A Tireless Endeavour Towards Excellent Evaluation Practices

Ahmad Zein Assi

Rektorska konferenca Republike Slovenije

22 October 2021

Swiss National Science Foundation

Plan

1. The SNSF Excellence Model
2. How to Evaluate Excellence Science?
3. The SNSF Evaluation Policy
4. The Way Forward and Beyond

The Swiss National Science Foundation (SNSF)

In its mission statement, the SNSF describes how it aims to act in the best interests of researchers, political authorities, the general public and its own employees. It defines the ambitions and values informing its work.

- Quality. We promote scientific excellence. We also expect our own performance to be of a high standard.
- Independence. We operate autonomously and evaluate impartially.
- Responsibility. We are a reliable partner. We fulfil our tasks judiciously and al-ways act ethically.
- Fairness. Our decisions are well-founded and we respect equal opportunities.

The Swiss National Science Foundation (SNSF)

In its mission statement, the SNSF describes how it aims to act in the best interests of researchers, political authorities, the general public and its own employees. It defines the ambitions and values informing its work.

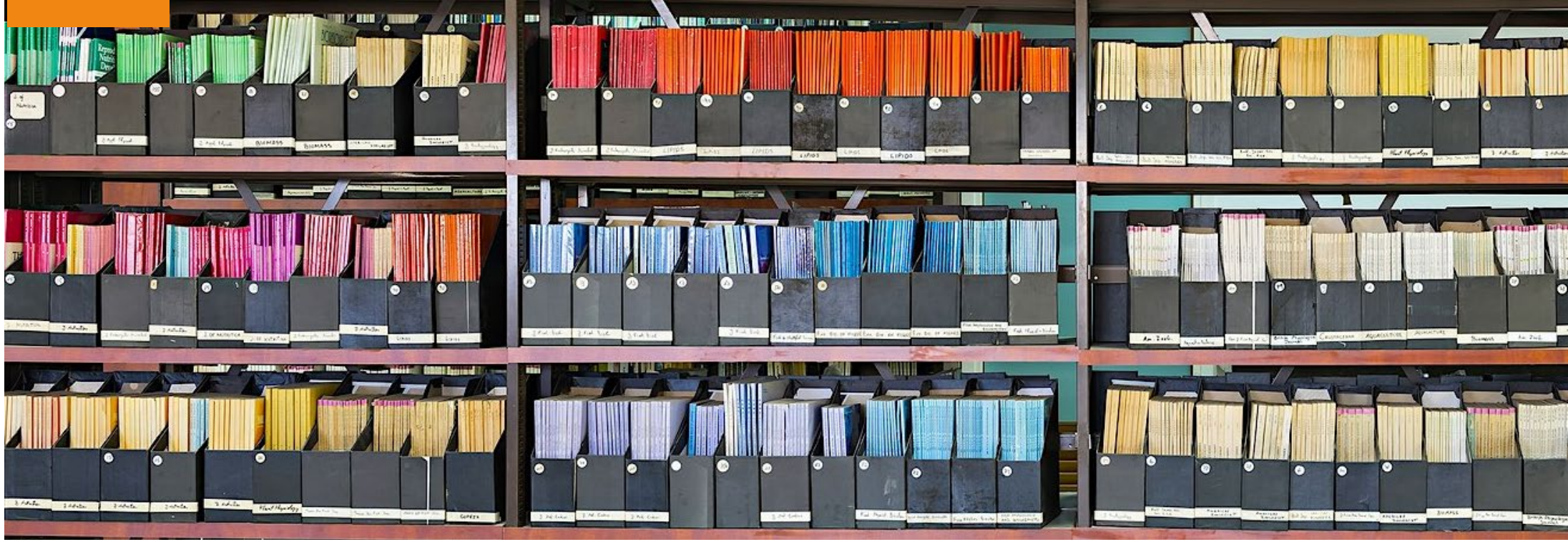
- Quality. We promote scientific excellence. We also expect our own
- Independence. We operate autonomously and evaluate impartially.
- Responsibility. We are a reliable partner. We fulfil our tasks judiciously and al-ways act ethically.
- Fairness. Our decisions are well-founded and we respect equal opportunities.

