

Judges Disagree: Engineering a Collective Decision Without Enforced Consensus

A proposal by Ahmad Issam Maadarani for diagnosing disagreement and documenting its resolution

Official article information

Article author Ahmad Issam Maadarani	Article number 8303	 Scan to verify the article
Article code —	Protection number —	
Reference —	International Union Of Arab Master Chefs	

By Ahmad Issam Maadarani

Panel agreement does not establish correctness, and disagreement does not establish failure. This article classifies disagreement before resolving it: sample, description, criterion application or weighting policy. Each type receives an appropriate procedure, with reasons for decision changes preserved. This professional design does not claim that organised discussion removes bias or guarantees fairness.

1. Why an average is insufficient

Widely separated scores can arise from different samples, different definitions of required texture or weights outside the rules. Averaging does not identify these causes and cannot remedy every disagreement. A task may require aggregation, but explanation must establish what is being aggregated. Collective judgment begins with diagnosis, rather than

silencing a strong voice or telling a dissenting judge that the majority suffices.

2. Four types of disagreement

Sample disagreement concerns what each judge received and how it was tasted.

Description disagreement concerns language used for the experience or different recorded sensory experiences. Criterion disagreement concerns the requirement's meaning and application. Weighting disagreement concerns importance or ranking rules. Types may coexist and should not be forced into one category. Personal preference remains separate information, entering the decision only to the extent permitted by the task.

3. Independent observation before discussion

Where the procedure specifies it, judges record initial observations and scores before seeing others' reports. A bare number is insufficient: discussion needs a criterion-linked description. Present reasons without treating seniority as substitute evidence. Speaking order can vary, and everyone should state what supports their observation and what could change it. Independence documents the starting point rather than proving freedom from every prior influence.

4. Hypothetical case: three scores for one dish

Suppose scores are 6, 8 and 9 on the prescribed scale. The first judge reports dryness, the second required cohesion, the third distinctive appearance. Review shows the first tasted an edge and the second the centre with sauce, while the third included appearance in a texture criterion. There is a possible sampling issue and a criterion-application issue. Review the third assessment against the rule; document variation by location or re-examine if permitted. Do not automatically relabel 6 as severity or accept 9 as creative enthusiasm: each decision needs its own reason.

5. Three resolution routes

For disagreement about a documented fact, review information or available samples within the procedure. For criterion application, return to the written description without revising it after seeing the dish. For remaining relevant sensory disagreement, preserve it and apply the declared aggregation or tie-breaking rule. No method is best for every scale:

a median may fit a particular ordinal treatment, and a mean may be a conventional index elsewhere. Both need explicit rules, and neither should be selected after inspecting which yields a desired result.

6. Recording change without artificial consensus

Preserve initial score or description, disagreement type, new information, final decision and reason for change or non-change. A changed opinion should identify changed evidence rather than merely cite panel agreement. Preserve unresolved disagreement internally. A single final outcome under the rule can coexist with multiple explanations. Productive agreement reveals how the problem was addressed; artificial consensus hides it and obstructs learning and review.

7. Testing procedural usefulness

Two panels could train on comparable cases, one using the disagreement record. Review reports against predefined indicators: clear reasons for changes, detection of unsuitable criterion application and preservation of evidentially unresolved disagreement. Also examine time cost and needless complexity. Greater agreement may reflect group pressure rather than improved fairness. Examine reasoning and adherence to the task, not just the final pattern of scores.

8. Conclusion: a panel need not be one voice

A panel's strength lies in turning multiple observations into an explainable decision. It need not eliminate every difference; it should stop differences in conditions and rules hiding within numbers. When judges know their contribution and why anything changed, collective judgment becomes clearer. Its legitimacy rests on declared procedures and documented reasons, not the number of people repeating the same sentence.

9. A concise working sheet

Field	Review question	What to record
Starting point	What did each judge record initially?	Preserve description and any initial score.
Sample	Did everyone taste comparable cases?	Document locations and conditions without assuming identity.

Field	Review question	What to record
Criterion	Was the same written requirement applied?	Correct application without rewriting the criterion afterwards.
Information	What did discussion actually add?	Separate new evidence from pressure to agree.
Change	Why did an opinion change or remain?	Link the decision to a reviewable reason.
Aggregation	Which final-result rule was declared?	Apply it while preserving unresolved disagreement.

Suggested citation: Maadarani, Ahmad Issam (2026). “Judges Disagree: Engineering a Collective Decision Without Enforced Consensus” Methodological article, version 1.0.

© 2026 Ahmad Issam Maadarani. All rights reserved, subject to applicable legal exceptions.



IUOAMC
International Union Of
Arab Master Chefs