Empirie und Experimentieren in der Softwaretechnik

M. Danz, T. Gräf, C. Michel*

Advisor: Andrei Miclaus†

Karlsruhe Institute of Technology (KIT)
Pervasive Computing Systems – TECO

*danz@teco.edu

tobias.graef@student.kit.edu

michel@teco.edu

† miclaus@teco.edu

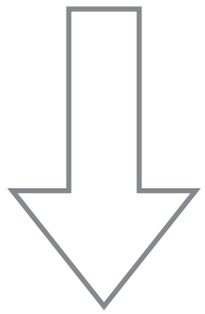
Karlsruher Institut für Technologie

Original Question

How can students compare two software architectures?

Students' Issues with Scientific Working[RHB06]

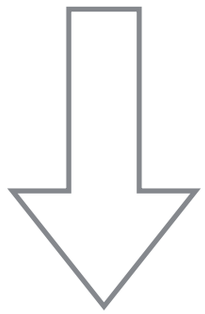
- "Students had problems constructing well-formulated […] questions."
- "Students used limited criteria for identifying the best or better evidence […]"
- "Students used a very limited number of search terms."
- …



Many students lack knowledge about scientific working and experiment design.

Original Question Revised

How can students compare two software architectures?

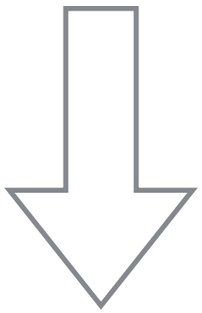


How to support students in scientific working and experiment design?

How to Support Students in Scientific Working and Experiment Design?

- Evidence-Based Software Engineering (EBSE) [KDJ04]
- Scientific Method [SMOP]
- SEED [JR08]
- Evidence Map [EBS]
- Workflow Graph [RB08]

Tools not tailored for students



We created a **process** and **guiding documents** for students.

Process as Guidelines for Students

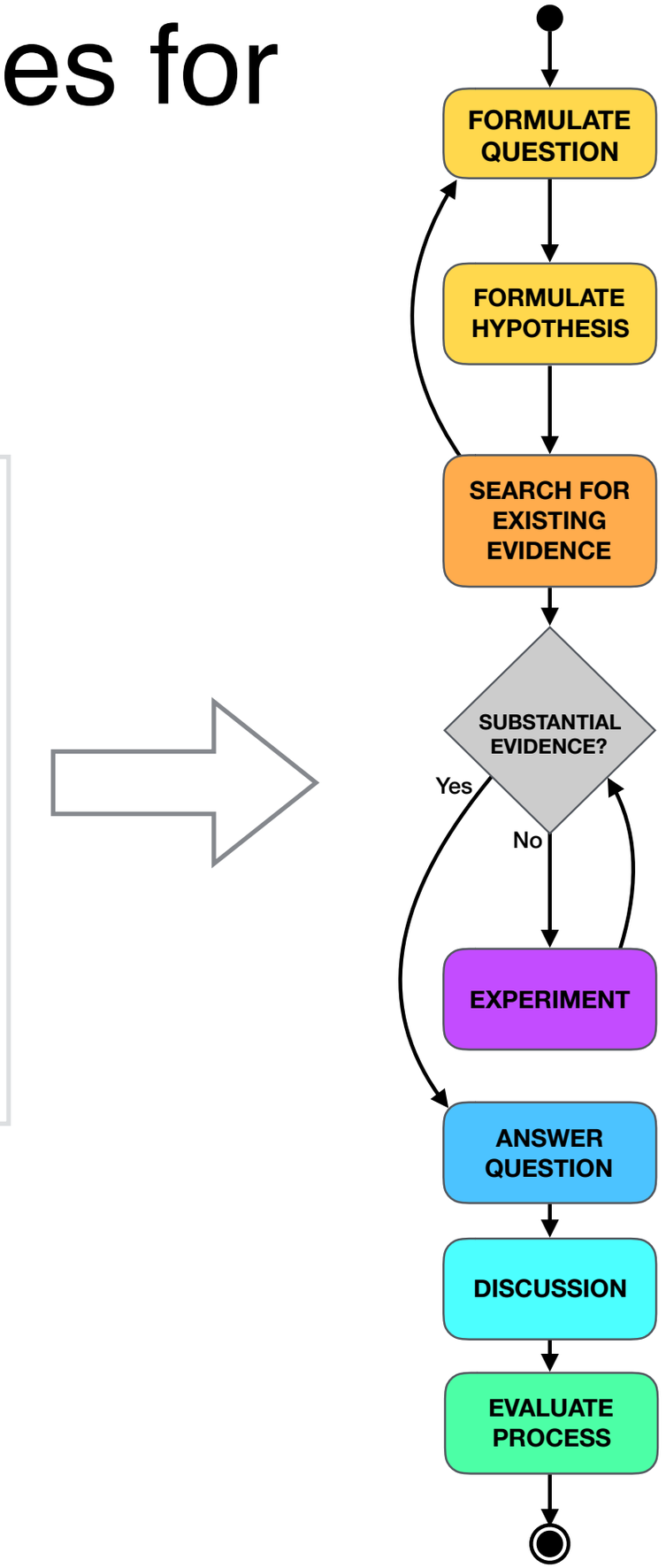
EBSE Process Steps:

1. Ask an answerable question.
2. Find the best evidence that answers that question.
3. Critically appraise this evidence.
4. Apply the evidence (and critical appraisal).
5. Evaluate the performance in previous steps.

