Not too long ago, I met a friend for a night out at the cinema. It soon became apparent that our film choices for the night were at odds: She wanted to see David O’Russell’s *American Hustle*, whereas I was more enthusiastic about Shane Carruth’s *Upstream Color*.

**[SLIDE 1]**

This was not the first time we wanted to see different films, and as with previous disagreements in our trips to the cinema, I called upon my background to try to get some leverage on the matter: I don’t know—I said—I read the reviews for both, and I’m not quite sure about *American Hustle*... My friend promptly got her phone out and said as she typed — I can prove to you, statistically, that *American Hustle* is a better film than *Upstream Color*— To which I could only stand there in bemusement. It turns out she had had enough of my tyranny of taste and found a way of getting her own leverage in the way of the review aggregator website Metacritic ([www.metacritic.com](http://www.metacritic.com)), where American Hustle scores an impressive 90%, and Upstream comes short at 81%.

**[SLIDE 2]**

As far as she was concerned, this settled the matter (as well as every future discrepancy in film selection), since she need not rely in her one, vaguely knowledgeable but ultimately biased friend, now that she had the balanced aggregate of many informed opinions at her fingertips.
As tempted as I was to point out the many problematic assumptions in the world of review aggregation, I resisted the urge to argue, and, instead, found myself in the dark a few minutes later thinking not about David O’Russell but about David Hume. More specifically, about his 1757 essay: The Standard of Taste.

[SLIDE 3]
This short essay has become one of his most cited works, and it is considered by many one the founding texts of modern aesthetics. In what follows I will be dealing with its central problem and Hume’s proposition to tackle it, namely, his solution to the paradox of taste. In terms of Metacritic, I am not interested at this point in either discrediting or validating its aggregation methods. Instead, my argument in this short presentation is that in Hume’s approach to address the problem of aesthetic judgment, there is a key intellectual gambit whose affordances are finding new substance in the wake of digital technologies and the networks they enable.

[SLIDE 4]
The paradox of taste

Perhaps one of the factors that elevated Hume’s Standard of Taste to a philosophical classic is the elegant simplicity of the paradox at the center of the essay, a conundrum that is easily observed in our encounters with art: while we concede that taste is subjective in nature, and therefore judgments of taste—as far as they are not accessed through natural properties of objects but rather sentiments that arise from our appreciation of these objects—cannot be compared in value, De gustibus non est disputandum [There is no accounting for taste].

---

1 There is some consensus among Hume’s scholars that his aesthetic thought is deeply embedded into his wider philosophical system—other essays like the Sceptic (1742), and particularly his Enquiry Concerning Human Understanding (1748)
Hume puts it as follows:

“each mind perceives a different beauty. One person may even perceive deformity, where another is sensible of beauty; and every individual ought to acquiesce in his own sentiment, without pretending to regulate those of others.”

And yet, Hume goes on to argue, common sense simultaneously tells us that some works of art must better than others, or at least we seem to equally endorse this proposition in our behavior:

[SLIDE 5]

“Whoever would assert an equality of genius and elegance between Ogilby and Milton, or Bunyan and Addison, would be thought to defend no less an extravagance, than if he had maintained a mole-hill to be as high as Teneriffe, or a pond as extensive as the ocean. Though there may be found persons, who give the preference to the former authors; no one pays attention to such a taste; and we pronounce, without scruple, the sentiment of these pretended critics to be absurd and ridiculous.”

Confronted with a seemingly irreconcilable variance in people’s taste that is somehow bound to an equally apparent consensus about some objects being more pleasing than others, Hume suggests the need for a standard: “a rule, by which the various sentiments of men may be reconciled; at least, a decision afforded, confirming one sentiment, and condemning another”. In Hume’s view, such a procedure would allow conflicting parties to come to some form of resolution in matters of taste, an ambitious agenda, and one that some critics consider a departure from his earlier accounts of either aesthetic skepticism, or radical empirical subjectivism.
Whatever his motivations, Hume’s solution to the paradox is highly original: simply put, he claims that although we cannot condemn sentiments, for they are real and refer to nothing beyond themselves, we can try to assess the critic’s ability to channel these sentiments in a consistent manner, and then, crucially, aggregate these judgments to reach a standard with which to evaluate any given work of art. He then spends the rest of the essay discussing the critical abilities a “true critic” must meet. Hume’s standard is thus revealed as the joint verdict of all the critics that meet these conditions:

