C U R A T E

# A framework for taste

*Taste is a practice, not a privilege.*

---

SIX DIMENSIONS · THIRTY RUBRICS · TWENTY-FOUR EXERCISES

FIRST EDITION

CURATEFRAMEWORK.COM

CONTENTS

# What is in here

PART

# I

# The judgment bottleneck

Production capacity became effectively infinite. Judgment capacity did not move. Everything else in this document follows from that single reversal.

For as long as anyone has made things professionally, making was the expensive part, and everything else arranged itself around that fact. Teams were sized for production. Budgets were written for production. Process was designed to protect production from waste, and quality control sat as a thin layer at the end, because the cost of building anything guaranteed that not much got built.

Judgment was never the bottleneck under those conditions. One creative director could review everything, because everything was not very much. A senior designer could set the standard by touching each piece of work personally. Taste stayed in people's heads and that sufficed, because the volume of decisions requiring taste was bounded by how slowly things could be built.

That constraint is gone. A team that produced twelve concepts a quarter now produces twelve before lunch, and the twelve are competent. The marginal cost of another variant, another draft, another asset, another feature is close to zero and

still falling.

What did not change is the number of people in the organisation who can look at any of them and say which one is worth building, or why the other eleven are not. That number is the same as it was three years ago. It may be lower, because the people who had that capability now spend their time reviewing generated output instead of developing the judgment that made them worth asking.

> Organisations used to have more judgment than they had output to apply it to. Now they have more output than judgment. That reversal explains something most teams have noticed without quite being able to name it: work got faster and quality got flatter. Not worse, in most cases. Flatter. More things arrive at a competent middle, and fewer arrive anywhere else.

## Why the usual responses fail

**Hiring more reviewers**

You hire two more senior designers. Output has gone up eightfold and review capacity by perhaps a third, and both new hires now spend their afternoons catching things a written standard would have caught before the work reached them. Headcount buys a linear improvement against a problem that stopped being linear some time last year, and it is the most expensive option available.

**Better prompting**

Prompting raises the floor, which is worth having. A team that writes careful prompts gets competent output more reliably than one that writes lazy prompts, and competent is a long way from distinctive. The model returns the middle of what it has seen, and a sharper instruction returns that middle more precisely. Teams who treat this as a prompting problem usually plateau within two quarters and conclude the tools are not ready yet.

**Design systems and brand guidelines**

Ask what your design system is actually for. It stops twelve people building twelve different buttons, and it does that job well. It will also ensure that everything you ship is coherent, on-brand, and very hard to tell apart from what your competitor ships, because they bought a system too and it solved the same problem the same way. Consistency was the right answer when the risk was fragmentation.

**Waiting for the tools to improve**

They will. They will get better at producing the middle, which raises the floor for every company at once and lowers the value of standing on it. Faster generation does not relieve a bottleneck in judgment. It widens it.

## What follows

If judgment is the constraint, then judgment is the thing to build capacity in, which means treating it as a capability rather than a personality trait: something

that can be described, distributed, taught, embedded in process, and configured into the tools themselves.

Taste has been allowed to remain mysterious, which is a different problem from being mysterious. The people who have it are not incentivised to explain it, and the people who lack it are not equipped to ask. So every organisation holds some quantity of taste inside particular individuals, undocumented, unmeasured, and structurally unable to influence more than a fraction of the decisions being made in its name.

No framework produces taste, and this one does not claim to. What a framework can do is move taste out of individuals and into the organisation, which is the only form in which it survives contact with the volume of work you are now producing.

---

Distinction is a relative property. It is not enough for your work to be good. It has to be identifiably yours in a market where everyone has the same tools, the same models, and the same reference material. That was always somewhat true. It is now the entire competitive question.

PART

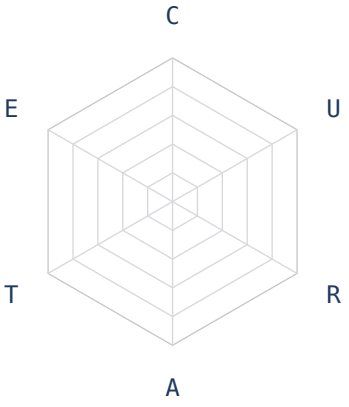
II

# The six dimensions

Taste, decomposed. Six behaviours you can observe in a finished piece of work, score against a rubric, and write into a brief or a hiring bar.

