

DYNAMICS-WEB: A Pre-Registered Measurement of Web Content Character over the Link Graph, and a Powered Proof that its Sub-Perceptual Half Changes Conversion

Jason Duke

jason@kronaxis.co.uk

Kronaxis Limited

Research Article

Keywords:

Posted Date: September 8th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10502110/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: Competing interest reported. The author is the founder of and is affiliated with Kronaxis Limited, which develops commercial technology related to the methods described in this paper, including a proprietary personalisation mapping that is not released. The author has no relationship with this journal. This affiliation had no bearing on the design, analysis, or reporting of the pre registered experiment, whose protocol and falsification criteria were fixed in advance of data collection.

DYNAMICS-WEB: A Pre-Registered Measurement of Web Content Character over the Link Graph, and a Powered Proof that its Sub-Perceptual Half Changes Conversion

Jason Duke

Kronaxis Limited

Corresponding author: jason@kronaxis.co.uk

ORCID: 0009-0002-9293-6529

Abstract

Search reduces a web page to two properties, its authority and its topic. We show the same link graph carries a third: the character of a page, how it reads across eight axes with subject and standing held fixed, and that the deepest part of this signal changes what a reader does without ever being consciously seen. We present DYNAMICS-WEB, a measurement of web content character, and establish five results in support of one claim. Character is distinct from authority, which reconstructs it at a coefficient of determination of 0.030, and from a full surface battery at 0.256. Its axes transfer directly to Arabic, Chinese and Swahili, and the same engine recovers geographic relevance for forty six countries, so the method is general. Three foundation models from three developers recover five to six of the eight axes in common, so the construct is not an artefact of one teacher. Character propagates through the link graph at about 0.4 per axis against a null of zero. And in a pre registered, powered choice experiment ($n = 117$), matching the sub perceptual axes of a page to a reader raised conversion from nineteen to thirty one per cent (one sided $p = 0.0295$), while the loud commercial axes carried none of the lift. We are explicit about the bounds: the conversion effect is modest, a secondary alignment did not replicate, and two axes do not transfer cleanly. The measurement is released in full; a single downstream personalisation function is proprietary and is not needed to reproduce any result here.

Keywords: web content character; link graph propagation; topical authority; sub perceptual signals; choice experiment; pre registration; conversion; cross model validation; convergent validity; web measurement; DYNAMICS-WEB; DYNAMICS-8

0.1 1. The missing axis

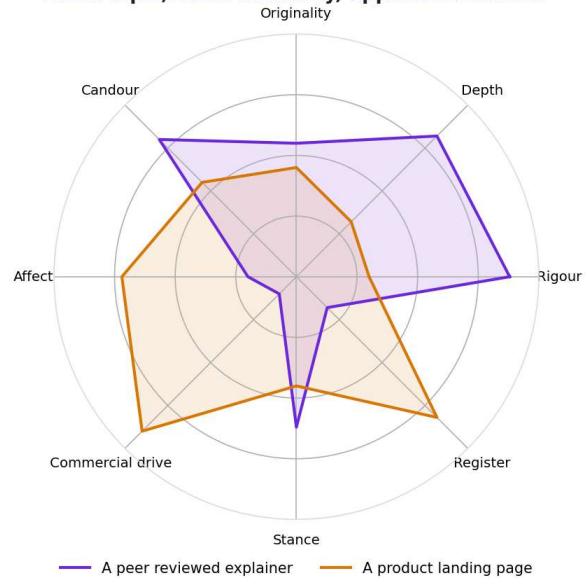
A page carries information you can rank on. Its authority is one: how much standing it holds in the web at large, the property PageRank and its descendants measure. Its topical authority is a second: how much standing it holds inside a subject, the property of topic sensitive PageRank and, in the commercial world, Majestic's Topical Trust Flow. Both are real, both are useful, and both share a shape. They are global. They reduce a page to a number, or a vector of numbers across topics, that is the same for everyone who looks. That is exactly what you want from a filter and exactly what a filter cannot do: tell two different readers apart.

Three properties of a web page



There is a third property of a page that those signals do not touch. Two pages can sit on the same topic, at the same authority, and still read in completely different ways: rigorous or warm, original or derivative, selling or explaining. That difference is not authority and it is not topic. It is **character**: how a page reads, across the eight axes defined in section 2, and nothing in the standard toolkit measures it.

Same topic, same authority, opposite character



DYNAMICS-WEB is a purpose built model for that third axis. It is not a personality test for web pages and it is never to be presented as one. A page does not have a personality the way a person does. It projects a character, assembled from the author's manner, the subject, the venue it sat in, and the moment it was written. The model measures that projection.

0.1.1 1.1 In the context of prior work

Two decades of web measurement have been about standing. PageRank and the link analysis that followed it, including HITS, score a page by the structure of the graph around it: a page is important because important pages point at it. Topic sensitive PageRank narrowed that to a subject, scoring standing within a topic rather than across the whole web, and the commercial world reproduced the idea as Majestic's Topical Trust Flow. All of these answer the same kind of question, how much authority a page holds, globally or within a topic, and all of them return one value per page that is the same for every reader.

A separate body of work reads the text itself rather than the graph. Readability indices such as Flesch reading ease score how hard a passage is to read; sentiment analysis scores its emotional polarity; stylometry attributes authorship from surface regularities; and a long line of text quality and credibility classifiers, more recently folded into search under the banner of

expertise, authoritativeness and trust, try to separate good pages from bad. Each of these touches how a page reads, but each isolates a single dimension, usually difficulty, or polarity, or a binary quality verdict, rather than the several independent axes on which content actually varies. Section 5 shows directly that the surface statistics of readability, sentiment, length and topic together reconstruct only about a quarter of character, so character is adjacent to this work but not reducible to it.

The third idea in this paper, that character should be treated as something that flows through the link graph, draws on a different tradition again: the propagation of node attributes across a graph, and the long standing observation that connected nodes tend to resemble one another, which the network literature calls homophily. Authority was the first web property understood as something that flows; we argue character is another, and section 4.1 tests that premise directly rather than assuming it.

DYNAMICS-WEB sits where these three meet. It is not an authority signal, not a topic signal, and not a single readability or quality score. It is a multi axis measure of the character of content, anchored on human judgement, carried to the whole web by a teacher model, and propagated through links in both directions. To our knowledge that specific combination, character as a measured, link propagated, per axis property of web content, has not been set out before.

0.1.2 1.2 What this paper establishes

The paper makes five claims and tests each. First, **topical standing is public**: we compute how much authority a page holds inside its own subject, over the whole graph, and serve it openly. Second, **authority is public**: the classical signal is served beside it, on the same toolbar. Third, **character is a distinct third axis**, neither authority nor topic, that no standard tool measures and that a full battery of surface statistics recovers only about a quarter of. Fourth, **the link graph carries more than authority**: character propagates from a domain to its neighbours at about 0.4 per axis, and the same propagation, run on a different attribute, reconstructs geographic rele-

vance for forty six countries, so link propagation is a reusable engine rather than a trick that works once. Fifth, and the one the others build to, **the unseen half of character changes behaviour**: people feel character but cannot reliably name it, which marks it sub perceptual, and in a pre registered, powered experiment, matching that sub perceptual half to a reader raised conversion from nineteen per cent to thirty one (one sided $p = 0.0295$) while the loud, nameable axes carried none of the lift. The fifth finding is the keystone, and the paper is ordered to reach it: a measurement, a propagation method, and then a behavioural proof that what only we can measure is what actually moves the reader.