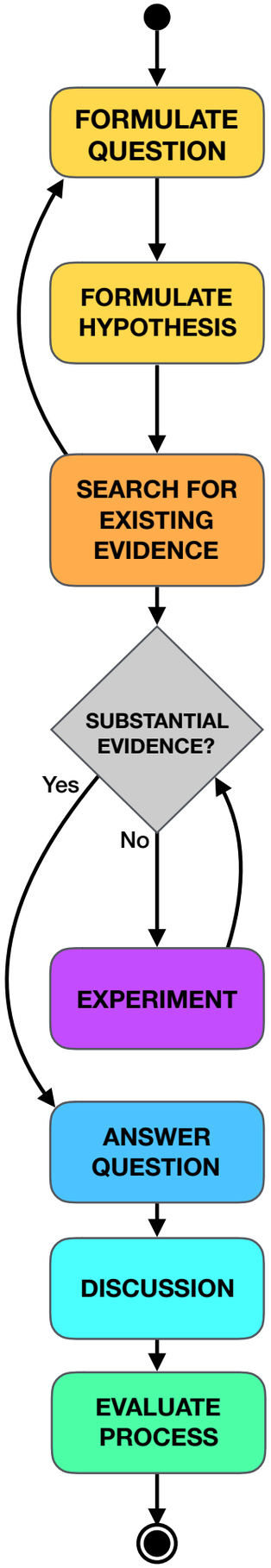
Process as Guidelines for Students

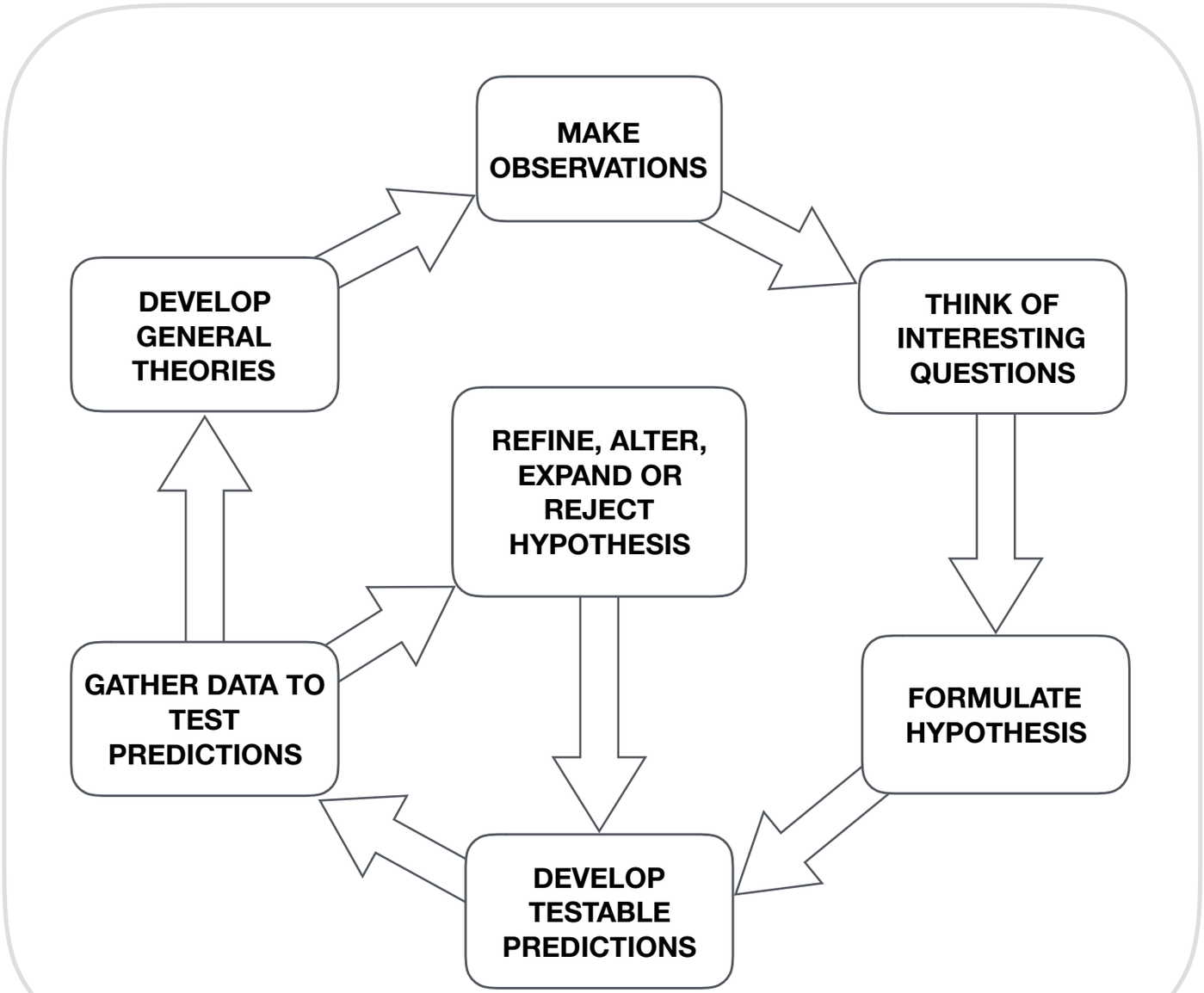
EBSE Process Steps:

1. Ask an answerable question.
2. Find the best evidence that answers that question.
3. Critically appraise this evidence.
4. Apply the evidence (and critical appraisal).
5. Evaluate the performance in previous steps.