"We promote scientific excellence"

The Excellence Model of the SNSF

Swiss National Science Foundation

What is Excellent Research?



The Evaluation of Excellent Research

Stages in the evaluation procedure (project funding)

Stages in the evaluation procedure (project funding)

The SNSF Evaluation Policy

Funding Policies

- Open Science
- Promotion of early-career researchers
- Gender equality
- International cooperation
- Collaborative research

The SNSF and DORA

- DORA declaration signed in 2014
- Agreement in the presiding board to abolish journal-based metrics
- Several modifications (Guidelines, proposal forms)
- Compliance report 2018 on the implementation of DORA: **Room for improvement** in Ambizione, Prima, Eccellenza – but most likely not only there …

Implementation of DORA – revisions in the career funding regulations

- Handling of **academic age** (and biological age)
- Role and evaluation of **track record**
- Role and concept of **mobility**
- Assessment of the candidate's **chances of success** for an academic career
- **Broader impact** as criterion for use-inspired proposals

DORA

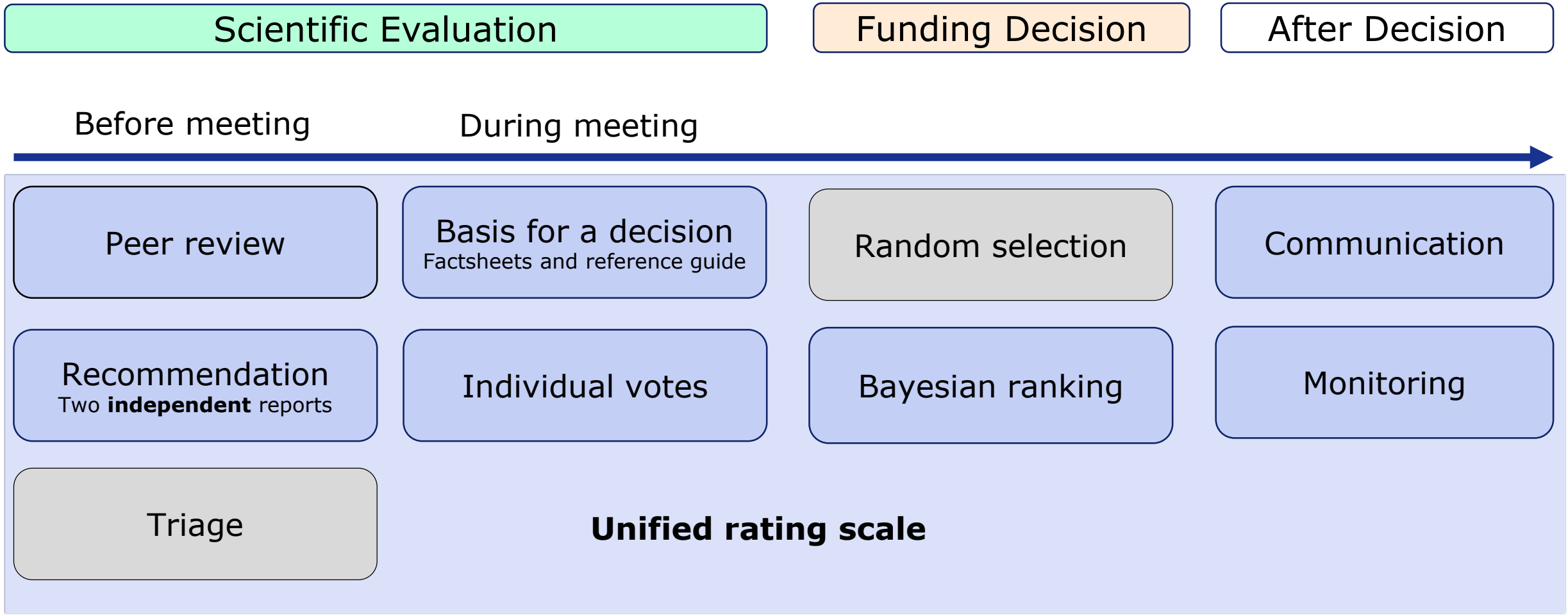
Best Practice

Consciously develop and regularly update one's approach to evaluation!

