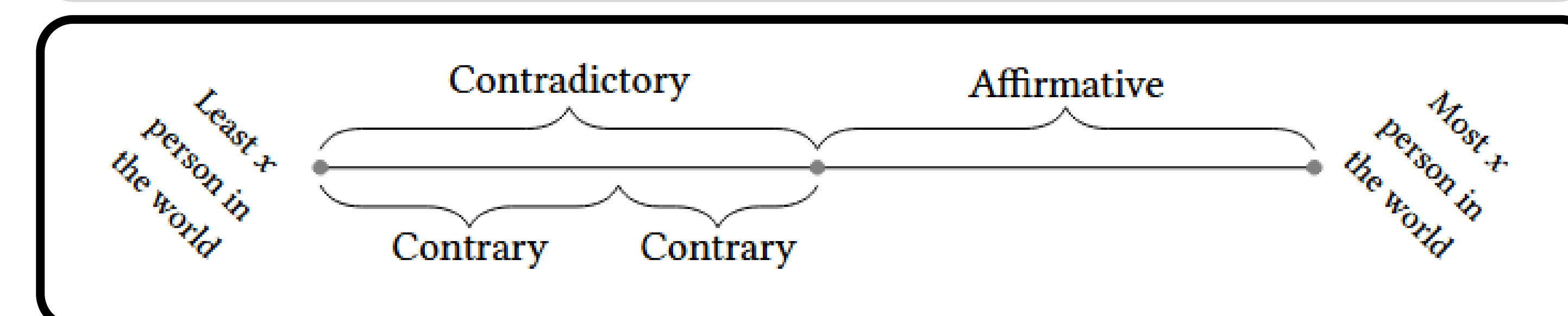
- Within subjects design, n=114
- **Negated Positive & Morphological Antonyms** not significantly different
- $\beta = -0.007$ ,  $t(17) = -0.39$ ,  $p = 0.701$



## THE CONTRARIETY CONFOUND

We identified a confound in Tessler & Franke's results, relying on a pesky feature of theorizing *un-* as contrary negation. Strictly speaking, all it takes to count as a **contrary**, in range terms, is to designate some subset of the contradictory range. Any arbitrary subset counts as a contrary.

So, if subjects select a random contrary range, and a random point within it, the output would be indistinguishable from **contradictory** judgments.



## CONTRARY VS CONTRADICTORY

Horn proposes that this difference in meaning is due to underlying different semantic contributions of *un-* and *not*. Where *not* is thought to function as the usual, **contradictory** (complement) **negation**, *un-* is thought to function as **contrary negation**. These are defined as follows:

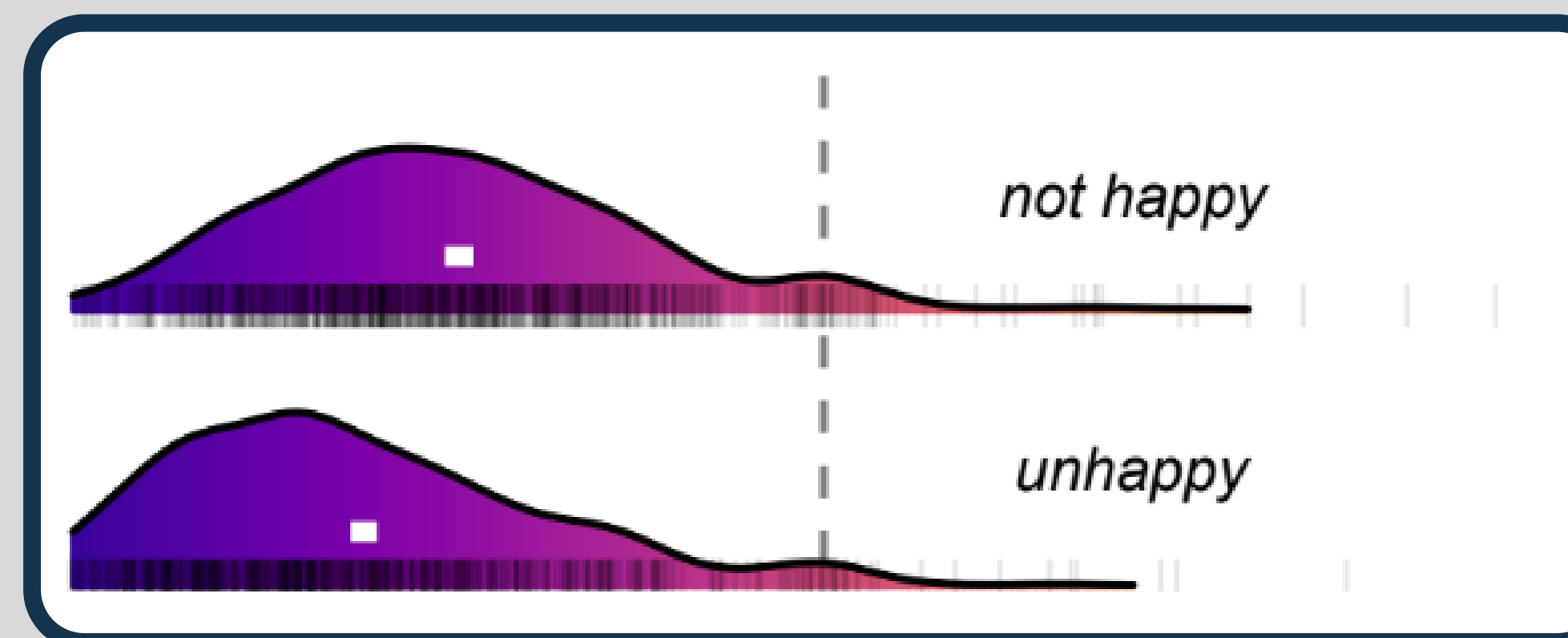
- Intuitively, happy and not happy exhaust the scale, but happy and unhappy leave room between - so someone could be neither.
- On a scale like the one above, a contrary of *P* just occupies any subset of the area occupied by the contradictory of *P*.