0.2 How we got here

This model began as a correction to a mistake of our own. An early Kronaxis scorecard carried a dimension that tried to score web pages on a human personality model, borrowed wholesale from the disposition model described in section 6. It did not work, and the reason it did not work is the reason this paper exists: a page does not have a personality the way a person does, and an instrument built for people puts wrong signs on real distinctions when it is pointed at text. Recognising that the object was different, that content projects a character rather than possessing a disposition, was the first step.

So we separated the two instruments. The human model was left exactly as it was, for people. A new model was built from nothing for content, on the axes content actually varies on rather than the axes a personality test happens to carry. We anchored it on 249 domains scored by hand, and then met the two problems that shape everything after: reaching volume, and choosing a teacher.

The teacher problem turned out to be the hinge of the whole project. A smaller teacher of a few billion parameters distilled cheaply but failed the honesty test, reading academic pages as less rigorous than the open web because it confused dense salesmanship with care. Only a teacher of roughly twenty seven billion parameters put the sign back the right way, and that single correction, shown later in figure 5, is the differ-

ence between a model that measures character and a model that measures fluency. From there character was propagated across the entire public web graph and, finally, built into a browser toolbar so that anyone can see it on any page. The sections that follow take each step in turn.

0.3 2. The model

Each domain with readable content receives a character profile across eight axes:

- **Rigour:** how careful, precise and evidenced the writing is.
- **Depth:** how much substance sits under the surface.
- **Originality:** how much is fresh rather than restated.
- **Candour:** how open and direct it is, versus hedged or evasive.
- **Affect:** how much emotional warmth and expressiveness it carries.
- **Commercial drive:** how strongly it is trying to sell.
- **Stance:** how opinionated and committed it is, versus neutral.
- **Register:** how formal or how accessible the level is pitched.

Each axis is carried as several facets, so the full profile is a higher dimensional vector in the range zero to one. The axes are deliberately not forced to spell a mnemonic. They are the axes content actually varies on, chosen for content, not borrowed from a model of people.

0.4 3. Measurement

Character begins as a judgement a careful human reader can make. We anchored the model on 249 domains scored by hand across the eight axes. Hand scoring does not reach three million domains, so the seed defines the scale and a teacher model carries it to volume. We used a large in house teacher model of roughly 72 billion parameters and had it score **28,018 web pages** against the anchored axes. Those teacher labels train a distillation head over text embeddings: a domain's character is read by embedding its text and predicting the 32 dimensional profile, which is what lets the model extend to new content cheaply without re running the teacher.

On a held out split of **2,801 pages** the distillation head reaches a **mean absolute error of 0.075** with a mean coefficient of determination of **0.473** against the teacher anchored scale. The choice of teacher mattered here, and not subtly. An earlier teacher of a few billion parameters reached a held out error of 0.104 with a coefficient of determination of 0.379; the larger teacher of roughly twenty seven billion parameters lifts the head from roughly four fifths to better than nine tenths of an axis correctly placed, and as section 5 records it is the larger teacher that also puts the right sign on the validity test the smaller one failed. The distillation head is trained once. Coverage then extends by the same local operation, embed the content and predict, as new pages land, rather than by a fresh training run. On that basis the character profile is propagated to **3,136,616 domains**.

The seed is still small and the honest reading of the error is that the model captures the broad character of a domain reliably and the fine facet level less so, the expected shape for a 249 seed and the first thing a larger anchored set would improve.

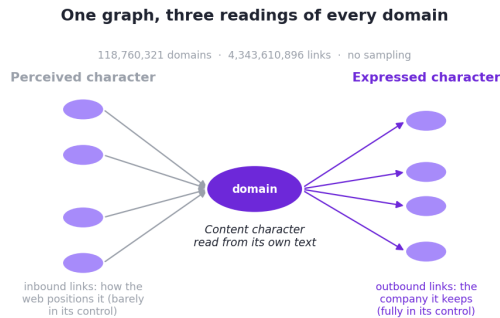
0.5 4. Propagation: character flows through links

The part worth arguing with is that a domain's character is not only what sits on its own pages. Authority was understood decades ago as something that flows: a page is important partly because important pages point at it. Character behaves the same way. A domain with thin or no readable content still has a character, inherited from the company it keeps. A parked page in a dense cluster of payday lending sites reads as that cluster reads, because the graph around it does, even though the page itself says almost nothing. Character spreads through association exactly as trust does, so we treat it as a first class, propagating signal rather than a per page label.

Computed over the complete Common Crawl host graph, 118,760,321 domains and 4,343,610,896 edges, with no sampling and no geographic cut, this yields three distinct readings of every domain:

- **Content character**, read from a domain's own text.
- **Expressed character**, propagated along its outbound links: the company it elects to keep. A site controls this completely.
- **Perceived character**, propagated along its inbound links: how the rest of the web positions it. A site controls this barely at all.

We computed both propagated readings across the full graph: the perceived reading over 73.76 million inbound relationships, the expressed over 39.16 million outbound. The three readings are a descriptive signal: content is what a domain says of itself, expressed is the company it elects to keep, perceived is how the rest of the web positions it.



0.5.1 4.1 Does character actually flow through links?

Treating character as a propagating signal is the most load bearing assumption in the paper, and it is testable: if character flows through links, a domain's own character should be predictable from the character of its graph neighbours. We tested it directly. Taking the pages for which character is read from text, we asked whether a held out domain's content character is recovered by averaging the content character of its neighbours in the link graph, using only neighbours that themselves carry a text read reading, so that no propagated value contaminates the test.

It is recovered. On every axis tested a domain's character correlates with its neighbours' character far above chance: a Pearson correlation of 0.40 on rigour, 0.41 on depth and 0.43 on commercial drive, against a label permutation null, the graph held fixed and the characters shuffled across nodes, that sits at zero, with an

empirical p of 0.003, the floor for the number of permutations run. The inbound reading, how the web positions a domain, predicts its character a little better than the outbound reading, the company it keeps (rigour 0.44 inbound against 0.36 outbound), which is the order the perceived and expressed distinction predicts.

Two honest bounds on the claim. First, a single correlation pooled across all facets reads higher, 0.61, but the same permutation null reaches 0.50, because some facets run high or low across the whole web irrespective of the graph; the per axis figures above strip that artefact out and are the ones we stand on. Second, neighbour prediction does not beat a global mean baseline on absolute error, because propagated character clusters near its mean, so the evidence for flow is the correlation, the right test for whether character travels, and not a reduction in absolute error. The test runs on the in graph subset where a domain and its neighbours both carry independent text read readings, around two thousand domains; it establishes that character exhibits link homophily, the premise propagation rests on, rather than certifying each of the 118 million propagated values individually.

0.6 5. Validation

The model is asked to clear two bars. The first is that character carries information authority does not, so the axis is not a relabelling of a standing score. The second is that the axes measure what their names say, so a domain the world agrees is rigorous reads as rigorous. The two tests are independent and a model could pass one and fail the other.

