- Thesis hand in: 11 Aug 2025
- Proposal: 17th feb 2025

- Send Julien literature folder

- Graph - threshold edges and

ToDo
- Try different clustering algorithms
- Try different similarities
- Try with embeddings
- Try thematic analysis
- Adding in higher order themes

Model inputs:

- Raw graph with all similarities
- Graph only the clusters and all the edges within one
- Graph only the clusters and the main edges
-

Clusters

- Category diff algo
- Clusters that I think are real clusters
- Check with ChatGPT to get ideas
- Thematic analysis, higher order words => how do they

Need for different runs of the model

- Run several times with different initialiazations which reflect some priors

!!! Mean != 50 because this means indifference (we don't care that the judgment of people is biased in one direction or the other) !!!

Adding experience vs not experienced

Model without any information about experience

! can remove liking from edge weight !

Algo test:

For each activity A:

- Train the model without activity A in the training graph
- Change liking of activity A to 0 or gibberish
- Put activity A into the graph (the normal one)
- Predict activity A liking
- If strongly wrong, then it relied on the wrong input
- If strongly right, then it didn't relied on the wrong input

Take each activity out of the model one by one and estimate the model without that activity. Then use the model to predict the liking of the activity using ONLY similarity (e.g. give it a liking of 0). Make sure you give in ZERO information about the liking of the held-out activity!