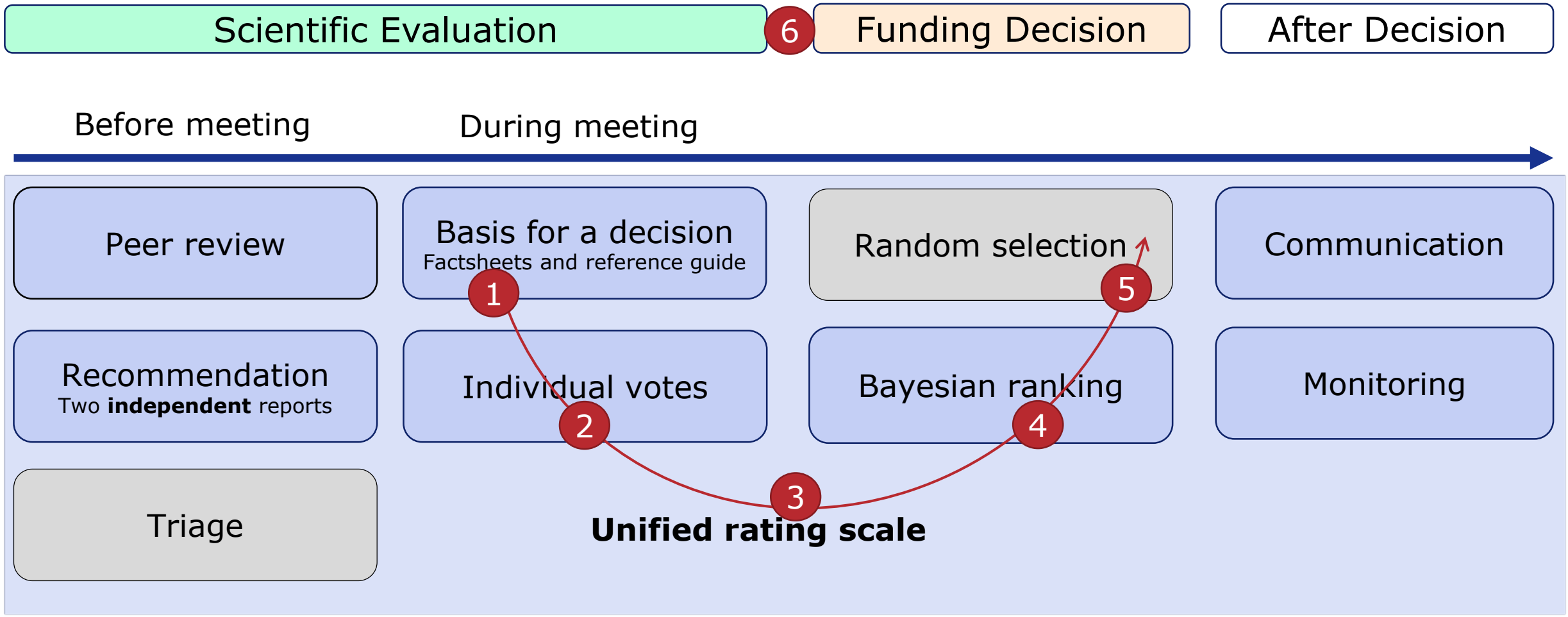
- Evaluation should be strategic and systematic, not haphazard, and it should be standardised across applications
- Metrics must be used with utmost caution, their content and limitations must be fully understood and comunicated. When used, they should be used consistently across applications and updated regularly
- Awareness of biases should be regularly refreshed, communicated and controled in a friendly manner within the group
- Inappropriate or informal commenting on proposals and smalltalk must be avoided
- Have the courage not to fund applications you are not convinced by
- Stand up for what you believe in while remaining open to the insights of others