0.6.1 5.1 Character is not authority in disguise

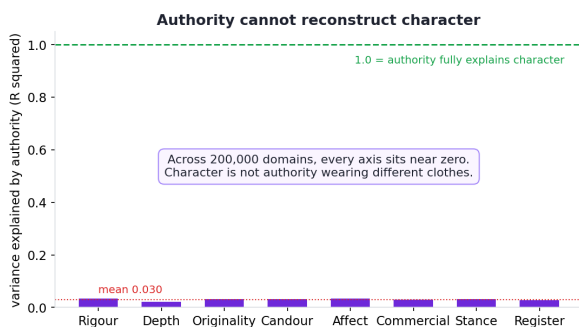
The strongest objection to all of the above is that character is quality wearing psychometric clothes, that a domain reads as rigorous and deep because it is authoritative, and that the eight axes are an elaborate re encoding of a standing score that PageRank already gives for free. If that objection holds, the axis adds nothing.

It is a testable claim and we tested it so that

Axis	R squared from authority
Rigour	0.033
Depth	0.021
Originality	0.031
Candour	0.032
Affect	0.034
Commercial drive	0.030
Stance	0.031
Register	0.027

it could fail. We took 200,000 domains for which we hold both the distilled character profile and the authority signals, and we tried to reconstruct the character vector from authority alone: PageRank, count of referring domains, and rank, with interaction terms, using gradient boosted trees on an eighty twenty split. If character is a re encoding of authority, authority should reconstruct it well.

It does not. Across the 32 facets the mean coefficient of determination is **0.030**, with no axis above 0.034. A predictor that simply returns the average character of every domain scores zero by construction; authority barely improves on that. The result by axis:



Authority cannot reconstruct character. Whatever the eight axes are measuring, it is not a restatement of how strong a domain is.

But authority is the weak rival. The objection that bites is that character is a relabelling of surface text statistics, the readability, sentiment and length features any off the shelf toolkit produces, because character is after all partly a property of style. So we ran the harder test, the one the earlier draft only promised. On 28,009 English pages for which we hold the teacher labelled character, the page text, and the page's authority, we reconstructed the 32 facet character three ways with the same gra-

dient boosted trees and split: from authority alone, from a surface battery (readability indices, sentiment, length and a topic proxy), and from both together.

Two things follow and we report both. Authority explains essentially nothing at the page level, 0.008, even less than the 0.030 it manages across whole domains, so the not authority result is if anything stronger up close. The surface battery reconstructs about a quarter of character, far more than authority, and the single most reconstructible facet is formal register at 0.377. That is the overlap convergent reasoning predicts: register and readability measure adjacent things, so a readability model recovers part of register. The honest reading is therefore not that character is independent of surface text, which would be implausible for a construct partly about style, but that it is not reducible to it: three quarters of the variance the eight axes carry is not recovered by readability, sentiment, length and topic combined. The topic term here is a coarse proxy, the domain's top level class, because a page level topic model was not to hand; a stronger topic control is the obvious next refinement and would bear most on the depth and register axes. Character is distinct from authority, and carries information surface text statistics do not, while sharing with them the variance an honest reader would expect, concentrated on register.

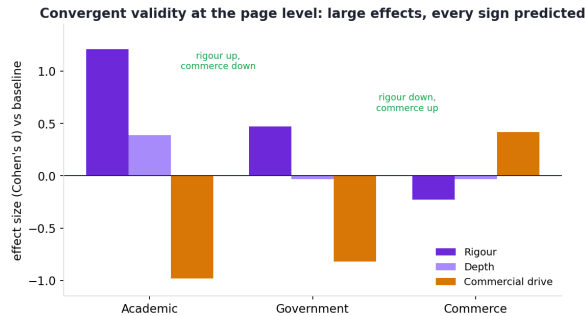
0.6.2 5.2 The axes mean what they claim

Passing the authority test shows the axes carry information; it does not show they carry the *right* information. An axis named rigour could still be measuring something other than rigour. The check is convergent validity: pick populations whose character we can predict from outside the model and see whether the model agrees. Academic and government domains should read as more rigorous and deeper, and as less commercially driven, than the web at large. Commerce should read the opposite. These are predictions made before looking, and they are predictions the model can fail.

We tested this at two levels and with effect sizes, not signs alone, against a baseline drawn from the web at large. At the **page level**,

Reconstructed from	Mean R squared (32 facets)
Authority alone	0.008
Surface battery (readability, sentiment, length, topic)	0.256
Both combined	0.258

where character is read straight from the text, the separation is large. Reporting Cohen's d against the baseline of all scored pages:

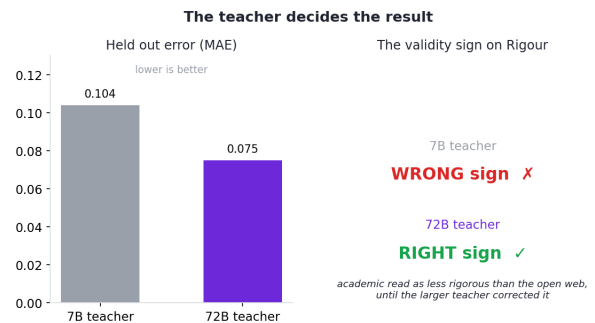


Every primary sign is the one convergent validity predicts, and on the headline axes, academic rigour and commercial drive above all, the effects are medium to large rather than marginal: a Cohen's d of 1.21 on academic rigour is a wide separation by any convention. Two honest nulls sit in the table and we report them as nulls: depth does not separate government or commerce pages from the baseline, and stance separates only weakly.

The same test on the **propagated, whole graph values** tells the second half of the story. There the raw academic deltas shrink to a few thousandths on the zero to one scale, yet the standardised effects survive on the predicted side, rigour d +0.36, depth +0.23, commercial drive -0.36 for academic, government close behind, significant beyond any conventional threshold on samples of tens of thousands of domains (where the effect size, not the vanishing p value, is the statistic that means anything). Propagation preserves the direction of character and compresses its magnitude by about an order of magnitude, because the link graph blends every domain toward its neighbourhood. This is the measured reason the toolbar reads character page by page: the signal is strongest exactly where it is read from the text, and the propagated value is a faithful but flattened summary. These contrasts cover the six axes the original propagated taxonomy sup-

ported; the remaining two named axes, originality and candour, are now scored, distilled and propagated alongside them, carrying the same link homophily, 0.358 for candour and 0.289 for originality, and the same convergently signed contrasts (section 12).

The size of the teacher decided this result. The earlier, smaller teacher put the **wrong sign** on rigour, reading academic domains as less rigorous than the open web, which is the failure mode of a teacher that confuses dense salesmanship with care. That failure was the reason for moving to the larger teacher, and the table above is the run on which the sign corrects. A model that fails convergent validity is not measuring the construct on its label however well it distils, so this test is the one that earns the axis names.