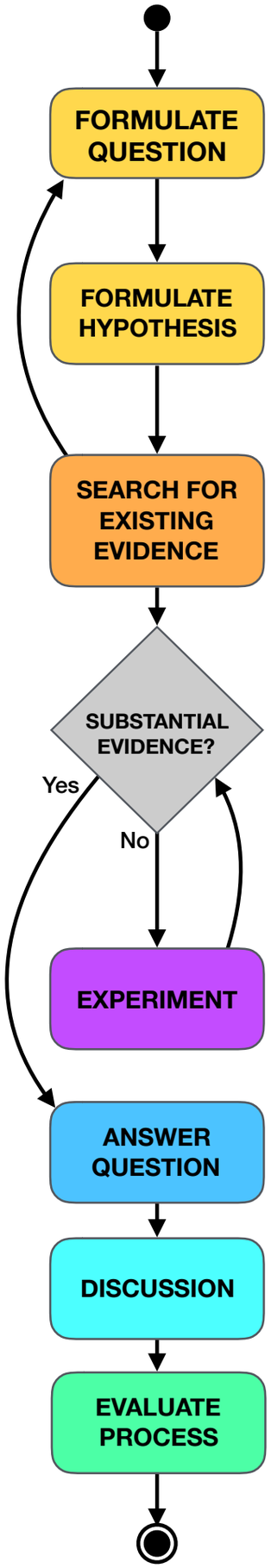


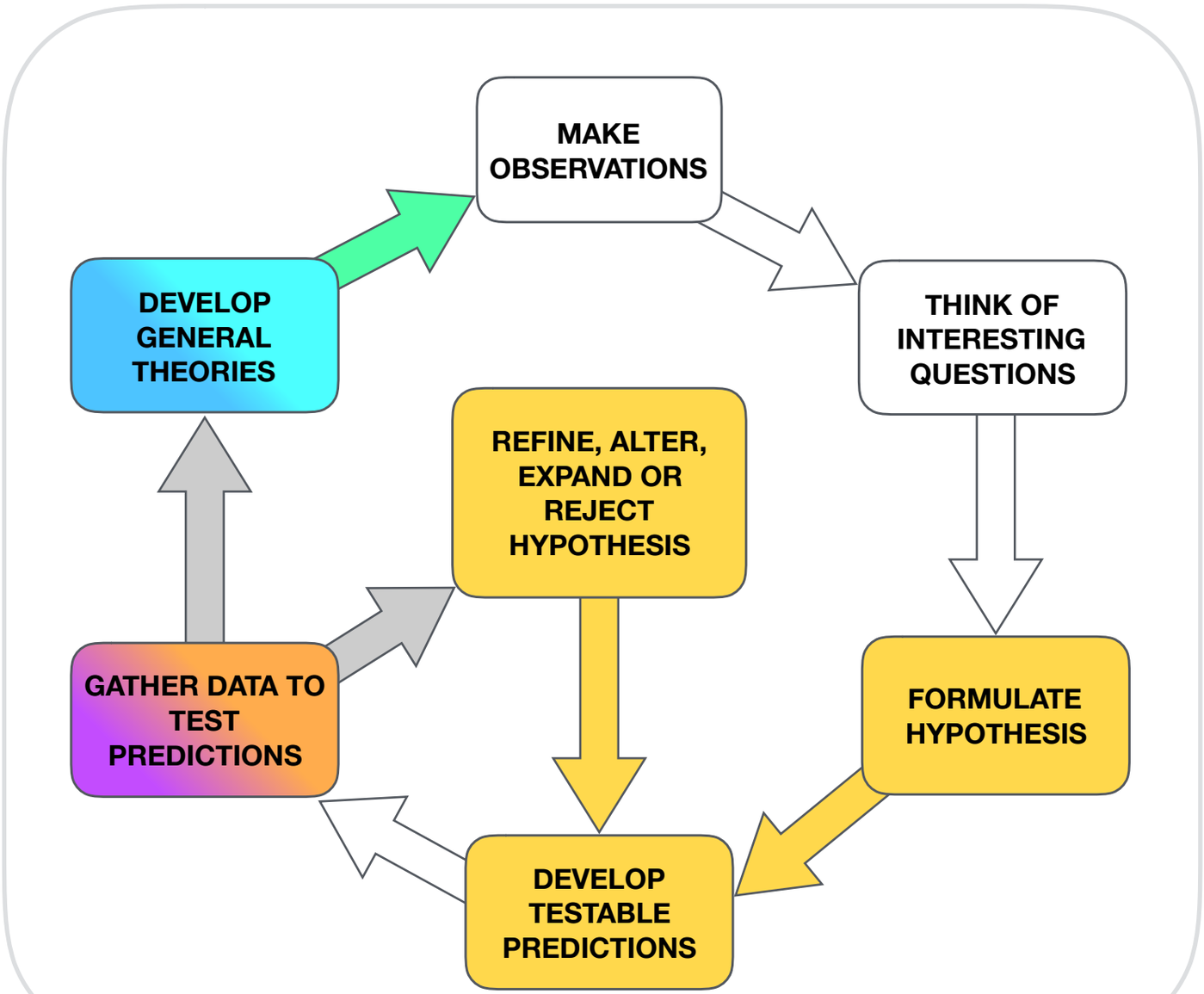
Scientific Method

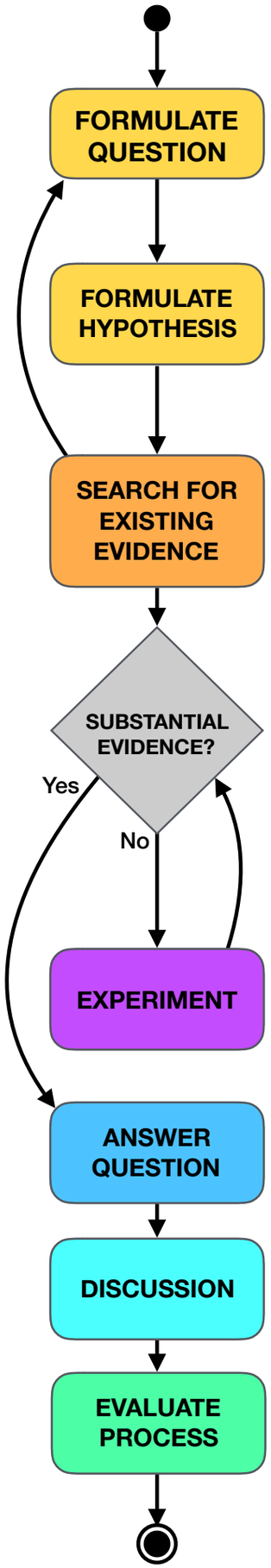




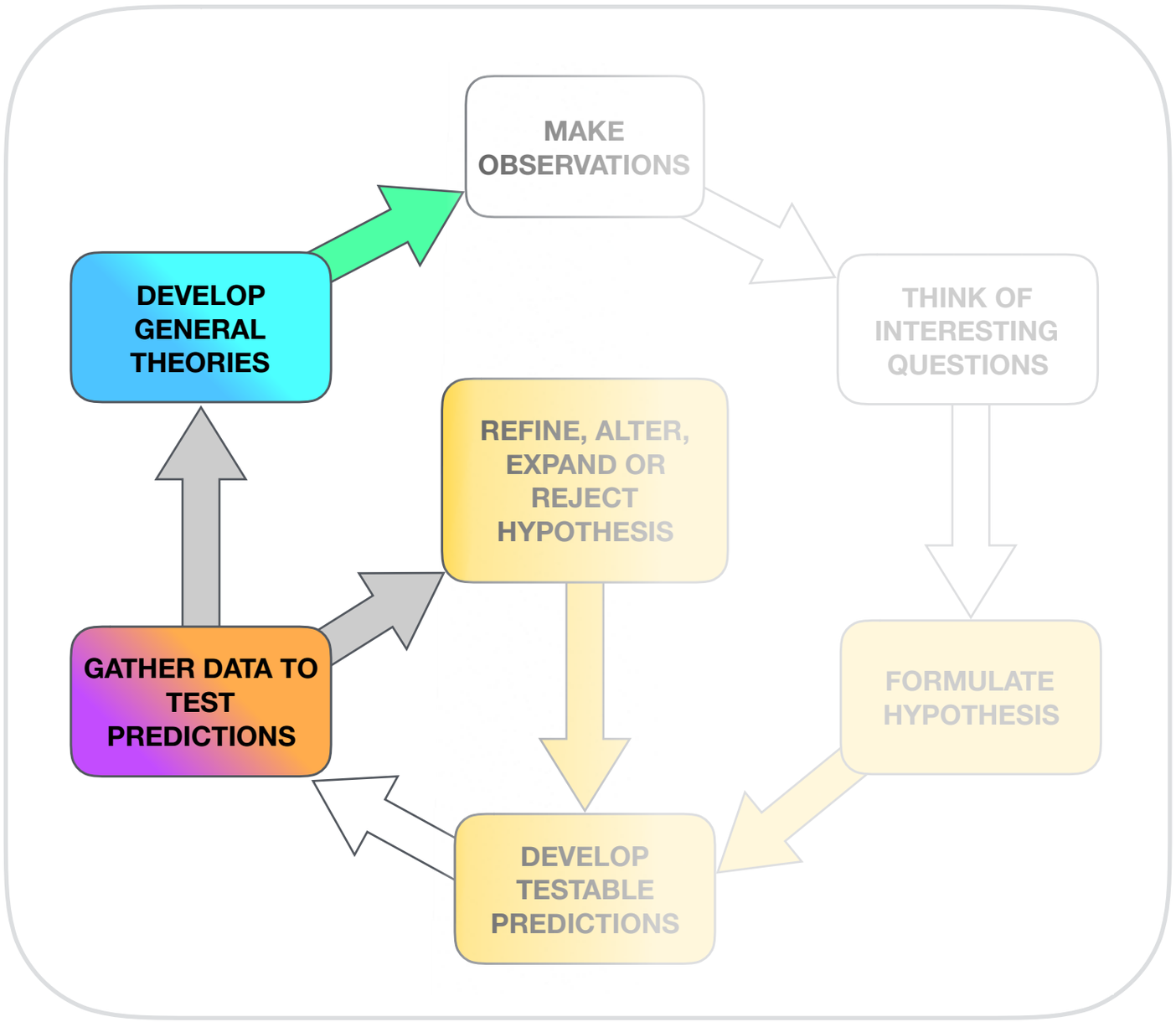
Mapping on Scientific Method





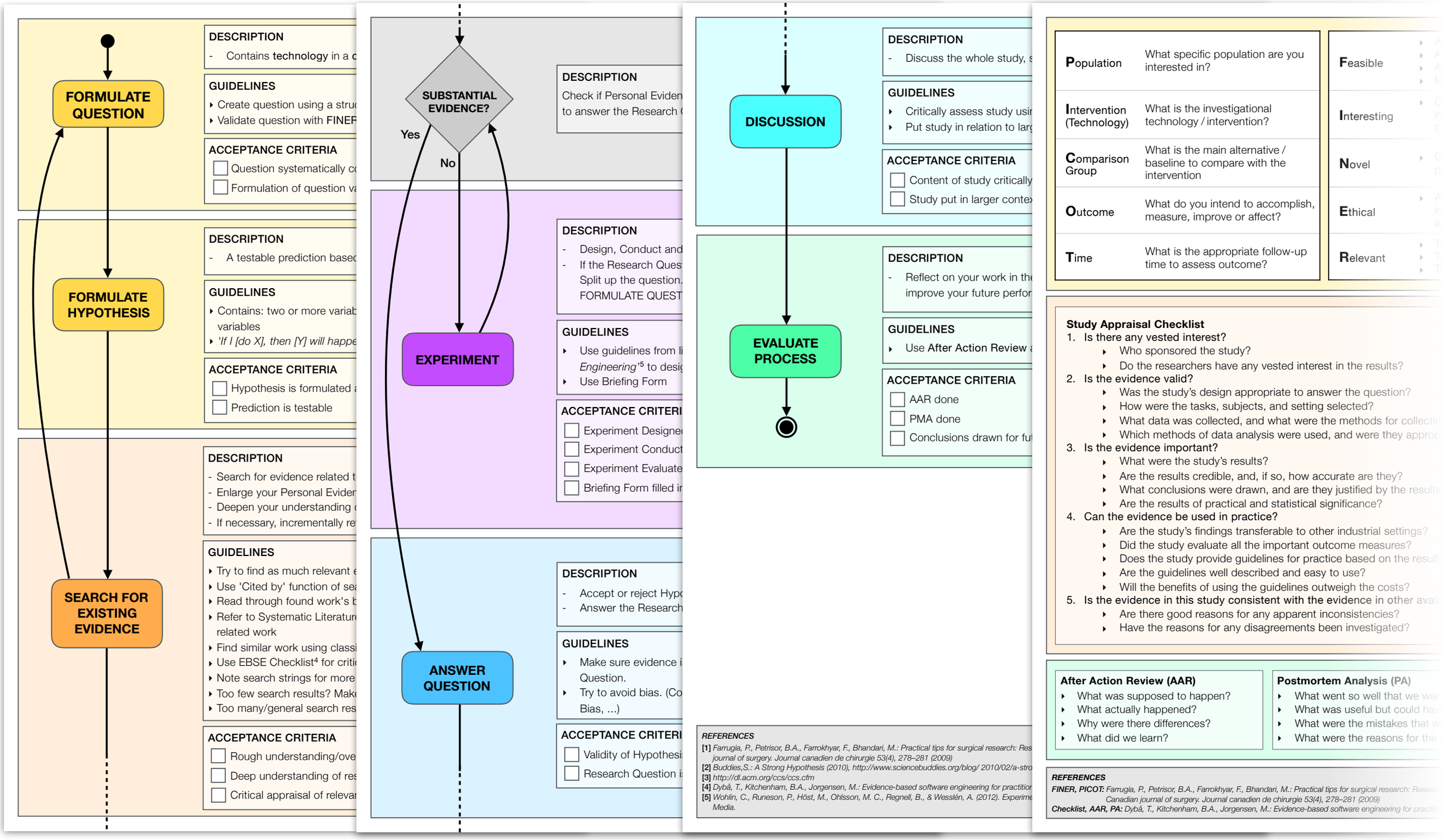


Mapping on Scientific Method

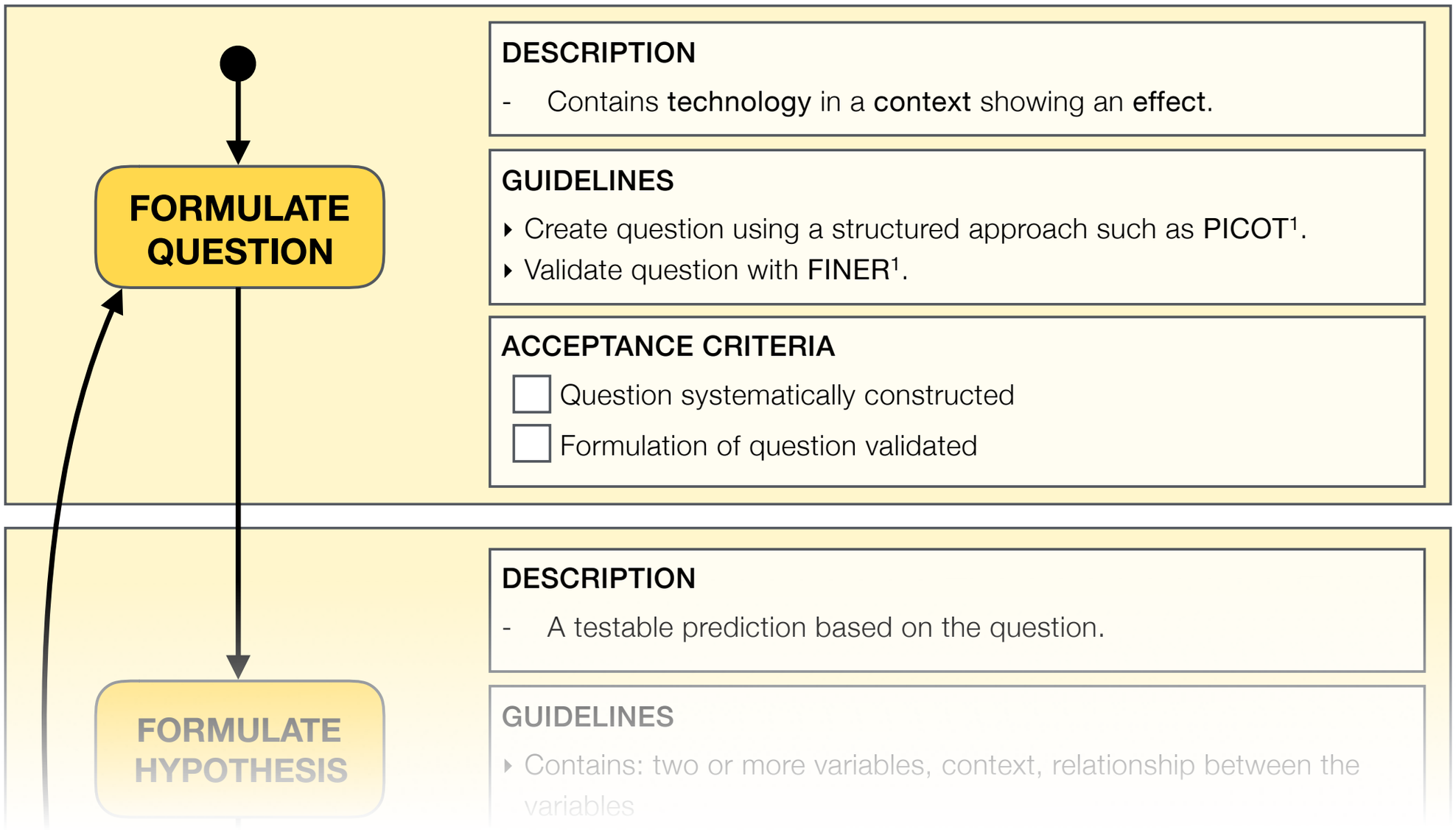


Focus on experimenting, therefore unrolled respective nodes.

Process Document - Overview



Process Document - Details



Process Document - Tools

Population	What specific population are you interested in?
Intervention (Technology)	What is the investigational technology/ intervention?