Novelties and Future Challenges

A Unified Evaluation Procedure at the SNSF



A Unified Evaluation Procedure at the SNSF



Motivation for a numeric rating scale

1. Works at all stages (external review, recommendation, panel)

   → Coherence and transparency, facilitates monitoring

2. Allows direct statistical analysis to establish ranking

   → No hidden mapping of categories to numbers

3. Works for all panel configurations

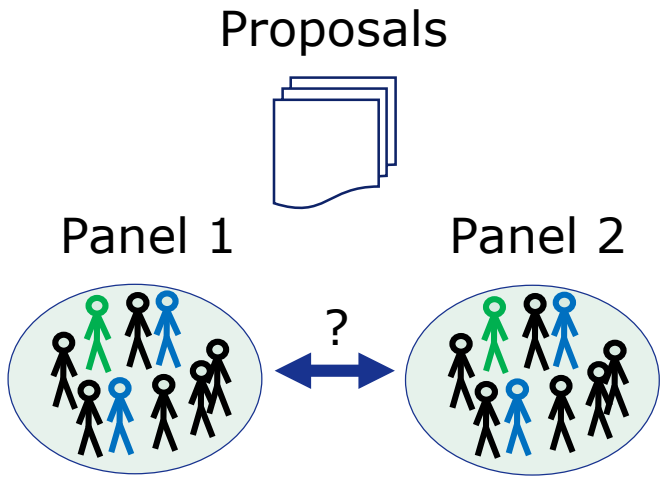
   → Allows for treatment of proposals in sub-panels

9-point numeric rating scale

Please provide a rating on the following scale regarding your assessment of [evaluation criteria]. 5 should be considered as the entry point; from that point, you should develop arguments to grade the [evaluation criteria] higher or lower.

9	Strong in all relevant aspects. No or negligible weaknesses.
8	
7	Strong in most relevant aspects. Few clearly identified weaknesses.
6	
5	Strong in several relevant aspects. Some clearly identified weaknesses.
4	
3	Some strengths in relevant aspects. Several clearly identified weaknesses.
2	
1	Few or no strengths in relevant aspects. Many serious weaknesses.

Randomness in evaluation?

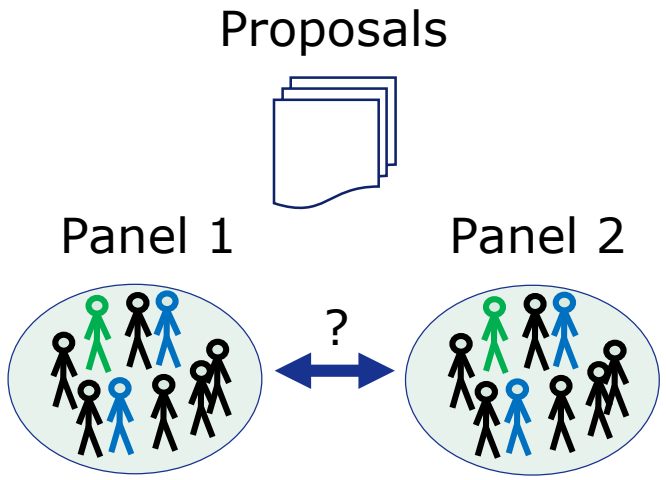


PANEL		Panel 2	
		fund	reject
Panel 1	fund	a	b
	reject	c	d

Agreement=(a+d)/(a+b+c+d)

Study	Setting	Type	Agreement
Cole, 1981	National Science Foundation	Grant proposals	
Hodgson, 1997	Canadian funding agencies	Grant proposals	
Fogelholm, 2012	Finnish Academy	Grant proposals	
Cortes, 2014	Machine learning conference	Abstracts	

Randomness in evaluation?



