FORM 2C - EVALUATION FORM RESEARCH MASTER'S THESIS RESEARCH PROPOSAL

Title of research project:	Assessing the performance of Occam's window for Bayesian model averaging
Student name:	David Coba
Student ID card number:	12439665
Reviewer name:	
Role:	Thesis committee member / Second assessor
Status (1st version or revision):	First draft

Please indicate your overall evaluation below, based on your evaluation of the specific points on the following pages. If your overall evaluation is 'not assessable' then the proposal is not approved and you can circle '0'.

Overall evaluation:

2 Approved (i.e., the research master's thesis project can continue; start data collection/analysis).
The student responds in a response letter point-by-point to the comments and submits this at CANVAS.

1 Conditionally approved (i.e., the proposal has some deficits that need to be revised before the research master's thesis project can continue).
The student revises the research proposal and writes a response letter with a point-by-point response how the comments are handles, and submits the revised proposal with the response letter at CANVAS.

0 Not approved (i.e., the research master's thesis proposal has severe deficits that need to be revised before the research master's thesis project can continue).
The student revises the research proposal and writes a response letter with a point-by-point response how the comments are handles, and submits the revised proposal with the response letter at CANVAS.

When completed with the evaluation, please upload your evaluation in CANVAS, using the Speedgrader in Modules - 'Submit research proposal' and attach the completed evaluation form. The assigned thesis committee member will compose a decision letter based on her/his evaluation and the second assessor's evaluation.

Please state your judgment on each specific item on this form using the denotation stated below (2, 1, 0, X), together with a (short) explanation. Obviously students are best served with concrete, concise and constructive feedback.

Your evaluation (per item, as well as the overall evaluation) can be either of the following:
2 – Approved
1 – Conditionally approved
0 – Not approved
X – Not assessable (Not clear enough / not enough information to be able to make a proper judgment)

A RESEARCH PROPOSAL

1. PROJECT DESCRIPTION

3.1 Prior research

- *Is the problem/research question described in a clear and concise manner?*
- *Is previous literature leading up to the research question adequately described?*
- *Is it clear what the contribution of the proposed research is in light of previous literature?*
- *Is the discussion of the previous literature well-structured and ordered?*
- *Is the theoretical background clear?*
- *Does the proposal in general have a clear structure and flow?*

Evaluation (2, 1, 0, X) :
Explanation and any suggestions for improvements:

I think the explanation of the method could be more clear. I realize that this is a difficult topic and a new one to the student, but here are a few suggestions that might help:

- In the summary you say you study Occam's window in general and specifically in the context of graphical models which sounds like a contraction
- If possible, already give some intuition for how Occam's window works

- In 3.1 you say that estimating a single model "essentially ignores the uncertainty …" why essentially? I think this statement could be cleared up
- In 3.1 "can help with the issue of single-model inference" be more specific; you already explained what the problem is
- Section 3.1 starts with the problem, but I think a paragraph before on the background is needed (what are graphical models, what are the general issues one has to tackle to estimate them)
- 3.1: why is it unclear how to implement the first group of methods?
- 3.1: "weighted average" what does this refer to? That the models not considered are weighted to zero?
- 3.1: I was a bit surprised by the paragraph "one method to combine …"; is this still about the second type of method? Or is this a general method? If the latter is the case, maybe this material should go first

- 3.1: "Occam's window algorithm first selects a set of models that fit the data reasonably well, and then discards all models that have simpler counterparts that fit the data equally well"; shouldn't we be selecting those models that fit equally well but are simpler based on the principle of parsimony?
- 3.1: I don't understand how the set A' is constructed; it's definition seems to depend on M_l, which is itself a member of A'
- 3.1: Can you provide a bit more intuition for how big A=A' \ B is? I would imagine one model fits best, which means B contains all models, which means A contains only one model. Clearly, I'm misunderstanding the method, but from the text it is not entirely clear to me
- 3.1: When first mentioning the "passes" it's not clear what they refer to, and why we need another pass backwards
- 3.1: General comment: A figure / example with 3-4 nodes would really help; I think this would also be nice for the final thesis