“Strong sense, united to delicate sentiment, improved by practice, perfected by comparison, and cleared of all prejudice, can alone entitle critics to this valuable character; and the joint verdict of such, wherever they are to be found, is the true standard of taste and beauty”

So when my friend and I found ourselves in disagreement as to the aesthetic merits of the films in question, we could, in order to resolve the dispute, gather the consensus of reliable critics, ruling out those who are prejudiced, or who lack intelligence (understood here as reasoning faculties); those who are inexperienced, or lack delicacy, attention to detail, or refinement. This procedure would work, according to Hume, because while a critic’s response is highly subjective and beyond dispute, her ability to judge a work of art is not, and her compliance of the precluding conditions as well as her place in the consensus of all her fellow critics, can be therefore ascertained through objective means.

From this example it follows that, in its design, Metacritic and other aggregators are remarkably close to enacting Hume’s standard, not only in their enabling of consensus by summoning the appraisals of critics “wherever they are to be found”, but also in trying in their own distinct ways to screen for “true critics”.

[SLIDE 7]
Metacritic in particular makes for an interesting case-study, since it distinguishes very clearly the reviews of critics from the reviews of non-critics, and it offers some kind of criteria for discrimination—not too distant from Hume’s conditions—to determine who is allowed into the critic’s list. This is, of course, hardly a transparent process: who makes the list is determined by the critic’s or publication’s “quality and overall stature”, a vague definition that nevertheless reminds us of Hume’s empirical verification of his conditions:

[SLIDE 8]

“we noticed that some critics consistently write better (more detailed, more insightful, more articulate) reviews than others. In addition, some critics and/or publications typically have more prestige and respect in their industry than others. To reflect these factors, we have assigned weights to each publication (and, in the case of movies and television, to individual critics as well), thus making some publications count more in the METASCORE calculations than others.”

In general terms, Metacritic works by pulling a number of reviews from this list of trusted sources, and then proceeds to average the results by either consolidating the scales into a percentage (e.g. four out of five stars becomes 80%), or by manually assigning percentage values when critics do not use a scaling system (SALON magazine or the New York Times for example). The result is aptly called a metascore, which is meant to reflect a more accurate picture of a film’s critical reception.

[SLIDE 9]
There are two important aspects note here:

1) Even in the face of potential deficiency, expressed as the inevitable range of prejudice or lack of consistency inherent to subjective judgments, the weighted average system of Metacritic aims to fend against critical aberration by accounting for it in the system, that is, by manipulating the maximum possible influence that any given critic has in the final metascore. This can, in principle, normalize certain prejudices, for example, if a critic favors Hollywood comedies over Latin American social dramas, and if we assume (and maybe even find proof) that he therefore tends to give higher scores to the former than to the latter, we could correct this prejudiced response by attaching something like a bias index function, that would adjust his overall degree of influence in the calculation of the metascore, thus subsuming a perceived critical deficiency into the aggregation system.

2) In the case of Metacritic, this formula to calculate the weighted average of reviews is something of a black box, what the website’s co-founder and editor Marc Doyle calls “Metacritic’s "secret sauce"”, when he refuses to disclose the formula itself.

Reading Of The Standard of Taste against a digitally rendered critical landscape, one is left under the impression that Hume could have been an eerily prescient advocate for data science. In his essay he eschews the transcendentalist attempt to define beauty in order to identify the beautiful, and provides us instead with a probabilistic system with which to compare and pass on judgment; a set of instructions (or an algorithm) and of conditions (or variables) designed for the very practical task of settling aesthetic disputes.

This is a fascinating move, not only because it shifts towards scientific discourse in both its normative aspirations and its empirical groundings; an attempt is made, as Noel Carroll writes, to “bridge the gap between fact and value by making the standard of taste a matter of empirical discovery” but also, as we
will see, because it helps us challenge assumptions genealogically inherited from this particular intellectual strategy.

A closer look at Metacritic sharply puts into focus the underlying tensions in taking this leap of breaking down a large subjective and unquantifiable problem, such as taste, into a set of seemingly objective and measurable problems: what happens, for example, if even after finding our true critics they are unable to come to an agreement? —As Hume himself foresaw. Or, what are we to think of extreme cases, where a film scores a full one hundred, or zero, in its metascore? In fact, what exactly is a metascore at the ontological level? And how do we, the meta-meta-critics, place this within our intellectual frameworks?