- *P* is the contradictory of *Q* iff
  - ***P* and *Q* must be false**, and
  - ***P* or *Q* must be true**.
- *P* is a contrary of *Q* iff
  - ***P* and *Q* must be false**.

## CONTEXT & COMPARISON

Tessler & Franke also found that the expected contrast between *un-* and *not* does emerge when subjects judge sentences of both types at the same time (presented on-screen simultaneously).

- Difference between **Negated Positive & Morphological Antonyms** significantly greater in **Multiple Utterances Condition** than in **Single Utterance Condition**.
- Between subjects design, n=715
- **Single Utterance**:  $\beta = 0.063$ ,  $t(6457) = 6.73$ ,  $p < 0.001$
- **Multi Utterance**:  $\beta = -0.064$ ,  $t(268) = -9.27$ ,  $p < 0.001$



## MEASURING RANGES

We aim to address this confound by soliciting **range judgments**, rather than **point judgments**, from participants. This will allow us to detect some of the precise signatures of contrary negation: if subjects select narrower ranges judging *un-* sentences, then *un-* might be a kind of contrary negation after all.

- More generally, measuring ranges should give us *strictly* more precise data compared to points.
- Once we have results for English, we plan to extend to Spanish (non-strict NC) and Polish (strict NC). Stay posted!

Your friend tells you about their friend: James.

"James is unhappy."

Where could James possibly be placed on the following scale?

# FINDINGS

### Design

- 2:1 ratio fillers to test items
  - Fillers included degree modifiers (*slightly, sort of, very, etc*)
- Four pseudorandomized lists
- 20 test items (10 *un-*, 10 *not*)
  - Negation type within participants, across lists