Comparison Group	What is the main alternative/ baseline to compare with the intervention
Outcome	What do you intend to accomplish, measure, improve or affect?
Time	What is the appropriate follow-up time to assess outcome?

Feasible	▸ Adequate number of subjects ▸ Adequate technical expertise ▸ Affordable in time and money ▸ Manageable in scope
Interesting	▸ Getting the answer intrigues investigator, peers and community
Novel	▸ Confirms, refutes or extends previous findings
Ethical	▸ Amendable to a study that institutional review board will approve
Relevant	▸ To scientific knowledge ▸ To clinical and health policy ▸ To future research

After Action Review (AAR)

- What was supposed to happen?
- What actually happened?
- Why were there differences?
- What did we learn?

Postmortem Analysis (PA)

- What went so well that we want to repeat it?
- What was useful but could have gone better?
- What were the mistakes that we want to avoid for the future?
- What were the reasons for the success or mistakes?

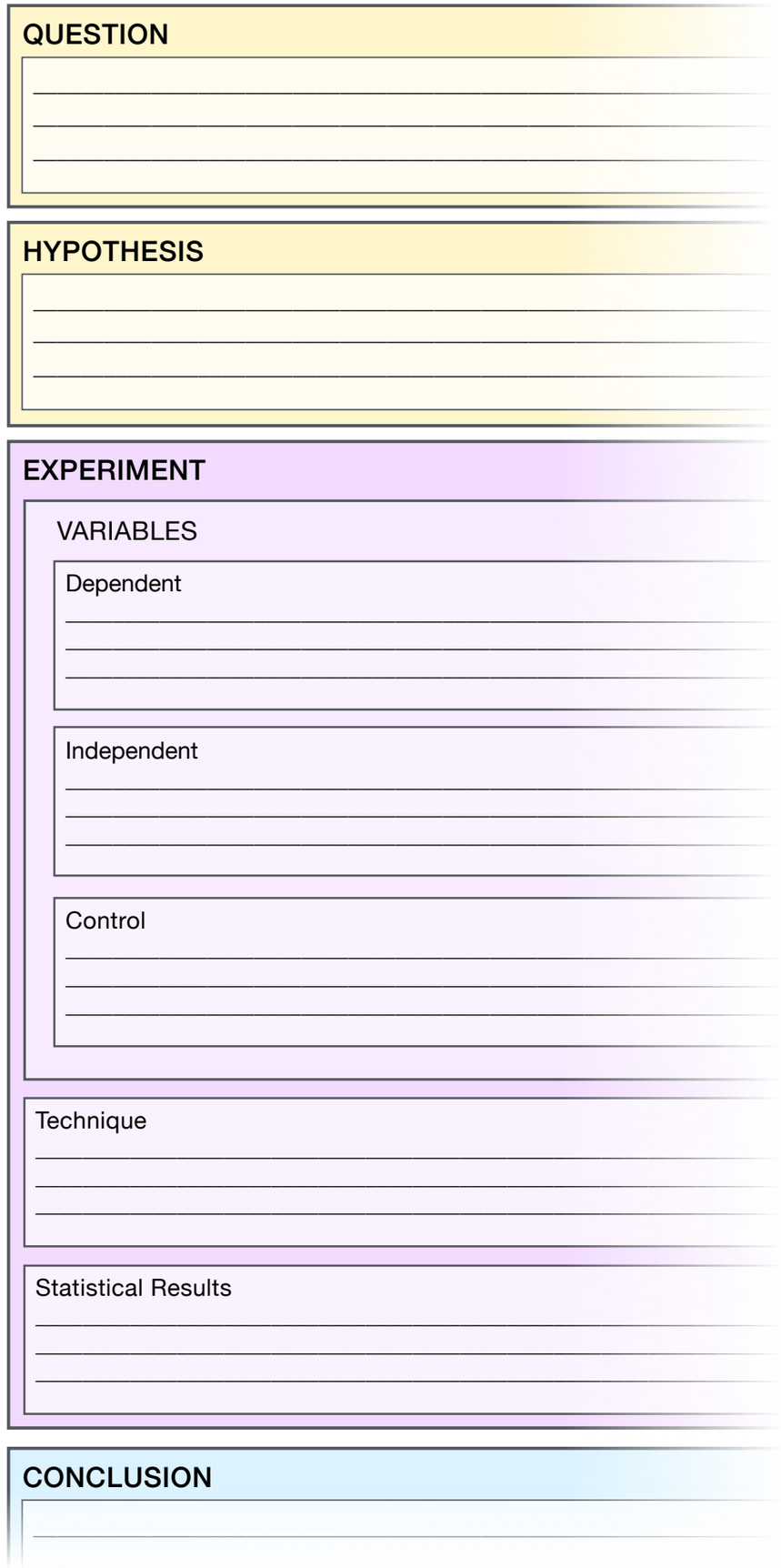
REFERENCES

FINER, PICOT: *Farrugia, P., Petrisor, B.A., Farrokhyar, F., Bhandari, M.: Practical tips for surgical research: Research questions, hypotheses and objectives. Canadian journal of surgery. Journal canadien de chirurgie 53(4), 278–281 (2009)*