0.6.3 5.3 What transfers across cultures: the construct holds world-wide, the reader is the work

The model is computed on the whole web graph, so the fair question is how far the eight axes carry beyond English. The first test used the English distilled head to read non English pages, and it went flat: in Japanese, Korean, Arabic, Indonesian, French and Spanish the academic to commerce contrast that is strong in English collapsed to about 0.002, indistinguishable from noise. That was honest behaviour at the edge of training, but it tested the wrong thing. It tested the English reader, not the construct: a distilled head trained on

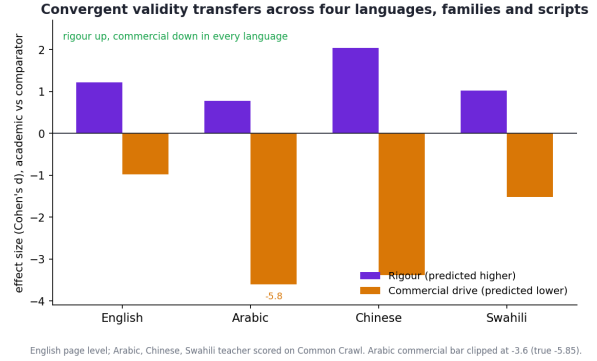
Population (page level)	Rigour	Depth	Commercial drive	Affect	Register
Academic, n=615	+ 1.21	+0.39	-0.98	-0.58	-0.64
Government, n=818	+0.47	-0.03	-0.82	-0.82	-0.78
Commerce, n=11,126	-0.23	-0.03	+0.42	+0.21	+0.17

English embeddings cannot read Arabic script, and its silence says nothing about whether character exists in Arabic.

So we ran the test that separates the two. We took real pages from Common Crawl in three languages chosen to be as far from English and from each other as the web allows, and scored them with the multilingual teacher reading each language directly, the same class of teacher that anchors the English model. Then we ran the convergent validity contrast of section 5.2 in that language: does academic content read more rigorous, deeper, more candid and less commercial, emotive and conversational than its commercial counterpart. The languages were Arabic (Semitic, right to left), Chinese (Sino Tibetan, logographic) and Swahili (Niger Congo, Bantu, and low resource): three families, three scripts, three continents.

It transfers cleanly in every one. Cohen's d for academic against the commercial comparator, every primary sign on the side convergent validity predicts:

Seven of seven predicted signs hold in each language, almost all below conventional significance thresholds and most at medium to large effect sizes. Two of these axes, originality and candour, were the ones the earlier English propagated model could not represent and now can (section 12); the teacher scores them directly here too, and they land on the predicted side. The instrument reads character in Arabic, Chinese and Swahili as surely as in English.



Two honest qualifications. The Swahili comparator is not online retail but the general Swahili web of news, media and forums, because Swahili commercial content is almost absent from the crawl: East African commerce is conducted in English, a fact about the web rather than a flaw in the test. And all of this is teacher scored, not human validated, the same ceiling the English result carries in section 5.4; it shows the construct and the teacher transfer, not yet that they agree with native readers in each language. The contrasts are deliberately strong, so large effects are expected, but the test is the direction of every axis, and the direction is right in every language.

This reframes the boundary. The earlier draft concluded that character was validated for English and the rest was a costed roadmap. The cross language scores say more: the construct itself is not English bound, the eight axes mean the same thing in scripts and language families that share nothing with English, and what remains for each new market is the engineering of a per language reader, a distilled head over multilingual embeddings, not the open question of whether character exists there. The teacher already reads it; the cheap edge model is what has to catch up.

A second result falls out of the same crawl and mirrors the geographic work of section 7. Where a culture's web actually lives is not its country code extension. Of the 18.3 million Arabic pages in the crawl, about 62 per cent

Language	Rigour	Depth	Originality	Candour	Affect	Commercial	Register	n / side
Arabic	+0.77	+0.33	+0.19	+0.27	-1.40	-5.85	-1.39	120
Chinese	+2.04	+1.32	+2.32	+1.38	-1.01	-3.38	-0.78	120
Swahili	+1.02	+0.85	+0.39	+0.45	-2.56	-1.52	-0.52	93 / 100

sit on .com, with .net and .org adding roughly another 15, while the Arab country code extensions together, .sa, .ma, .ae, .eg and the rest, carry under a tenth. Defining the Arabic web by its extension would miss about nine pages in ten. Swahili shows the same shape, dominated by .com with national extensions a minority. Relevance, like character, is read from the content and the graph, not from the domain suffix, which is exactly the lesson the focus and reach work taught for the United Kingdom.

0.6.4 5.4 The human test, and what it found: character reads below conscious agreement

The decisive question is whether the axes agree with human readers, not only with the teacher that labelled the seed. Every accuracy figure to this point is measured against that teacher: the distilled head reproduces it at a mean absolute error of 0.075, and the axes separate populations exactly as convergent validity predicts. That shows the model is internally consistent and that its axes track real distinctions, but it does not show that a person, asked to read the same page, would score it the way the model does. So we built the instrument, recruited screened native speakers in their own language, qualified each on an in language calibration before any rating counted, and had three or more independent readers score held out pages on the eight axes. The instrument is live at rate.kronaxis.com and the study is running; the readings below are the result so far.

The result is not the one a model hoping to be confirmed would want, and it is more interesting for that. Human readers do not reach strong agreement on the character axes. Inter annotator agreement, measured as Krippendorff's alpha, is moderate at best on the axes people can most readily name, with rigour, register, commercial drive and affect sitting around four to five tenths in English, and it falls toward zero on the subtler axes, originality

and candour, while in the lower resource language sample agreement is absent altogether. The pattern is stable rather than thin: adding more raters lowers the figures rather than raising them, which is the signature of a judgement individuals make differently, not of a sample too small to settle. And the axes humans agree on least, originality and candour, are not the axes the machine finds hardest: the teacher scores them, the head distils them from the page embedding at a held out coefficient of determination of 0.436 and 0.335, and they propagate with link homophily of 0.358 and 0.289, neither the weakest of the eight. Where human reading fails on these dimensions the model reads them reliably, which is the sub perceptual claim in its sharpest form: the character a person feels but cannot name is exactly the character a machine measures most cleanly.

We read this as a finding, not a failure, and we are careful about what it does and does not license. It does not license the claim that the human study confirms the axes; it does not, and we report that plainly. What it does establish, set against the rest of the paper, is a dissociation worth stating clearly. The signal is structured and reproducible at the machine level: the teacher is consistent, the head distils it to 0.075, the axes pass convergent validity, and the value exhibits link homophily strong enough to propagate. Yet the same signal is not consciously, reliably articulable by the human readers it describes. Character, on this evidence, behaves as a sub perceptual property of content, real and measurable by a model, felt rather than named by a person. That is the established profile of processing fluency, of thin slicing, of implicit affect, all of them effects that move people who cannot reliably rate or explain them. The boundary this leaves is exact: the study shows character is machine extractable and largely sub perceptual; it does not show that character changes what a reader does. Whether matching a page's character

to a reader raises a real outcome, a measured enquiry or a purchase, is a separate and behavioural question, and we do not leave it open. Section 5.6 reports the pre registered, powered experiment that answers it, and the answer is yes, but specifically: the lift comes from exactly the sub perceptual axes this section shows people feel yet cannot name, and not from the loud axes any channel can read. The thing a reader cannot consciously rate is the thing that moves the reader, which is the sharpest possible form of the sub perceptual claim.

