Step 0: Clearly outline substantive research question(s). Is there a semantic priming effect? Does the effect differ across languages?

Step 1: Expert selection

* Purposive sampling: research team identifies potential contributors based on their knowledge of the field
* Research team launches a more broad call (e.g., many analyst projects). In the context of the PSA, they can also contact everyone involved in the “main” study
* Meta-analysis-type literature search with certain keywords (e.g., semantic priming, continuous lexical decision), scan articles for appropriateness, and contact corresponding authors.
* Snowball procedure: as part of the Step 2 questionnaire, contributors can nominate other researchers (similar to suggesting reviewers in a manuscript submission procedure)

Step 1.5: Identify Analysis Degrees of Freedom

* Using the pathways from papers selected in Step 1, identify common pathways in the following areas of analysis: data cleaning/data processing, assumptions, analysis.
* Use these steps of analysis to design pathway solicitation as described below, potentially with open ended questions first (no bias) and then directed questions within these analysis checkpoints.

Step 2: Pathway elicitation/selection

* Ask potential contributors to describe as concretely as possible (potentially with pseudo-code) the pathways they consider suitable. One might argue/object that access to the data is necessary. One can accommodate this to some degree by using existing datasets (semantic priming project; available CLD priming data). Critically, the idea is that the pathways can’t be influenced by the eventual outcome (e.g., “is there a significant priming effect”), because confirmation bias and hindsight bias might influence pathway selection. This questionnaire could also include questions regarding expertise etc. but maybe we want to cast a wide net initially?
* Select the pathways from the papers selected in Step 1 (the meta-analysis-type literature search). Only select empirical papers with clear analytic pathways (who should code this? two coders to reduce possibility of bias). With regard to the latter, we don’t have access to the data at this point, so we can’t be biased in that respect. Generally speaking though, many multiverse analyses are conducted on secondary data, so it might be a worry. Could be solved by using coders without access to the data.

Step 3: Combining pathways

* Research team processes the data from step 2 and identifies which pathways can be combined. This results in the “full” multiverse.

Step 4 Pathway validation

* All the choices from the full multiverse analysis will be presented to potential collaborators. They could evaluate them based on criteria by Del Guidice and Gangestad (2021), order them in terms of suitability. Here (and also in Step 2?), it is of crucial importance to identify the goal(s) of the multiverse: assessing robustness of the effect, testing boundary conditions, and transparency. More on that in a separate section below (not included yet). Maybe it might even make sense to include a short vignette for each goal that could be included in the questionnaire. Here, I would include questions about expertise, confidence,… I think this should be doable from the point of view of contributors. However, the big question is how to translate this output in the eventual multiverse. I don’t think it is feasible to do the entire multiverse constructed in Step 2, at least for SPAML. If it would, you can then do that multiverse and weigh the pathways based on the validators’ assessment of suitability. However, in SPAML’s case, I suspect, it will not be feasible to exhaustively consider all possibilities. In that case, one could use the validators’ assessment of suitability and select the top X pathways. Alternatively, or additionally, one could draw a random sample from the pathways (and use validators’ assessment of suitability as a weight, either for the sampling, or later when processing the outcomes of the pathways).

Some questions (some are already in the descriptions):

* Should we limit Step 0 to “is there a semantic priming effect”?
  + I think for this example, yes, mostly because it’s a very broad question already.
* Same researchers for Step 2 and 4? Or different “validators”?
  + I would vote for the same people. Or at least sample from the same pool and you might not get the same exact people twice but at least a good sample for both.
* How to go about this practically? I could maybe see it as a Registered report: First submit proposal for procedure, and then put it in practice? Advantage is that we could get started with a small group (maybe involve a few other experienced multiverse people?), and get a green light before involving collaborators. The “outcome” would be the pathways or also the outcome of the multiverse? In the latter case, the deadline (for Stage 2) depends on that of the main SPAML project. Other disadvantage is that the purpose/take-home message might get blurry: the procedure to arrive at a multiverse systematically + a multiverse that can be applied to different kind of semantic priming studies (and beyond) + whether semantic priming is robust. Maybe the outcome could be just the first two, and the result of the multiverse analysis itself could be included in NHB (as exploratory)?
  + I could see this going either way. If we do the RR, maybe use the PCL <https://rr.peercommunityin.org/> unless you have a specific journal in mind. To me, this seems like a great Psych Methods or BRM paper, but neither do RRs.
  + <https://www.apa.org/pubs/journals/met/call-for-papers-tutorials> Psych Methods is specifically asking for tutorials, so might be a good shot to aim high?