Checklist, AAR, PA: *Dybå, T., Kitchenham, B.A., Jorgensen, M.: Evidence-based software engineering for practitioners. IEEE Software 22(1), 58–65 (2005)*

Briefing Form

- Analog concept
- Template for paper summary
- Assistance for experiment design
- Easier to search existing work

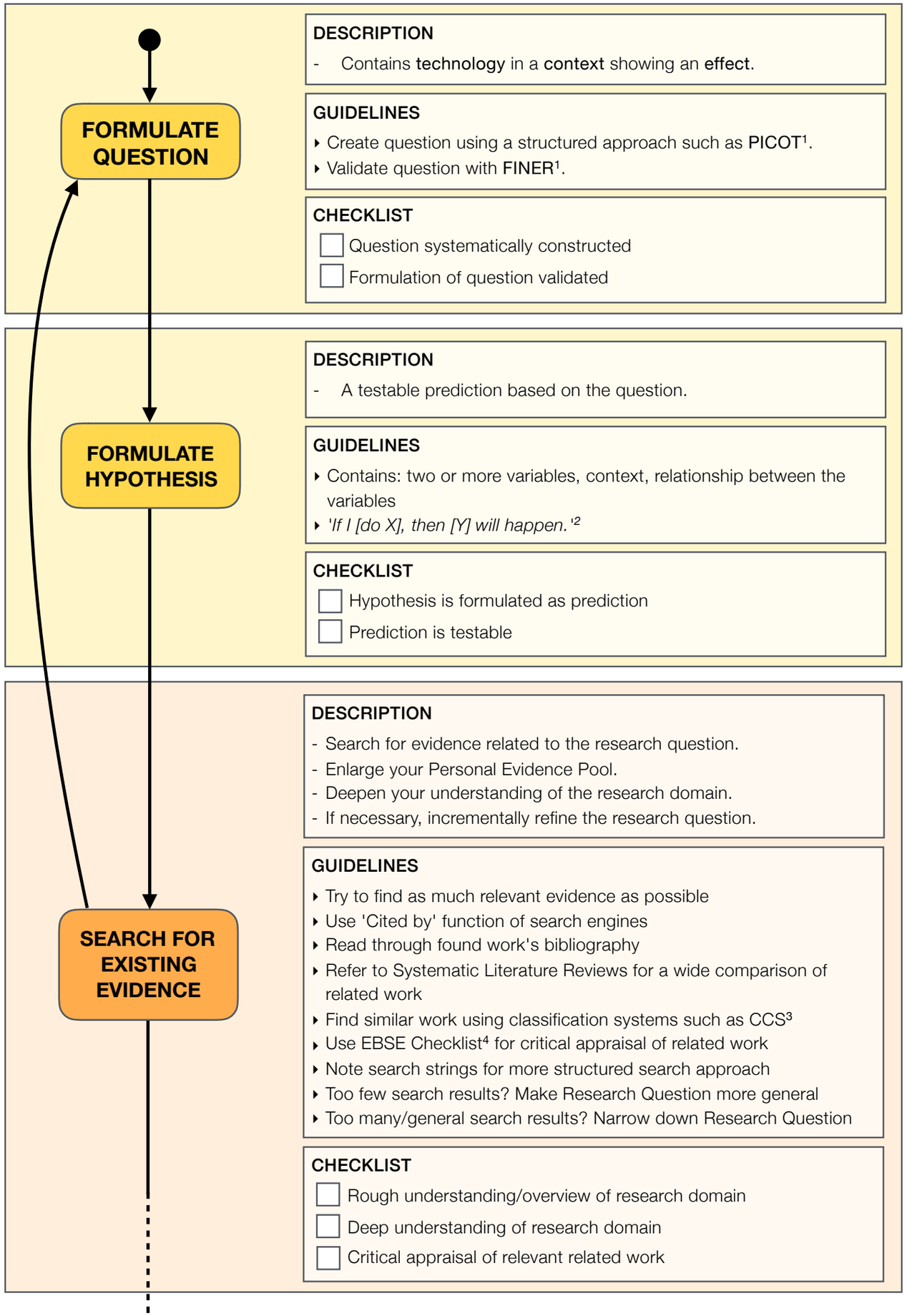