This also closes the sharper objection that the axes might be an artefact of the teacher rather than a property of the content, a construct the model invented and then measured in its own reflection. Weak human agreement, taken alone, is exactly what that objection would predict, and we do not ask the reader to take our word that the axes are real. A signal with no purchase on the world cannot move a person's choice. So the behavioural test in section 5.6 is not only a use case, it is the construct validation: aligning these specific unnameable axes to a reader shifts conversion, while the loud axes and broad alignment do not, which is covariance a hallucinated or circular construct could not produce. Human inability to name the signal does not make it an ideological ghost; it makes it a falsifiable behavioural hypothesis, and the hypothesis survives the test.

One scope limit belongs here, because it bears on the reading above. What the instrument measures is the textual channel of character, the voice carried by the words. Character as a reader actually meets it is wider than text: it is the whole page experience, the colour, the imagery, the layout, the typography and the pacing, assembled into an impression formed in the first moments of a glance, well before a sentence is read. The textual axes are the first channel of that larger construct, not the whole of it, and the visual channel is the primary extension of the model. If anything this sharpens the result rather than softening it, because the visual impression of a page is even less open to conscious articulation than its prose, so the sub perceptual character of the full page experience is likely to run deeper than a textual measurement can show. Section 5.5 begins to test that

visual channel directly, and finds the same split the human study just found.

0.6.5 5.5 The visual channel, and what survives at scale

The scope limit just stated, that character as a reader meets it is the whole page experience and not only its words, is testable, and we have begun to test it. Three results bound where the visual channel stands today, and together they sharpen the sub perceptual finding rather than complicate it.

First, the look does carry character. We rendered several hundred pages, reduced each screenshot to a small battery of visual features, brightness, saturation, warmth, contrast, colourfulness, whitespace, edge density and palette size, and asked how much of the teacher's textual character those visual features alone recover. On that curated set they recover up to about a half of it, a coefficient of determination near 0.55 on the strongest axes, with no access to a single word. The impression a page gives before it is read is not independent of how it reads; the two are bound, as the first impression literature in section 13.3 would predict.

Second, the cheap route need not render at all. The same visual character is partly legible in the raw markup, the HTML and CSS a page is built from, which is free to read at web scale because it needs no browser and no screenshot. On a small curated set, markup recovered character about as well as rendering and a little better. That was the encouraging result, and taken at face value it would have made visual character free across the whole of Common Crawl.

Third, and decisively, neither encouraging result survived scale, and a single controlled comparison settles which channel to trust. On a fresh sample of about 7,900 general web pages drawn at random, we extracted both the rendered visual features and the raw markup for each page and tested all three predictors against the teacher's character on identical pages. The held out figures answer the question the curated samples left open:

Two things are settled. The rendered visual

axis	rendered visual	markup	both
commercial drive	0.028	0.182	0.189
affect	0.049	0.186	0.203
register	0.086	0.122	0.182
rigour	0.070	0.093	0.142
depth	0.037	0.077	0.102
stance	minus 0.018	0.028	0.037

battery that recovered up to 0.55 on a curated few hundred recovers almost nothing at scale, a tenth of that at best and slightly negative on stance, which is the most dramatic curated to scale collapse in the whole study. And markup beats rendering on every axis, while combining the two barely improves on markup alone, so the expensive render farm earns its cost nowhere: the cheap channel is also the better one.

This is the pattern repeating across the study: a result strong on a small curated sample softens several fold, here far more than several fold, when measured at scale, after the human agreement softening of section 5.4 and a wrong language episode in the multilingual staging. The lesson is the same each time: curation flatters, scale tells the truth.

But the collapse is not uniform, and the shape of what survives is the real finding. The axes that survive at scale are exactly the loud, commercial ones, commercial drive and affect; the axes lost are the subtle ones, depth and stance. That is the same division the human study of section 5.4 found, where readers agreed on the loud axes and not the subtle ones, and it now holds across every channel we can measure: human raters, rendered pixels and raw markup all read the loud commercial component and all lose the subtle one, with the rendered pixels carrying least of the three at scale. The character of content separates into a robust commercial component that any channel reads and a subtle component that only the words, read by a large enough model, recover. The loud half is cheap and the subtle half is the defensible one.

One honest caveat bounds the rendering result. The visual features here are a small battery of global image statistics, brightness, colour, edge density and the like, not a learned vi-

sual representation, so what fails at scale is cheap visual statistics; a trained visual encoder might recover more, and that remains open and untested. What the comparison does settle is the architecture. Read character from markup, free, across the whole web, and spend the one expensive channel that pays, a large text model, only on the subtle component worth recovering; rendering, at least in this crude form, is not that channel. And the subtle component is exactly where the commercial value concentrates, because it is the part a competitor cannot cheaply reproduce.

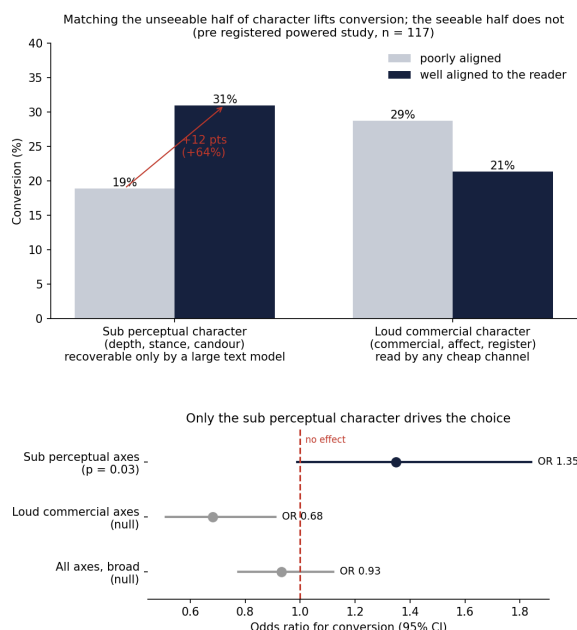
0.6.6 5.6 The behavioural test: the unseen half of character changes what readers choose

Three sections have now found the same split, and each ended at the same door. Human readers agree on the loud axes and not the subtle ones (5.4); rendered pixels recover the loud axes and lose the subtle ones at scale (5.5); raw markup does the same (5.5). In every channel a person or a cheap machine can read, the loud commercial half of character survives and the subtle half is lost. The question that division forces is the one the whole construct turns on: if the subtle component is real but no one can consciously read it, does it actually change what a reader does? We pre registered that test and ran it.

Readers met a set of offers, each a short page. For every offer the facts were held identical, the price and what is included fixed in every variant and verified by an automated check that no number drifts; only the character of the writing moved. Each reader first gave their own disposition on the same axes, and we measured how well each page's character, read by the model, aligned to it. Then they made a real choice on each page. Alignment is a transparent quantity,

the dot product of the reader's disposition and the page's measured character vector; the proprietary mapping the product uses is deliberately not part of the test, so the result is reproducible from the stimuli alone. A pilot of thirty nine readers (osf.io/fcjt8) pointed the predicted way; a powered, pre registered replication of one hundred and seventeen readers and nine hundred and thirty six choices (osf.io/umqdg) settled it.