PANEL		Panel 2	
		fund	reject
Panel 1	fund	20	15
	reject	15	50

Agreement = 70 /100 (!)

Study	Setting	Type	Agreement
Cole, 1981	National Science Foundation	Grant proposals	70-76%
Hodgson, 1997	Canadian funding agencies	Grant proposals	73%
Fogelholm, 2012	Finnish Academy	Grant proposals	69%
Cortes, 2014	Machine learning conference	Abstracts	74%

Potential advantages of a random selection

- **Acknowledge** the **limitations** of **peer review**
- **Remove bias** against risky research
- **Reduce unconscious bias** (e.g., against women applicants) that resides in panel
- Correct for the "Matthew effect" whereby recipients of one grant are more likely to get another.

  "The system is already in essence a lottery without the benefit of being random"

- Increase efficiency

Bayesian Ranking

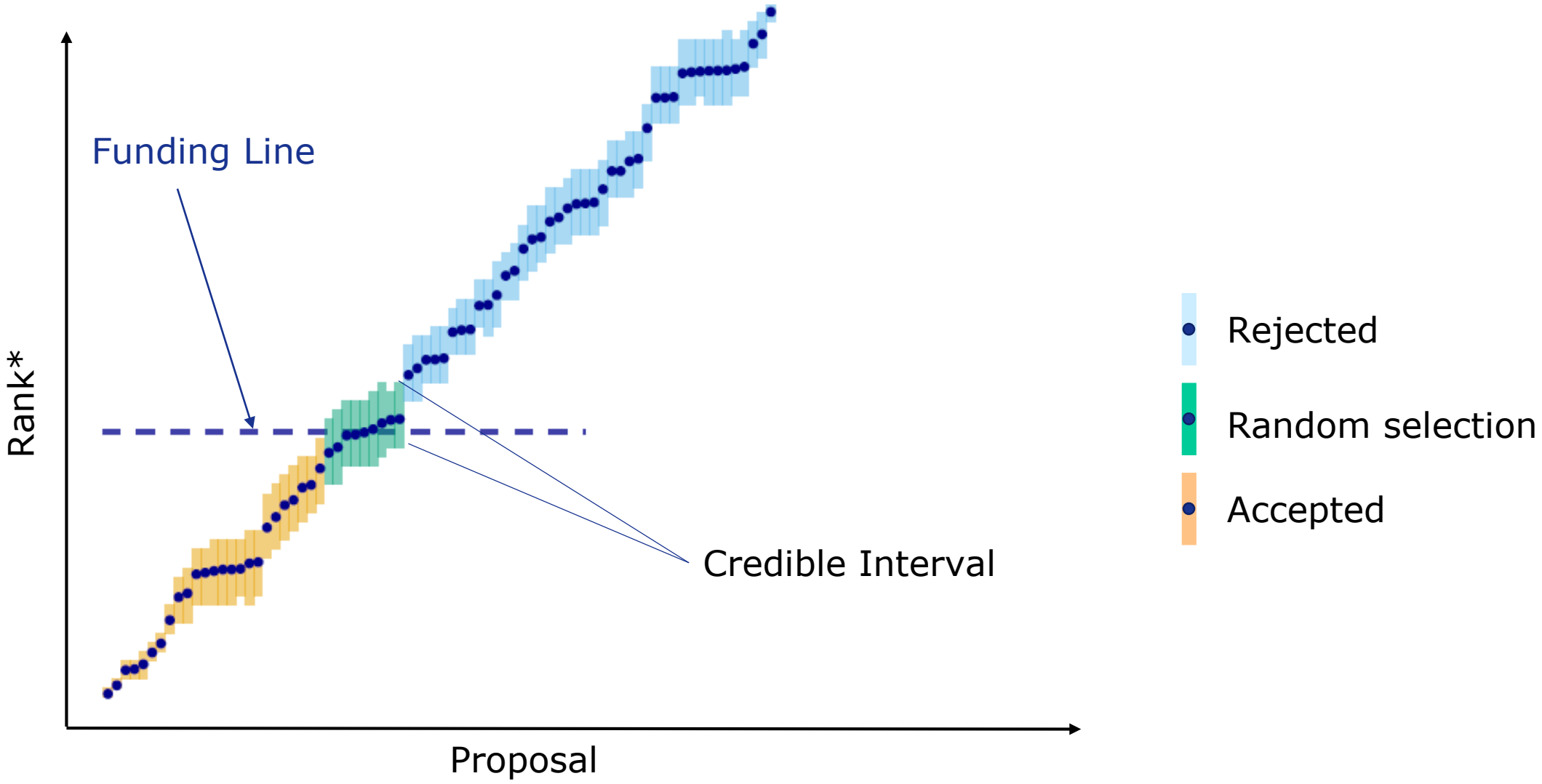
- There are many ways to create a ranking out of individual votes
- All have different advantages / disadvantages
- E.g. averages, intuitive, but not optimal