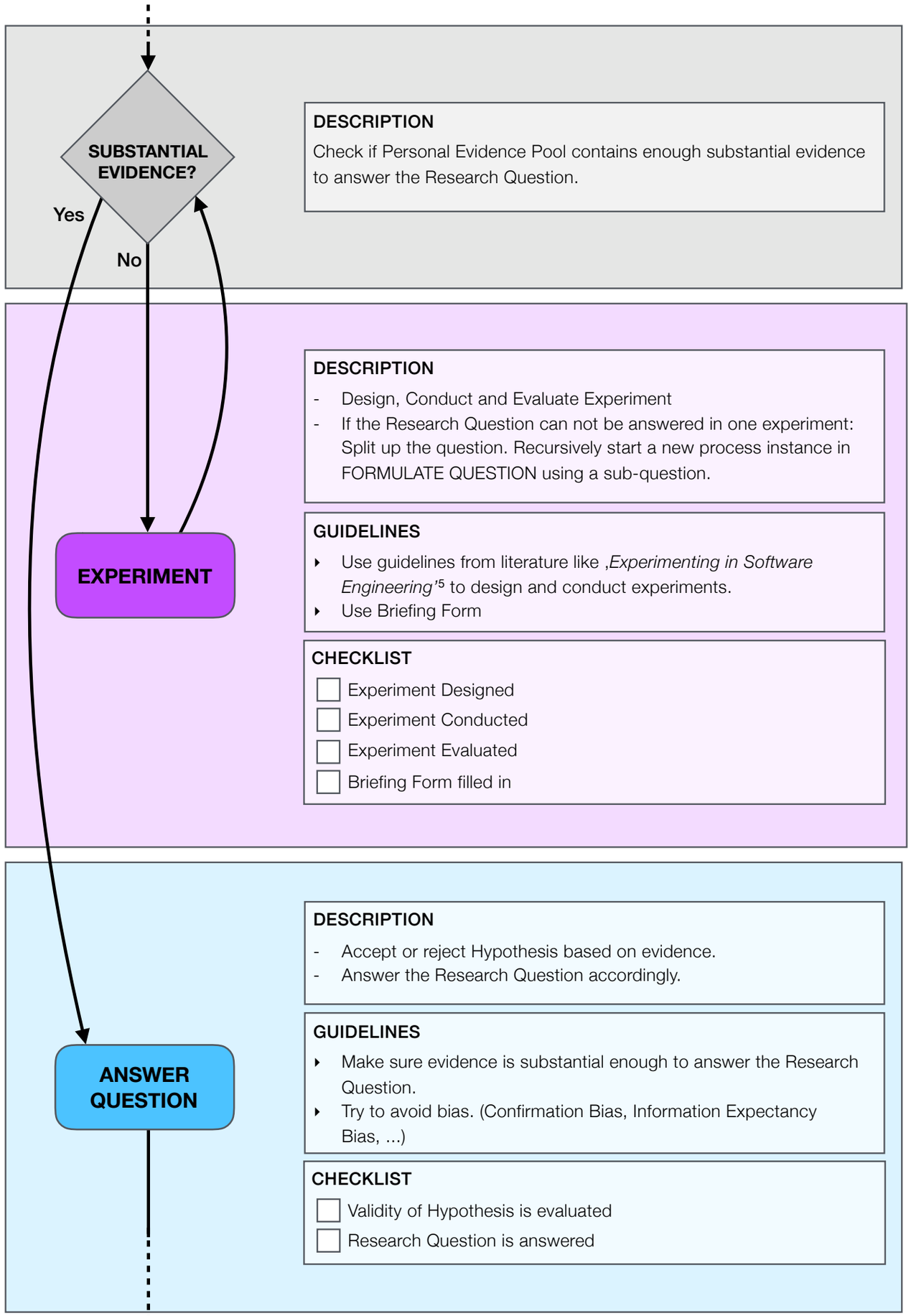
Future Work

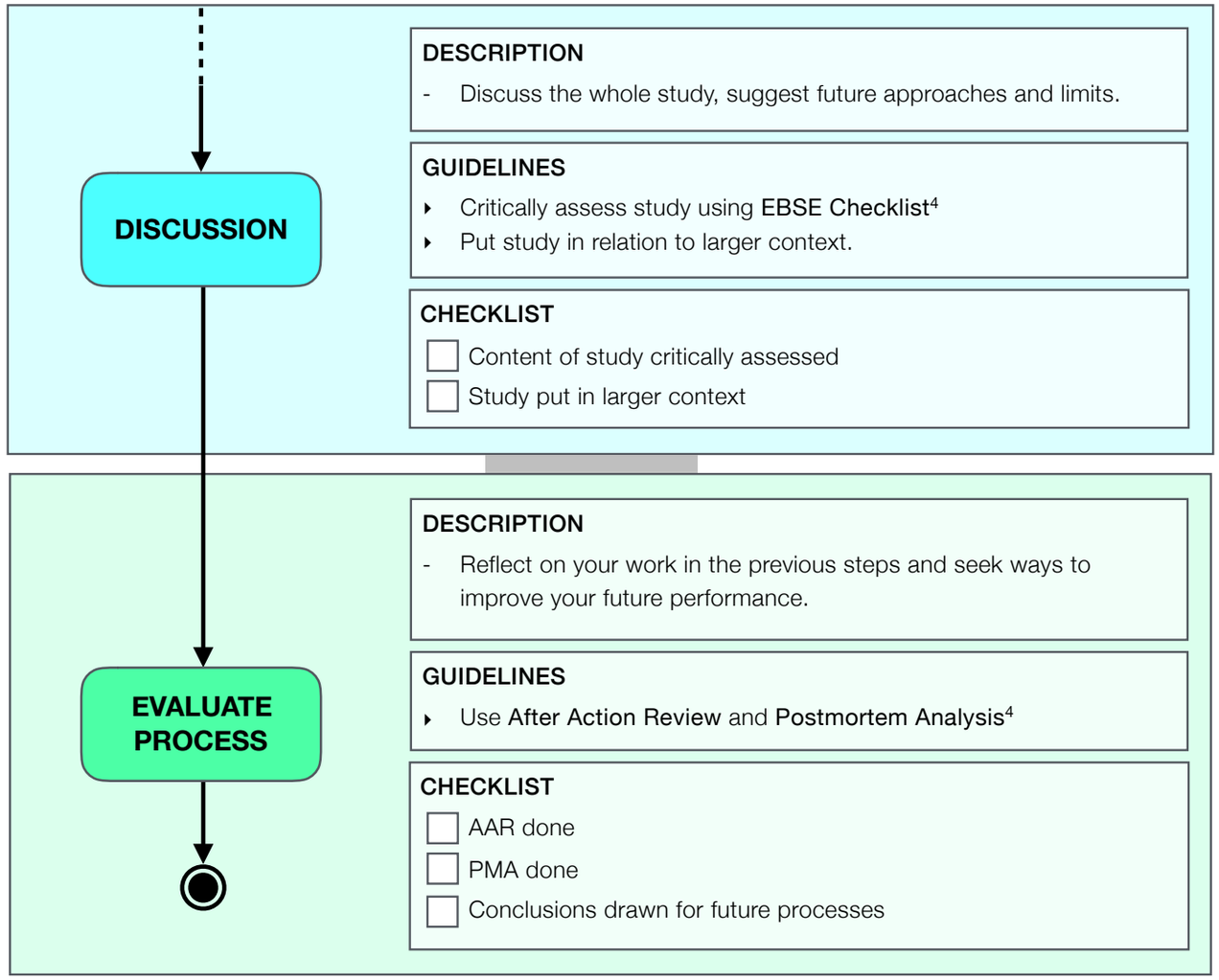
- Digitalize Briefing Form
- Should be evaluated using students' thesis
  - (Two upcoming, maybe four)
    - ➡ Revise Process Document tools
    - ➡ Refine Briefing Form
- Scientific Working Basics missing in Education?