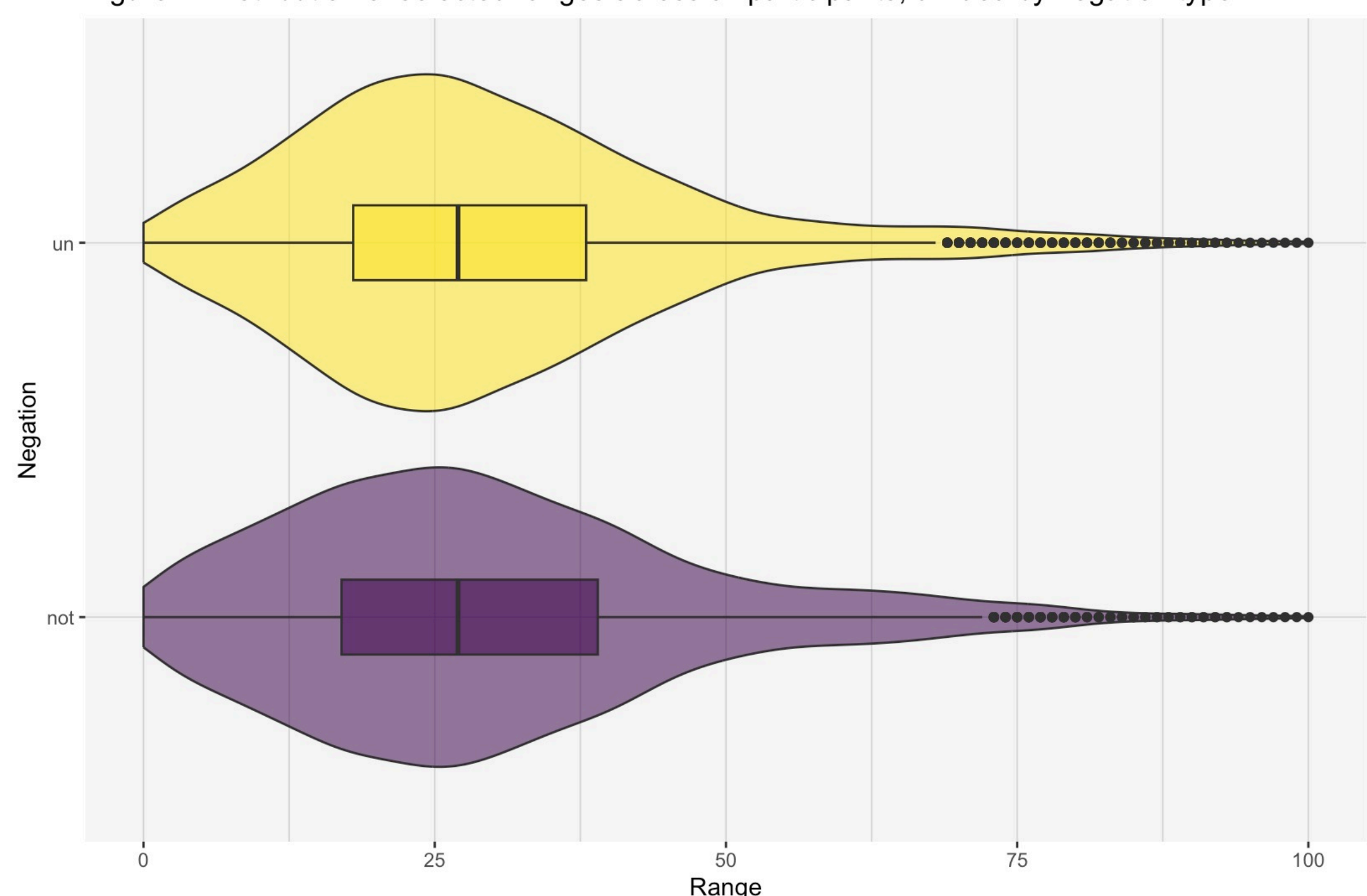
### Participants

- 46 Native English speakers, recruited via Baruch SONA
- 16 Excluded (failed knowledge check/skipped negated adjs.)
- Of 30 analyzed:
  - Ages 18-27 (mean 20.4)
  - Gender 18F 12M