- The **Bayesian Ranking** (BR) is a statistical model that *increases fairness*

Answer 1
- BR compares each proposal with all the others to produce a relative ranking

Answer 2
- It provides a sound method to define random selection groups

→ by using credible intervals



A comment & further information

- BR is a help, a practical tool, a sound tool, a pretty exciting tool …

Rethinking the Funding Line at the Swiss National Science Foundation: Bayesian Ranking and Lottery

Rachel Heyard, Manuela Ott, Georgia Salanti, Matthias Egger

Funding agencies rely on peer review and expert panels to select the research deserving funding. Peer review has limitations, including bias against risky proposals or interdisciplinary research. The inter-rater reliability between reviewers and panels is low, particularly for proposals near the funding line. Funding

Submitted to Statistics and Public Policy, currently revised.
https://arxiv.org/abs/2102.09958

Drawing the Funding Line: Split Evaluation model

Strict separation of scientific evaluation and financial considerations:

Responsible Committee

(Division, Specialised Committee, SNSF Presiding Board)
Responsible for placing funding line

Scientific Panel

Responsible for
attributing of all proposals to quality groups

Drawing the Funding Line

The SNSF approved the splitting of scientific evaluation procedure in December 2019.

Advantages

- Clear distribution of tasks
- No mixing of scientific quality with financial/strategic elements
- Reduction of COI
- Fairer and more transparent outcome

Statistical procedure to help setting the **Funding Line**

Challenges inherent in Peer Review and Evaluation

- Bias (e.g. status and proximity bias; gender bias etc.)
- School of Thought
- Chance
- …

Implemented and planned measures

SciCV
FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION

DORA

About

In an effort to improve the evaluation of CVs, the SNSF will run a pilot programme to test a new, standardised CV format called SciCV. The pilot concerns all applicants for project funding in biology and medicine in this year's first call for proposals (submission deadline 01 April 2020). All applicants participating in this call, are required to submit their CV in the new format as part of their application.

At present, the formats of CVs submitted by applicants are heterogeneous and not always in line with international best practice. The aim of SciCV is to remedy this situation by allowing researchers to compile their CV in a structured way and to present their most important contributions to science in brief narratives, rather than only as lists of publications. This approach will help make other academic outputs, beyond publications, visible and valued and promote equal opportunities. SciCV will also introduce a uniform way of calculating the academic age of applicants, which indicates how long they have been active researchers as opposed to their biological age. The new format will no longer include any journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles. Rather, the actual content of articles and their citation impact will be considered

Implemented and planned measures

Pilot SciCV

New standardised CV format to improve evaluation

- Focus on all academic output not only publication list
- Focus on content of articles, not JIF

STEP Programme

SNSF **T**raining on **E**valuation **P**rogramme

- Coordinated training programme running over 2 years
- Input by international experts on research assessment
- Topics incl. bias, conflict of interest, ethics
- Resources publicly available for use at home institutions

THANK YOU!

Swiss National Science Foundation