Sources

- [SMOP]: Garland, Jr., Theodore. "The Scientific Method as an Ongoing Process". U C Riverside. (http://idea.ucr.edu/documents/flash/scientific_method/story.htm)
- FINER, PICOT: Patricia Farrugia, Bradley A. Petrisor, Forough Farrokhyar, and Mohit Bhandari. Practical tips for surgical research: Research questions, hypotheses and objectives. Canadian journal of surgery. Journal canadien de chirurgie, 53(4): 278–281, 2009.
- PMA, AAR: Tore Dybå, Barbara A. Kitchenham, and Magne Jorgensen. Evidence-based software engineering for practitioners. IEEE Software, 22(1):58–65, 2005.
- [RHB06]: Austen Rainer, Tracy Hall, and Nathan Baddoo. A preliminary empirical investigation of the use of evidence based software engineering by under- graduate students. 10th International Conference on Evaluation and Assessment in Software Engineering (EASE 2006), 2006.
- [KDJ04]: Barbara A Kitchenham, Tore Dyba, and Magne Jørgensen. Evidence-based software engineering. In Proceedings of the 26th international conference on software engineering, pages 273–281. IEEE Computer Society, 2004.
- [JR08]: David S. Janzen and Jungwoo Ryoo. Seeds of Evidence: Integrating Evidence-Based Software Engineering. In Software Engineering Education Conference, Proceedings, pages 223–232, 2008.
- [EBS]: Evidence Based Software Engineering. https://community.dur.ac.uk/ ebse. accessed 2017-02-25.
- [RB08]: Austen Rainer and Sarah Beecham. Supplementary Guidelines , Assessment Scheme and evidence-based evaluations of the use of Evidence Based Software Engineering, Version 2. (February 2008):1–27, 2008.


FORMULATE QUESTION
DESCRIPTION
- Contains technology in a context showing an effect.
GUIDELINES
Create question using a structured approach such as PICOT[1].
Validate question with FINER[1].
CHECKLIST
Question systematically constructed
Formulation of question validated
FORMULATE HYPOTHESIS
DESCRIPTION
- A testable prediction based on the question.
GUIDELINES
Contains: two or more variables, context, relationship between the variables
'If I [do X], then [Y] will happen.'[2]
CHECKLIST
Hypothesis is formulated as prediction
Prediction is testable
SEARCH FOR EXISTING EVIDENCE
DESCRIPTION
- Search for evidence related to the research question.
- Enlarge your Personal Evidence Pool.
- Deepen your understanding of the research domain.
- If necessary, incrementally refine the research question.
GUIDELINES
Try to find as much relevant evidence as possible
Use 'Cited by' function of search engines
Read through found work's bibliography
Refer to Systematic Literature Reviews for a wide comparison of related work
Find similar work using classification systems such as CCS[3]
Use EBSE Checklist[4] for critical appraisal of related work
Note search strings for more structured search approach
Too few search results? Make Research Question more general
Too many/general search results? Narrow down Research Question
CHECKLIST
Rough understanding/overview of research domain
Deep understanding of research domain
Critical appraisal of relevant related work


SUBSTANTIAL EVIDENCE?
Yes
No
DESCRIPTION
Check if Personal Evidence Pool contains enough substantial evidence to answer the Research Question.
EXPERIMENT
DESCRIPTION
- Design, Conduct and Evaluate Experiment
- If the Research Question can not be answered in one experiment: Split up the question. Recursively start a new process instance in FORMULATE QUESTION using a sub-question.
GUIDELINES
Use guidelines from literature like ‚Experimenting in Software Engineering'[5] to design and conduct experiments.
Use Briefing Form
CHECKLIST
Experiment Designed
Experiment Conducted
Experiment Evaluated
Briefing Form filled in
ANSWER QUESTION
DESCRIPTION
- Accept or reject Hypothesis based on evidence.
- Answer the Research Question accordingly.
GUIDELINES
Make sure evidence is substantial enough to answer the Research Question.
Try to avoid bias. (Confirmation Bias, Information Expectancy Bias, ...)
CHECKLIST
Validity of Hypothesis is evaluated
Research Question is answered



REFERENCES

[1] *Farrugia, P., Petrisor, B.A., Farrokhyar, F., Bhandari, M.: Practical tips for surgical research: Research questions, hypotheses and objectives. Canadian journal of surgery. Journal canadien de chirurgie 53(4), 278–281 (2009)*

[2] *Buddies,S.: A Strong Hypothesis (2010), http://www.sciencebuddies.org/blog/ 2010/02/a-strong-hypothesis.php*

[3] *http://dl.acm.org/ccs/ccs.cfm*

[4] *Dybå, T., Kitchenham, B.A., Jorgensen, M.: Evidence-based software engineering for practitioners. IEEE Software 22(1), 58–65 (2005)*

[5] *Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). Experimentation in software engineering. Springer Science & Business Media.*

Population	What specific population are you interested in?
Intervention (Technology)	What is the investigational technology/ intervention?
Comparison Group	What is the main alternative/ baseline to compare with the intervention
Outcome	What do you intend to accomplish, measure, improve or affect?
Time	What is the appropriate follow-up time to assess outcome?

Feasible	▸ Adequate number of subjects ▸ Adequate technical expertise ▸ Affordable in time and money ▸ Manageable in scope
Interesting	▸ Getting the answer intrigues investigator, peers and community
Novel	▸ Confirms, refutes or extends previous findings
Ethical	▸ Amendable to a study that institutional review board will approve
Relevant	▸ To scientific knowledge ▸ To clinical and health policy ▸ To future research

Study Appraisal Checklist

1. Is there any vested interest?
   - Who sponsored the study?
   - Do the researchers have any vested interest in the results?
2. Is the evidence valid?
   - Was the study's design appropriate to answer the question?
   - How were the tasks, subjects, and setting selected?
   - What data was collected, and what were the methods for collecting the data?
   - Which methods of data analysis were used, and were they appropriate?
3. Is the evidence important?
   - What were the study's results?
   - Are the results credible, and, if so, how accurate are they?
   - What conclusions were drawn, and are they justified by the results?
   - Are the results of practical and statistical significance?
4. Can the evidence be used in practice?
   - Are the study's findings transferable to other industrial settings?
   - Did the study evaluate all the important outcome measures?
   - Does the study provide guidelines for practice based on the results?
   - Are the guidelines well described and easy to use?
   - Will the benefits of using the guidelines outweigh the costs?
5. Is the evidence in this study consistent with the evidence in other available studies?
   - Are there good reasons for any apparent inconsistencies?
   - Have the reasons for any disagreements been investigated?

After Action Review (AAR)

- What was supposed to happen?
- What actually happened?
- Why were there differences?
- What did we learn?

Postmortem Analysis (PA)

- What went so well that we want to repeat it?
- What was useful but could have gone better?
- What were the mistakes that we want to avoid for the future?
- What were the reasons for the success or mistakes?

REFERENCES

FINER, PICOT: *Farrugia, P., Petrisor, B.A., Farrokhyar, F., Bhandari, M.: Practical tips for surgical research: Research questions, hypotheses and objectives. Canadian journal of surgery. Journal canadien de chirurgie 53(4), 278–281 (2009)*

Checklist, AAR, PA: *Dybå, T., Kitchenham, B.A., Jorgensen, M.: Evidence-based software engineering for practitioners. IEEE Software 22(1), 58–65 (2005)*

QUESTION

HYPOTHESIS

EXPERIMENT

VARIABLES

Dependent

Independent

Control

Technique

Statistical Results

CONCLUSION