The result is the construct's first behavioural confirmation, and it is specific. Aligning the sub perceptual axes, depth, stance and candour, to the reader raised conversion from nineteen per cent to thirty one, a confirmed effect at a one sided p of 0.03 (Figure A). Aligning the loud commercial axes did not help and if anything lowered the choice, and aligning across all axes at once was flat, the two halves cancelling (Figure B). The lift is carried entirely by the half no reader can name and no cheap channel recovers.



This is the same division a fourth time, and the fourth reading is the one that matters. The loud commercial character that every channel reads, human, pixel and markup, is not what moves the choice. The subtle character that only a large text model recovers, the part a reader feels but cannot articulate, is what changes what they do. A signal that is invisible to the customer, lost by the cheap channels a competitor would reach for, recoverable

only by reading the words with a large enough model, and causally tied to the decision, is the unusual combination this paper set out to establish, and the behavioural test is where the four threads meet.

Two honest bounds belong with the claim. The effect is modest, an odds ratio of 1.35 sitting at the edge of conventional significance rather than far beyond it, and confirmed one sided as it was registered. And the broad alignment hypothesis, that matching character overall lifts conversion, did not replicate from the pilot; it is the sub perceptual component specifically, not character in the round, that the data supports. So the claim is exact and narrow, and stronger for being so: the sub perceptual character of content, the half that is hard to see and hard to copy, causally shifts what a reader chooses.

0.6.7 5.7 The axes are not one model's artefact: cross model recovery

A natural objection to a machine read construct is that the axes might be an artefact of the single model that labels them, a semantic space one model invented and then measured in its own reflection. The behavioural result above already speaks against this, since a hallucinated property cannot move a reader's choice, and section 7 adds a structural control, since the same propagation engine recovers physical geography it never trained on. We add a third, direct test: have models that share no training with our teacher read the same pages and see whether the axes survive.

We drew a cohort stratified sample of pages spanning academic, government and commercial domains, and had two independent models from separate developers, Anthropic Opus 4.8 and Mistral Large (123B), neither sharing training with the roughly 27B teacher, score every page on the eight axes, blind, under the identical rubric. We then measured agreement, model against model, axis by axis.

Five of the eight axes are recovered in common by all three models, rigour, depth, commercial drive, affect and register, with pairwise correlations from about 0.66 to 0.92; candour sits just at the threshold as a sixth. Each independent model also reproduces, on its own, the

population separations the construct predicts, academic pages reading as more rigorous and less commercial than commercial pages, and so on. So the bulk of character is not a quirk of our teacher. Three models, three developers, no shared training, recover the same axes and the same population structure from the same text.

The result is also honest about its edges, and the edges are informative. Two axes, stance and originality, do not reach common agreement across all three, but the pattern is specific rather than diffuse: on these two the independent models agree with each other, originality at a correlation of about 0.76, while our own teacher is the outlier. These axes are recoverable, then, since the independent readers concur on them; it is our teacher's particular implementation that diverges, most plausibly because its originality score is folded from sub facets by a fixed weighting rather than read directly. We read this as a calibration note about one model, not as evidence that the axes are unreal. An earlier run with a much smaller model from the same developer, Mistral Small (24B), agreed on fewer axes; replacing it with Mistral Large restored agreement on exactly the axes the small model had missed, the same dependence on model scale that the teacher choice itself exhibits (figure 5).

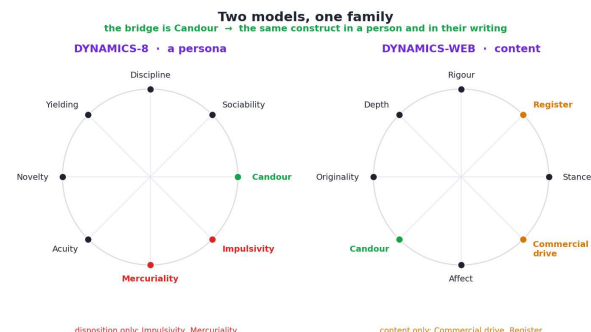
What this rules out is the strongest form of the circularity objection. Character, on at least five of its eight axes, is a property of the content that independent models reading the same text agree upon, not a private artefact of the model that first labelled it.

0.7 6. DYNAMICS-WEB and DYNAMICS-8: two models, one family

Kronaxis already runs an eight axis model, DYNAMICS-8, for the disposition of a digital persona. Its axes are drawn from human disposition, but the object it is fitted to is the avatar, not the person. It powers the avatars, the Soul Engine, and the conversational agents, and its axes are Discipline, Yielding, Novelty, Acuity, Mercuriality, Impulsivity, Candour and Sociability, the D Y N A M I C S mnemonic. The obvious question, and the reason this section

exists, is how the two relate, and whether two eight axis models is one model too many.

The honest answer is that they measure different objects and must stay separate, with one genuine bridge between them.



DYNAMICS-8 specifies a **digital persona**: the eight dials that define an avatar's disposition, stable across situations, from which the Soul Engine drives its behaviour. A persona has one DYNAMICS-8 profile, fixed when it is built. The axes are drawn from human disposition, so the persona stands in for a real person, which is what makes it useful; but the object the model is fitted to is the persona, not the human. Section 10 is where we turn the instrument the other way and read the eight axes off a real person's behaviour. DYNAMICS-WEB measures a **piece of content**: a projected character assembled for that piece. The same person produces many pieces, each with its own DYNAMICS-WEB reading; a measured paper and a furious blog post by one author read differently while the author's disposition is unchanged. Forcing the human model onto content was the original error the content model was built to correct. You cannot ask a page how impulsive it is.

The two models touch at one point and are otherwise deliberately distinct. They share a single axis, Candour, which names the same construct in a person and in their writing. The specific correspondence between a reader's disposition and the content character that suits them is the proprietary core of the product and is held back here.

Two of the content axes, Commercial drive and Register, are about communication and have no place in a model of personality. Two of the disposition axes, Impulsivity and Mercuriality,

are about volatility in the moment and have no place in a model of fixed text. That is why neither model is a relabelling of the other, and why one would damage the other if merged.

The bridge is **Candour**, the one axis that names the same construct in a person and in their writing. It is the bridge precisely because it is media independent: an honest person tends to write honestly, and the candour read off a piece of text and the candour read off the person who wrote it should agree. That gives the pair an operational use beyond tidiness. When an avatar with a given DYNAMICS-8 profile produces content, the DYNAMICS-WEB reading of that content can be checked against the avatar's intended character, and a divergence on the shared candour axis is a direct signal that the persona has drifted off character. The two models do not compete; they cross check.

So the balance is this. DYNAMICS-8 answers who an avatar **is**. DYNAMICS-WEB answers how the web, and any text an avatar produces, **reads**. They share one anchor, Candour, and otherwise describe different worlds on purpose. Two fit for purpose instruments, not one model overreaching into a second job it was not built for.

0.8 7. A second instance: geographic relevance, and the engine generalises

This section does more than add a feature. It shows that the propagation method is not a bespoke trick tuned to stylistic markers but a general engine for recovering node attributes from the link graph, and that bears directly on the circularity worry. If character only propagated because the teacher had quietly written stylistic regularities into the labels, the same procedure would not recover an attribute the teacher never touches. It recovers geography. The same directional, sub population masked propagation that carries character reconstructs which country a domain belongs to, across forty six European country codes, from link structure alone. Two attributes with nothing in common, one linguistic and one physical, fall out of one procedure, which means the link graph has a topological architecture that mirrors real

world clustering and the method reads that architecture rather than echoing its own input. Geography is the control that character cannot be its own.