### Procedure

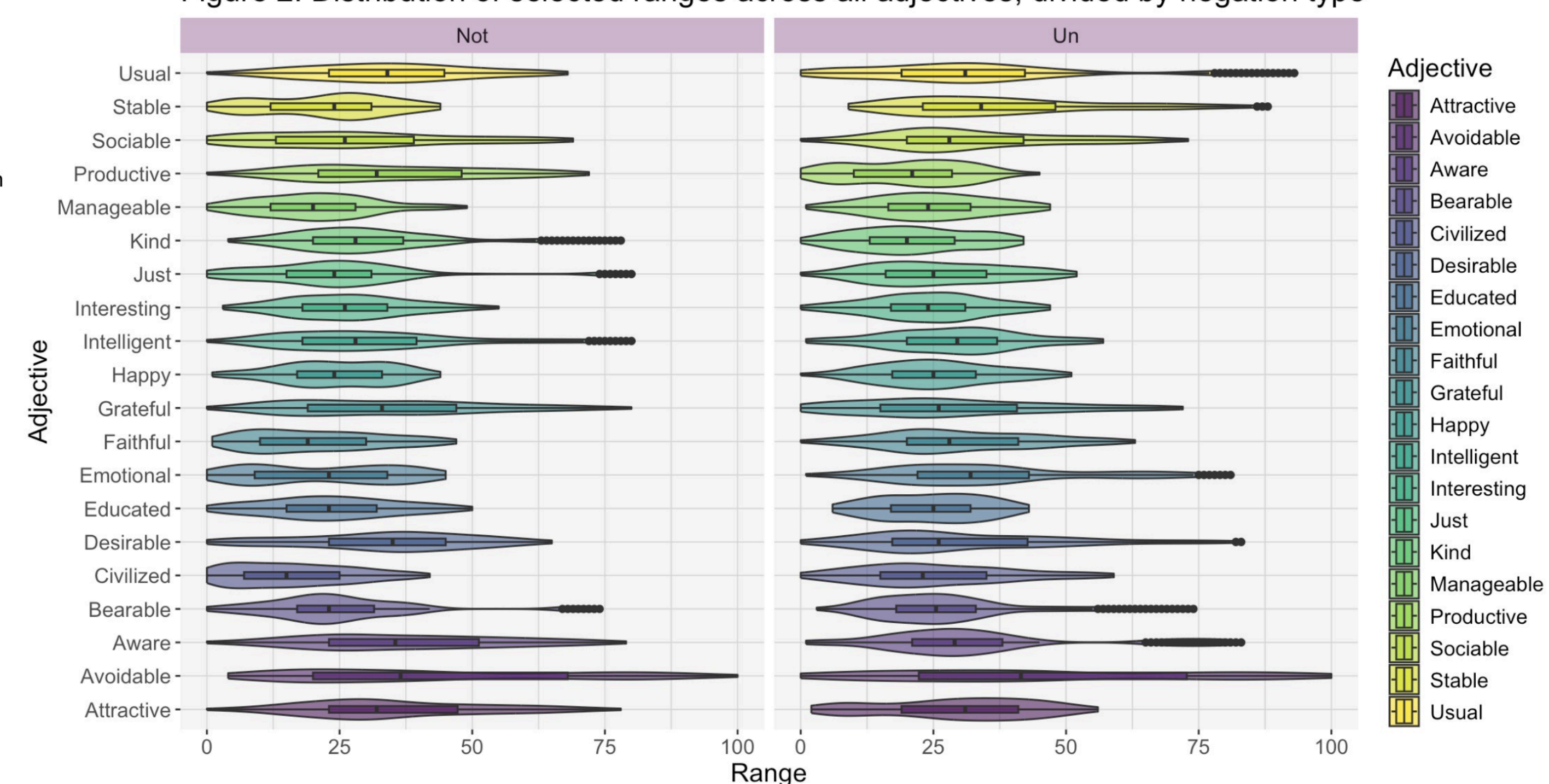
- Demographic Qs → 2
- Training Items → Knowledge Check → Survey
- Survey included 60 total items and took 15 minutes

Figure 1: Distribution of selected ranges across all participants, divided by negation type



Mixed effects logistic regression fitted with lme4, negation sum coded (with un: 1 and not: -1)

Figure 2: Distribution of selected ranges across all adjectives, divided by negation type