In examining the way these aggregators work, a number of assumptions arise about both the nature of criticism and the possibilities of computing. One the one hand no respectable critic will be satisfied with an exploration of a work of art that yields a “60% good and 40% bad” verdict, and rightly so, for these are not terms that are in any way useful for them, nor would we wanted them to be. I think it is fair to say we seek criticism in an attempt to be persuaded by nuance and critical insight, by sets of informed and well-organized propositions that we know are created to test our beliefs, and yes, our taste, and so we engage these propositions in playful negotiation; a game that very often finds fertile ground in the irreducible ambiguity of critical language, the very same ambiguity that computer science, on the other hand, seems to abhor.

Computational systems, we are told, preclude contradiction at their core, and so their design goes to great length to sample, summarize, break down and, ultimately contain paradox, if only to avoid the much feared “blue screen of death”, or fatal system error, produced when a program attempts an undefined operation such as dividing by zero.

[SLIDE 10]
From this standpoint, it makes sense to think of paradox as something uncomfortable, crippling, or otherwise problematic in computing. And it bears saying that for the purposes of disambiguation, isolation, automation, normalization, and other operations regarded as fundamental to scientific discourse, the underlying instabilities at the center of many paradoxes can indeed be undesirable.

Maragaret Cuonzo foregrounds this inclination to contain ambiguity in her Bayesian appraisal of Hume’s paradox of taste, this approach, she argues:

“allows for disagreements about equally acceptable valuation systems but also allows that some super-values may bind the different acceptable valuations together. Thus, the supervaluationist way of solving the paradox of taste holds that although it may be adjudicating among matters of taste is not possible, we can, using supervaluationism, still determine objectively whether some artwork is good or bad”.

Without going into the otherwise fascinating details of supervaluationism, note that the solution put forth by Cuonzo is still constructed under the assumption of discrete oppositional terms, as she herself explains:

“Respect for theoretical and observational connections is needed in order to avoid conflict with experience and the specifications of other terms; for example, an interpretation that treated the term beautiful as referring to a particular object must not also treat the term ugly as applying to the same object”.

"beautiful" and "ugly" seem to indicate terms belonging to the same category, but they are actually talking about different concepts.
Yet, art criticism seems to operate as much through language as a set of self-identical terms, as through its array of instabilities, take this example from Peter Bradshaw's review of Upstream Color:

“It is invigoratingly freaky and strange, with a Death-Valley-dry sense of humour somewhere underneath — though a little derivative sometimes. More than once, Carruth gives us a close-up on a hand ruminatively stroking a surface: very Malick. And the shots of creepy creatures swarming under the skin are very Cronenberg.”

If one is accustomed to film criticism, it is clear that Bradshaw’s stance on it is generally positive, even when there is very little semantic exclusivity in his terminology. “Freaky”, “Strange”, Malick-like, Cronemberg-like, are all indeterminate qualifiers in themselves, and only derive their expressive powers from context. Criticism is evocative in this manner not in spite of language’s vagueness, but rather because of it. Or put in a different way, criticism’s precision is paradoxically linked to its skillful handling of infinite possible worlds without giving into the temptation of breaking them apart.

At this point Carroll’s reading of Hume becomes very useful, as he identifies that the Scottish philosopher conflates two distinct operations in his *Standard of Taste*: liking and assessing. One can assess, Carroll argues, “the symbolic density of a given poem with no special sensation of pleasure or enjoyment and with no disposition to be moved to like the poem […] just as one can note that a given checkmate is a masterpiece of ingenuity while neither being moved by it nor taken with chess in general”. In my view, this distinction is critical, for it throws the problem back to the realm of language: liking invokes opposition; either one likes a film or not, considering perhaps the degrees of liking in between, like Cuonzo suggests and Metacritic to some extent enacts. Assessing however, is not linear
in this way, and tends to invoke inter-subjective possible worlds, as exemplified in Bradshaw's review, and to an extent in Hume's own qualms about his desired standard.

Aesthetic assessment is thus revealed as a profoundly human activity, one that is fraught with paradox and very resistant to formalisation. In order to convincingly simulate an art critic through computation, I argue, we need a different understanding of what computation is; at a the philosophical level, we need to countenance the fundamental problems of ambiguity, vagueness, and paradox, which haunt the relation between counting and naming, between reference and representation.