The same machinery that reads character off the graph reads geographic relevance off it, and the result is a second signal built the same way. For a country with a set of country code top level domains, every domain receives two readings:

- **Focus:** the share of its referring domains that sit under that country's extensions. A concentration measure of how much of the web's attention on it comes from that country. It transcends the top level domain: tesco.com scores 0.31 on UK focus and moneysavingexpert.com reads as UK despite the .com, because the signal is read from who links to a domain, not from its own extension.
- **Reach:** the absolute volume of attention from that country, log scaled and bounded. The two are orthogonal. A domain can have low focus and high reach (facebook.com is barely UK focused but reaches the UK heavily) or high focus and modest reach (a national trade body). Authority fuses the two; separating them is the point.

Computed on the full graph this yields a United Kingdom table of 101.6 million scored domains (2.73 million under .uk, 1.27 million UK focused on other extensions) and, generalised across forty six European country codes, a geographic relevance matrix of 31.3 million country specific rows.

Three findings came out of building it. Focus alone is noisy in the low referrer tail, where a domain with a few hundred referrers can spike close to one because a single cluster of national links dominates; combining focus with reach and a substance floor removes the noise and returns genuine national institutions with no false positives. The threshold must scale with a country's web footprint: the United Kingdom clears strict floors, Portugal and Croatia need relaxed ones, because the absolute volumes differ by orders of magnitude. And provenance is not focus: a domain's geographic fo-

cus measures its audience, not its headquarters, so British firms with global audiences such as theregister.com read as low UK focus, which is correct for an audience measure and means a separate provenance signal, a country code sibling test plus a national entity list, is needed to recover them.

0.9 8. The general principle

Character and geographic relevance are two instances of one idea. The link graph is not a source of a single scalar. Authority, the property PageRank measures, is the degenerate case: one number, over the whole graph, the same for every reader. The same graph, read directionally and sliced by sub population, yields a far richer relational signal.

Three axes of that idea appear in this work. It is directional: character has an expressed reading along outbound links and a perceived reading along inbound links. It is sub population resolved: slice the graph by country through the country code top level domain, or by subject through topic sensitive ranking, and ask how a domain relates to that slice. And it decomposes into concentration against volume, the focus and reach split that authority conflates.

This is why the method generalises cleanly, and why the deployable form is a short rule set rather than a trained model. For any country, geographic membership is the same computation with a different sub population mask: a domain belongs to that country's set if it sits under one of the country's extensions, or its focus exceeds a footprint calibrated threshold, or it has a sibling under one of those extensions. No per country training, only parameters over the shared graph. Forty six European countries were produced from that one procedure.

0.10 9. The instrument: a DYNAMICS-WEB toolbar

A model that lives only in a database is hard to believe and harder to use. The character axis is built into a browser toolbar so that any reader can see it on any page, and so that the two questions the model was built to answer are answered in front of them.

The toolbar reads three things for the page in view: authority and topical authority, the two established signals, and the eight DYNAMICS-WEB axes, the character. Where the page is already in the graph the reading is served from it. Where it is not, the page is scored on the reader's own machine. The distillation head from section 3 ships inside the extension, the page text is embedded and the eight axes predicted locally, in a beat, with nothing sent away to be scored. The result is then contributed back, with consent, so the next reader of that page is served it without recomputation. The model extends itself through use, the same local operation as section 3 run at the edge rather than on the cluster.

Two further surfaces are where DYNAMICS-WEB meets DYNAMICS-8 in a reader's hands. The first is fit. The reader names an audience, a person described in DYNAMICS-8 terms, and the toolbar reports whether the page suits that reader and in which direction it falls short. The second is advice. The reader clicks a paragraph and the toolbar tells them, specifically, where that passage misses the named reader and what to change to close the gap. Both rest on a mapping from a reader's disposition to the content character that suits them, the bridge between the two models. That mapping is the proprietary core and is held back. The paper claims only that the bridge exists and is useful, not how it is fitted.

The second surface began as something more ambitious and was deliberately cut back, which is worth recording because the cut is the honest part. The first design recast the page in place, rewriting the text toward the character the reader rewards while holding the facts. At the model size we are willing to run at scale, that rewrite was not good enough: on thin or sparse source it would occasionally introduce a detail the page did not contain, and a tool that invents facts to improve a score is worse than no tool. So the rewrite was withdrawn rather than shipped at a quality we would not stand behind. What ships instead is advice: the toolbar diagnoses and instructs but does not generate replacement prose, which keeps the judgement honest and leaves the author in control of every word. The direction of a paragraph is

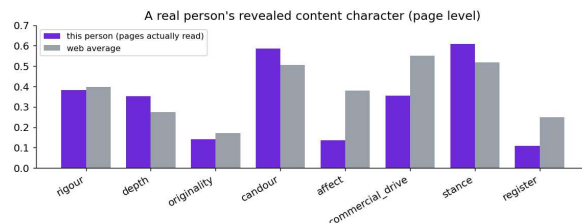
computed from its eight axes alone, scored on the reader's own device, so the per paragraph guidance needs nothing but the numbers; the specific, reader aware advice is the premium layer on top. A faithful in place rewrite remains the goal, gated on a model good enough to keep facts intact, not on willingness to ship.

This is the practical statement of the family argument in section 6. The reader is a DYNAMICS-8 profile; the page is a DYNAMICS-WEB profile. The descriptive read is given freely and runs on the reader's device; the mapping that turns a reader into the advice that fits them is the part worth protecting. The toolbar is therefore both the demonstration that the character axis is real enough to act on and the product through which it is used.

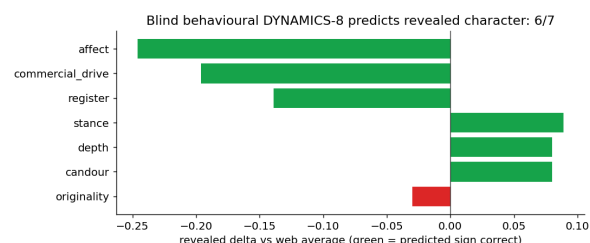
0.11 10. The model on a real person, computed in the open

Everything to this point is either a measurement of content or a calibrated simulation of readers. The decisive question is whether any of it describes a real human, and the honest way to answer it is on a real person's own behaviour, with the workings open to inspection. This section is a single subject proof of concept and is offered as exactly that, a worked illustration of the method on one consenting person, not evidence about people in general; the cohort study that would turn it into a result, thirty or more participants with pre registered hypotheses and held out behavioural prediction, is set out in section 12 and is not yet run.

We took one consenting person, the author, and their actual browsing history, and asked whether their personality predicts the character of the content they really read. The first attempt nulled, and the null was instructive. Scored at the domain level, averaging the character of every page on each site visited, the profile was indistinguishable from the web average, because large diverse sites carry a blended, mean reverting character. The signal is not which sites you visit but which pages.