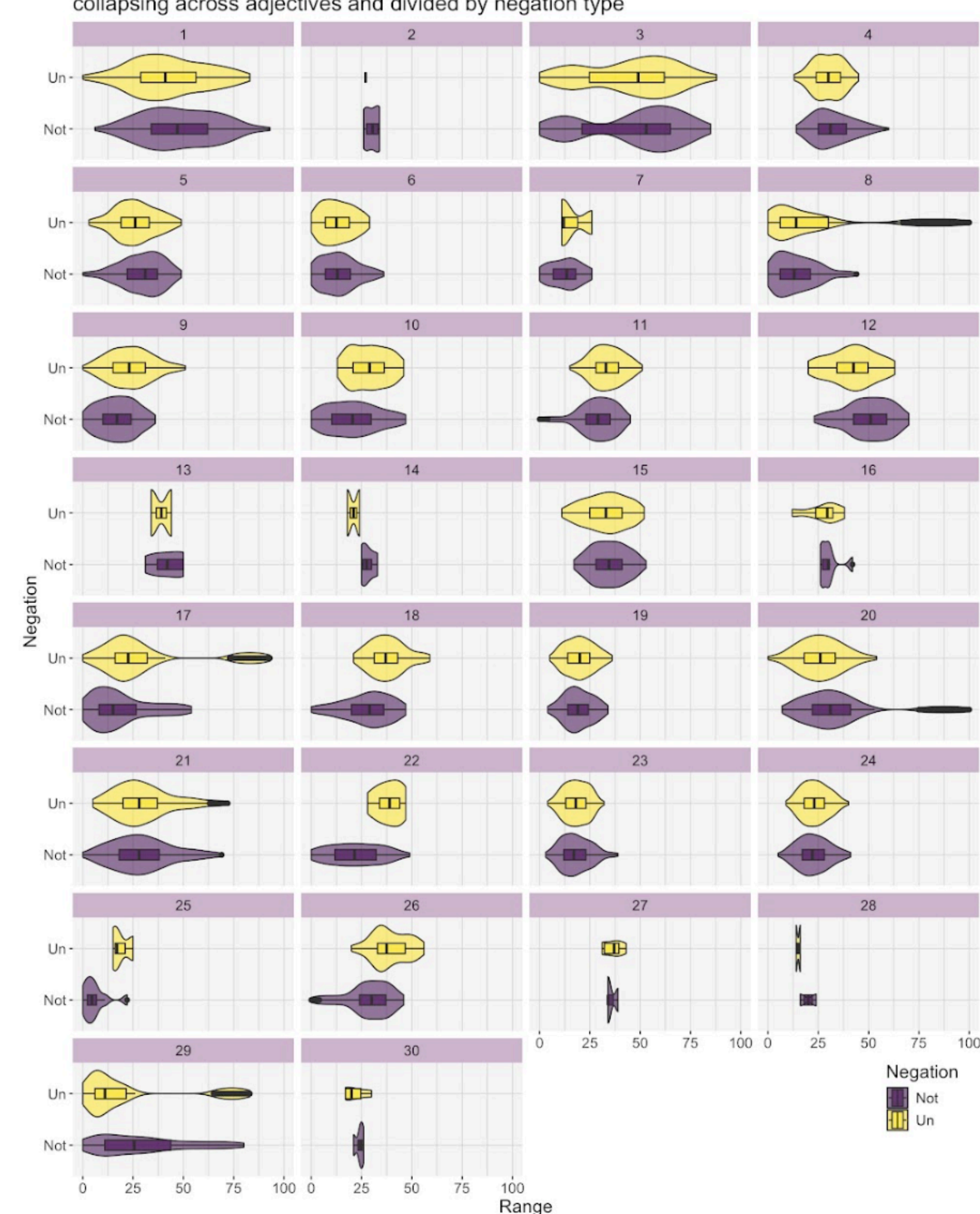
Predictors	Midpoint				Width			
	Estimates	95% CI	p	df	Estimates	95% CI	p	df
(Intercept)	27.83	24.48 - 31.18	<0.001	420.00	14.34	10.26 - 18.41	<0.001	420.00
Negation (sum)	0.14	-0.95 - 1.23	0.802	420.00	-0.28	-1.08 - 0.53	0.502	420.00
<b>Random Effects</b>								
$\sigma^2$	127.82			68.57				
$\tau_{00}$	69.53	ResponseId		114.60	ResponseId			
	4.24	adjective		5.54	adjective			
ICC	0.37			0.64				
N	30	ResponseId		30	ResponseId			
	20	adjective		20	adjective			
Observations	425			425				
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.000 / 0.366				0.000 / 0.637			

### Levene's Test for Homogeneity of Variance in Range Width

	Un vs. Not negation		F	p
	Variance (Un)	Variance (Not)		
Range Width	154.4	178.47	0.597	0.44

Variance reflects the spread of range widths (rangeMax - rangeMin) across participants.

Figure 3: Distribution of selected ranges by individual, collapsing across adjectives and divided by negation type



## DISCUSSION

Preliminary results suggest that we are successfully replicating Tessler & Franke, ruling out the contrariety confound and increasing the evidence that *un-* and *not* are not semantically distinct. This is significant not only for theories of negation in linguistics, but also in cognitive science more broadly. It undermines arguments to the effect that contrary negation is the *proto-concept* from which contradictory negation is built (Feiman et al, 2022)

This is, however, a primarily negative result; Horn & Krifka are wrong, but what explains the apparent difference in interpretation when *un-* and *not* are considered in comparison? The mechanism must be context sensitive, but providing a satisfying explanation remains an open problem.

For now, this result is also totally limited to English. Whether the pattern holds cross-linguistically, and in particular across variation in negative concord, is our next experiment!

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*.  
 Feiman, R., Mody, S., & Carey, S. (2022). The development of reasoning by exclusion in infancy. *Cognitive Psychology*, 135, 101473.  
 Horn, L. R. (1989). A natural history of negation. Center for the Study of Language and Information (CSLI) Publications.  
 Horn, L. R. (2020, March). Negation and opposition: Contradiction and contrariety in logic and language. In *The Oxford handbook of negation*. Oxford University Press.  
 Krifka, M. (2007). Negated antonyms: Creating and filling the gap. In U. Sauerland & P. Stateva (Eds.), *Presupposition and implicature in compositional semantics* (pp. 163-177). Palgrave-Macmillan.  
 R Core Team. (2025). R: A language and environment for statistical computing. R Foundation for Statistical Computing.  
 Tessler, M. H., & Franke, M. (2019). Not unreasonable: Why two negatives don't make a positive