Appendix

Table A1

Response pattern for each multiverse option, expressed in percentages (N = 56)

Options	Appropriate	Not appropriate	Don't know
Exclusion Age			
Remove participants younger than 18	84	12	4
Keep participants regardless	36	55	9
Exclusion Language			
Remove non-native speakers	89	7	4
Keep participants regardless	27	70	4
Exclusion Multimodal			
Exclude participants with multimodal RT distribution according to Silverman's test	32	18	50
Keep participants regardless	50	21	29
Exclusion Number of Trials			
Remove participants with fewer than 100 trials	80	20	0
Keep participants regardless	48	50	2
Exclusion Same Responses			

Remove participants who always use the same response button	93	7	0
Keep participants regardless	9	88	4
Exclusion Alternating Responses			
Remove participants who always alternate responses after every trial (word, nonword, word, nonword,. . .)	89	5	5
Keep participants regardless	16	75	9
Exclusion Participants Accuracy			
Across trials: Participants with an error rate above 10% removed	36	59	5
Across trials: Participants with an error rate above 20% removed	46	48	5
Across trials: Participants with an error rate above 40% removed	71	25	4
Across trials: Participants with more than 70% of the trials being errors or time-outs are removed	89	7	4
Across trials: Calculate each participant's accuracy and remove those whose accuracy is 3 SD below the mean	84	12	4
Across trials: Calculate each participant's accuracy and remove those whose accuracy is more than three scaled MAD above and below the median accuracy, with scaled MAD defined as c*median(abs(accuracy-median(accuracy))), where c=-1/(sqrt(2)*erfcinv(3/2))	55	12	32
For nonwords: Participants with an error rate above 25% removed	55	38	7

For words: Participants with an error rate above 25% removed	59	36	5
Per lexical status (words vs. nonwords): Participants with an error rate above 30% for either lexical status are removed	61	30	9
Per lexical status (words vs. nonwords): Participants with an error rate above x% removed, where x is determined based on a one-sided proportion test to see whether participants performed above chance (alpha level = .05, chance level means p = .50)	55	25	20
Keep participants regardless	21	75	4
Exclusion Items Accuracy			
Across trials: Items with an error rate above 25% removed	52	43	5
Across trials: Items with an error rate above 50% removed	86	11	4
Keep items regardless	38	59	4
Exclusion Trials Accuracy			
Exclude trials with an incorrect response	75	21	4
Exclude trials with an incorrect response and trials following an incorrect response	32	57	11
Keep trials regardless	29	64	7
Exclusion First Trial			
Exclude the first trial of each block	61	30	9

Keep trials regardless	73	23	4
Exclusion Negative RTs			
Exclude negative RTs	82	11	7
Keep trials regardless	20	71	9
Outliers Participants short RTs			
Across trials: Remove participants who responded quicker than 250 ms on more than 25% of the trials	57	32	11
Keep participants regardless	57	36	7
Outliers Participants timeouts			
Across trials: Remove participants with more than 50% time out trials (i.e., responses outside of the 3s window)	84	12	4
Across trials: Calculate each participant's proportion of time outs and remove those whose proportion is 3 SD below the mean	68	29	4
Keep participants regardless	34	62	4
Outliers Participants long RTs			
Across trials: Calculate each participant's mean RT and remove those whose mean RT is 2 SD above the grand mean	39	52	9
Across trials: Calculate each participant's mean RT and remove those whose mean RT is 2.5 SD above the grand mean	62	29	9

Across trials: Calculate each participant's mean RT and remove those whose mean RT is more than three scaled MAD above and below the median of participant's mean RTs, with scaled MAD defined as c*median(abs(mean RTs-median(mean RTs))), where c=-1/(sqrt(2)*erfcinv(3/2))	52	18	30
Keep participants regardless	54	43	4
Outliers Items long RTs			
Across trials: Calculate each item's mean RT and remove those with a mean RT of 2.5 SD above the grand mean	62	34	4
Keep items regardless	61	38	2
Outliers Trials short RTs absolute			
RTs $<$ 50 ms removed	84	11	5
RTs $<$ 100 ms removed	86	11	4
RTs $<$ 150 ms removed	70	20	11
RTs $<$ 160 ms removed	61	29	11
RTs $<$ 200 ms removed	52	36	12
RTs $<$ 250 ms removed	32	54	14
RTs $<$ 300 ms removed	18	71	11
Keep trials regardless	23	71	5
Outliers Trials long RTs absolute			

RTs > 1000 ms removed	12	75	12
RTs > 1500 ms removed	25	64	11
RTs > 1600 ms removed	25	66	9
RTs > 2000 ms removed	43	50	7
RTs > 2500 ms removed	54	39	7
RTs > 3000 ms removed	79	18	4
Keep trials regardless	30	64	5
Outliers Trials relative			
Across trials: 5% fastest and 5% slowest RTs are removed	27	64	9
Across trials: RTs +- 2 SD from the mean are removed	23	64	12
Across trials: RTs +- 2.5 SD from the mean are removed	48	43	9
Across trials: RTs +- 3 SD from the mean are removed	61	30	9
Across trials: RTs +- 3 SD from the mean are replaced by the mean +- 3 SD (i.e., the end of the distribution)	20	66	14
Across trials: RTs beyond third quartile + 3*interquartile range are removed	30	39	30
Per item: RTs +- 2.5 SD from the item-specific mean are removed	48	38	14
Per item: RTs +- 3 SD from the item-specific mean are removed	52	36	12

Per item: RTs +- 2 SD from the item-specific mean are replaced by the mean +- 2 SD (i.e., the end of the distribution)	14	68	18
Per participant: Remove 5% fastest and 5% slowest RTs	27	59	14
Per participant: RTs +- 2 SD from the participant-specific mean are removed	23	68	9
Per participant: RTs +- 2.5 SD from the participant-specific mean are removed	46	45	9
Per participant: RTs +- 3 SD from the participant-specific mean are removed	55	32	12
Per participant: RTs +- 5 SD from the participant-specific mean are removed	48	39	12
Per participant: RTs +- 2 SD from the participant-specific mean are replaced by the mean +- 2 SD (i.e., the end of the distribution)	7	77	16
Per participant: RTs +- 2.5 SD from the participant-specific mean are replaced by the mean +- 2.5 SD (i.e., the end of the distribution)	12	70	18
Per condition (related vs. unrelated): RTs +- 2.5 SD from the condition-specific mean are removed	25	62	12
Per condition (related vs. unrelated): RTs beyond third quartile + 1.5*interquartile range are removed	20	66	14

Per condition (related vs. unrelated): RTs +- 2 SD from the condition-specific mean are replaced by the mean +- 2 SD (i.e., the end of the distribution)	9	80	11
Per condition (related vs. unrelated) and item combination: RTs +- 2.5 SD from the condition-by-item specific mean are removed	27	57	16
Per condition (related vs. unrelated) and participant combination: RTs +- 2 SD from the condition-by-participant-specific mean are removed	18	70	12
Per condition (related vs. unrelated) and participant combination: RTs +- 2.5 SD from the condition-by-participant-specific mean are removed	25	59	16
Per condition (related vs. unrelated) and participant combination: RTs +- 3 SD from the condition-by-participant-specific mean are removed	32	54	14
Per block (of 100 trials) and participant combination: RTs +- 3 SD from the block-by-participant-specific mean are removed	30	50	20
Keep trials regardless	32	61	7