They describe how a decision was arrived at, and say nothing about what the decision should have been. They read the same way whether a person or a model produced the draft, which is why each one carries a note on how it fails in generated work. What gets scored is how your team decides, not which tools it uses.

They are deliberately non-overlapping. Two concern what happens before work begins, two describe properties of the artefact, one concerns the encounter with an audience, and one concerns time. Where they seem to touch, the distinction is one of scope rather than of kind.



DIMENSION C

# Confidence

## *Does the work trust itself?*

Hedged work is rarely wrong, which is why it survives review. The extra label, the second call to action, the explanatory subtitle, the tooltip that exists because nobody was quite sure the interface was clear: each addition was defensible on its own, and the accumulation reads as anxiety. Audiences register hesitation before they register content.

**In AI workflows**

A model satisfies every clause of a prompt, so its default output is hedged and the constraint has to be negative and explicit. Telling a model to be clear produces more explanation. Telling it to cover a single interpretation and omit the alternatives produces less. Element caps and prohibitions belong in your system prompts rather than in your review comments.

**Cost of movement: Low**

Confidence moves faster than any other dimension because it requires no reorganisation and no new authority. It is a review-discipline change, and the results are visible within one cycle.

REFERENCE SCORE 2.8 / 5.0 RANK 2 OF 6 COST LOW

Reference values are the framework's own baseline, not measured averages. See the appendix.

DIMENSION U

# Unique

## *Could only you have made it?*

Nothing is wrong with generic work. Nothing identifies it either. The failure is not incompetence but interchangeability: replace the logo and no one notices, including the team that made it. This is the dimension most reliably dismissed by teams whose work looks professionally acceptable.

**In AI workflows**

Convergence on the centre of the distribution is the mechanism, not a defect in it, so distinctiveness has to be supplied as input rather than requested as an instruction. A prompt asking for originality returns the most conventional available form of it. What changes the output is grounding: your own prior work, your own reference set, your own constraints, provided as material the model works from. Teams that treat this as a prompting problem stay at Level 2 indefinitely.

**Cost of movement: High**

Unique requires changing where the team looks, which is a hiring and habit question rather than a process question. Expect four quarters before movement is visible externally. It is also the dimension with the largest commercial consequence, which is why it is worth the timeline.

REFERENCE SCORE 2.4 / 5.0 RANK 5 OF 6 COST HIGH

Reference values are the framework's own baseline, not measured averages. See the appendix.

DIMENSION R

# Refusal

## *What did you decide not to make?*

When production costs approach zero, declining becomes the only scarce act. Refusal sits upstream of editing, since it concerns what never enters the pipeline at all. Its absence is invisible, because nothing that was never built appears in any review.

**In AI workflows**

Generating another variant, feature, or asset now costs nothing, which removes the friction that used to enforce restraint. The constraint has to be reintroduced deliberately, at the point of proposal rather than the point of production. The question has changed from whether the team can build something to whether anyone would have asked for it when building was expensive.

**Cost of movement: Moderate**

Refusal requires that someone hold a position against a requester, which makes it partly an authority question. It moves faster than Unique and slower than Confidence. It is also the dimension most improved by a single structural change: a standing forum where "no" is an available outcome.

REFERENCE SCORE 2.3 / 5.0 RANK 6 OF 6 COST MODERATE

Reference values are the framework's own baseline, not measured averages. See the appendix.

END OF SAMPLE

# That was the argument.

You have read why judgment, rather than production, is now the thing in short supply, and three of the six dimensions that decompose it. What the sample does not contain is the part you would actually use.

**Thirty rubric entries**
Five levels for each dimension, written to be scored from work rather than from self-report. They describe behaviour in enough detail that a disagreement has somewhere to land.

**Twenty-four transitions**
One bounded exercise for every step up, on every dimension, with an honest note on what each costs in time and political capital. Some are free. One is a reorganisation.

**The maturity model, in full**
Which level is worth aiming at, why the top one is not, and the specific failure mode of a team that has systematised its taste and stopped noticing the system has expired.

**Configuring AI systems**
The six dimensions rewritten as explicit constraints for a system prompt or an evaluation rubric. Generic instructions toward quality return the average of everything already considered good.

**The instrument**
All eighteen questions, the scoring arithmetic, the band boundaries, and the reference profile.

> `curateframework.com`
>
> The assessment is free and takes about four minutes. It scores your team across all six dimensions and returns one exercise for whichever dimension needs it most. The complete framework is available on the same site.
>
> *Taste is a practice, not a privilege.*