3.2 **Research question and hypotheses**

- *Does the research question follow logically from the theoretical background?*
- *Does the research question translate into specific, clear and testable research model/ expectations/ hypotheses?*
- *Are the research model/ expectations/ hypotheses well-founded and logically related to underlying psychological mechanisms?*

Evaluation (2, 1, 0, X) :
Explanation and any suggestions for improvements:

These questions are very vague without the specifications in Section 4; maybe you can address this by giving a bit more information and referring to Section 4; for example, it should be clear in the research question whether the focus is regression or graphical models

2. **PROCEDURE**

- *Are the research model/ expectations/ hypotheses operationalized in a clear and correct manner in an experimental design?*
- *Are the proposed measures and DVs properly described and justified (incl. validity and reliability; are proposed interventions adequate)?*
- *Is the suggested sample (category, control groups, sample size) properly justified?*
- *In case new or adjusted instruments or measures (e.g., questionnaires) are used, is the new instrument first validated before use in the main experiment? If not, is there a proper justification and indication on how the data from this measure can be interpreted?*
- *Is the choice of statistical techniques (per hypothesis) properly justified and explained?*
- *Is there room for modification of the intended procedure, or is there ample motivation why not?*

Evaluation (2, 1, 0, X) :
Explanation and any suggestions for improvements:

- Which different marginal likelihood approximations you will consider?

- How did you select the alternative model selection procedures? Are they the best-performing ones? Or the ones available?
- How will you assess performance?
- What are the simulation conditions you will consider? I realize that this depends a bit on how far you will get in the project, but I think you can give a rough description

B ADDITIONAL INFORMATION RESEARCH PROPOSAL

3. **INTENDED RESULTS**
   - *Are the intended results clearly formulated?*
   - *Are the intended results related to the specific and more general research questions?*
   - *Is there a clear reflection on alternative interpretations in case results turn out differently?*
   - *Is it clear what results will lead to falsification of theory and research model/ expectations/ hypotheses?*
   - *What is the innovative value of the research in scientific and / or practical and societal terms?*

Evaluation (2, 1, 0, X) :
Explanation and any suggestions for improvements:

- Are there any theoretical expectations with respect to which method will do better than another? Maybe condition on the characteristics of the data generating model?
- Do you expect trade-offs in estimation performance vs. computation time in the considered methods? If yes, which?

4. **WORKPLAN AND FEASIBILITY**
   - *Is the proposed planning realistic and expedient?*
   - *Is the proposed research feasible and sensible in terms of size and investment in terms of time?*
   - *Is the proposed research realistic in terms of recruitment of participants?*
   - *Is the proposed research realistic in terms of infrastructural needs (e.g., place of execution, used materials)?*

Evaluation (2, 1, 0, X) :
Explanation and any suggestions for improvements:

This looks good. I agree that this is very flexible and that one could start writing up the thesis at any time.

C LEARNING AIM / EFFECT AND OVERALL JUDGMENT

5. **LEARNING AIM / EFFECT**
   - *At the end of this research project, is it your judgment that the student will be able to independently think of a research question, and set up / run an empirical study to test research model/ expectations/ hypotheses derived from the research question?*

Evaluation (2, 1, 0, X) :
Explanation and any suggestions for improvements:

I think the student will learn a lot about Bayesian inference, and will learn to execute and report a simulation study. I consider both crucial skills in the area psychological methods.

6. **FINAL JUDGMENT**
   - *Are the previous sections generally positively evaluated?*
   - *Are the previous sections in balance with each other?*
   - *What is the quality of the proposal as a whole?*
   - *What is your judgment on the structure, readability, and length of the proposal?*

Evaluation (2, 1, 0, X) :
Explanation and any suggestions for improvements:

The proposal is overall clear, interesting and feasible. However, the method needs to be explained better, some choices in terms of compared methods need to be motivated more and some specifics regarding the implementation should be provided.