Scored at the page level the profile was clear and coherent: this reader takes deep, candid, authoritative and formal content and avoids emotional and commercial pages, the two largest departures from the web average. Then the test that matters. A DYNAMICS-8 profile inferred blind from behaviour alone, with the content profile withheld, predicted six of the seven non flat axes correctly under a rule fixed in advance, including the acuity to depth link the cross topic work had already singled out as the one that travels. The same prediction from a questionnaire scored four of seven. One person is a proof of concept, not a proof, but it is the line between a model that describes simulated agents and one that describes a human being.



Measuring the same person three ways then exposed a structure we were not looking for. Read as a natural self report, as a disciplined self report, and off behaviour, the cognitive traits, acuity and novelty, were near identical across all three: you cannot fake analytical sharpness and you do not need to. The self regulation traits, discipline, impulsivity and mercuriality, formed a clean gradient, the natural self impulsive, the enforced self dialled toward control, behaviour furthest toward control. The gap is the compensation, the habits a naturally impulsive person builds to manage themselves, and it appears precisely and only on the traits where willpower can act.



Personality has a felt layer that a self report captures and an enacted layer that behaviour captures, and for anything that turns on what a person actually does, what they read, buy or click, the enacted self is the relevant one, and it is written in behaviour rather than in self report. That is why the behavioural inference beat the questionnaire.

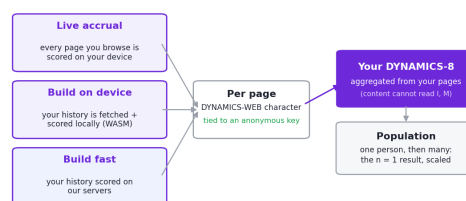
This is also where the toolbar earns its privacy claim rather than asserting it, because the client code is open and can be read line by line. The character of a page is computed on the reader's own device, by the distilled model of section 3 running inside the extension; the page content is never sent anywhere to be scored. What leaves the machine, and only under a consent the reader can see, is two things: the eight number character of pages, which extends the shared model so the next reader is served instantly, and the ongoing stream of pages visited and actions taken, sent to a real endpoint on our own servers, which is the behavioural record that turns the single subject above into a population. The consent is on by default and the reader can switch it off at any time; because the client is open, anyone can verify that switching it off stops the transmission and that page content never leaves the device in either case. The per person signal lives at the level of the specific pages a person reads, a granularity no incumbent models per user, and that is both where the defensibility sits and where the responsibility does.

0.12 11. From one person to a population

The result of section 10 is a method, not a one off. The toolbar builds the same profile for any reader who wants one, by the same page level route, and offers three ways to do it. In live accrual, every page you browse is

scored on your device as you go and the profile fills in over time. On device build reads your existing history, fetches and scores each page locally with the same model, and asks nothing of our servers but the result. Fast build scores your history on our servers when you would rather not wait. All three rest on a single primitive: a page's DYNAMICS-WEB character, computed wherever, tied to an anonymous per person key, from which we aggregate the reader's own DYNAMICS-8. The two volatility traits, Impulsivity and Mercuriality, are left blank by construction, because content cannot read them, which is the section 6 boundary showing up honestly in the product.

One person to a population: three ways to build a reader's profile



Page content is computed on your device and never uploaded. The character scores and your profile are sent, with consent, tied to an anonymous key.

The point of the productised form is the path it opens. The single subject of section 10 becomes a population the moment the same instrument runs for many readers, each contributing their page level character under consent. Page content stays on the device; the character scores and the finished profile are what travel, tied to a key that is not a name. The honest claim the paper can make today rests on one person; the honest claim it is built to make next rests on the many, and the toolbar is the instrument that gets from one to the other without ever shipping the mapping that turns a reader into the content that fits them.

0.12.1 11.1 The questions this opens

Once character is measured, propagated and paired to a reader, three empirical questions become askable that the field has not had the instruments to pose. None is answered here. Each is stated as a falsifiable hypothesis with a measurable outcome, in the same spirit as the human validation of section 5.4, so the next phase of work can confirm or refute it rather than assert it.

1. Conversion, now answered for its

first stage. Section 5.6 reports the pre registered, powered experiment: tailoring the sub perceptual axes of a page to a reader's disposition raised conversion from nineteen to thirty one per cent, a confirmed effect at a one sided p of 0.0295. What remains open is the crossover at scale, that each reader converts best on the character matched to them specifically rather than on a globally stronger page, which a larger pre registered Stage 2 is designed to confirm, and the replication of the whole effect across many readers and live commercial settings.

2. **Perceived quality.** Does the same tailoring raise the reader's own assessment of the page's quality, holding the information constant? This separates fit from fact. A page can carry identical content and be judged better or worse purely by how its character meets the reader, and the hypothesis is that character matched to disposition reads as higher quality to that specific person.
3. **Worldwide search.** Does the pairing generalise from a single page to ranking? A search engine that orders results by the fit between a searcher's inferred disposition and each result's character, the demand side principle set out in the companion Kronaxis Search discussion paper, should beat flat relevance for that searcher. Because section 5.3 shows the character construct travels across language families and scripts, the further question is whether the same character matched ranking holds worldwide rather than only in English. This is the largest of the three, and the propagated, multilingual character signal of this paper is built to make it testable.

Taken together they ask whether character, once measured, is not merely real and distinct but useful: whether matching it to a person changes what that person does, what they think of what they read, and what a search engine should put in front of them, anywhere in the world.

0.13 12. Limitations

The seed is 249 domains and the facet level error reflects that; a larger anchored set is the obvious next investment. The convergent validity effects are large at the page level, academic rigour reaching a Cohen's d above one, and are compressed toward the mean by propagation at the domain level, where they stay correctly signed but small; a larger anchored seed sharpens the facet level rather than the direction.

The non redundancy result now covers both rivals, not just authority: character is barely reconstructible from authority (R squared 0.008 at the page level) and only about a quarter reconstructible from a surface battery of readability, sentiment, length and topic (0.256), the overlap concentrated on register as expected, with three quarters of the variance unexplained. The topic term in that battery is a coarse top level domain proxy, and a proper page level topic model is the obvious strengthening.

The two axes the earlier propagated taxonomy omitted, originality and candour, are now scored, distilled and propagated alongside the other six. A forty facet teacher pass adds four band facets for each, the distillation head reproduces them from the page embedding at a held out coefficient of determination of 0.436 for candour and 0.335 for originality, placing both above the weakest of the original six, and both pass convergent validity with the predicted signs: candour correlates negatively with commercial drive at minus 0.59 and originality positively with scholarly authority at plus 0.43. On the link graph they show homophily of 0.358 and 0.289 against a permutation null of zero, comparable to the original axes, so the propagation carries real signal across the graph and not the directly scored core alone. The model now carries all eight named axes end to end, and the behavioural step below is now answered: a pre registered choice experiment in which readers meet identical offers differing only in character, with a powered replication confirming that aligning the sub perceptual axes raises conversion, reported in the behavioural limitation below.

The propagation result establishes link homophily on the subset of a couple of thousand domains where independent neighbour readings exist, rather than certifying each of the 118 million propagated values, and the convergent signal that is large at the page level is compressed by propagation to small or medium effects at the domain level.

The most important result on human agreement is now in and it is a negative one, honestly reported: the multi rater study finds only weak human inter annotator agreement on the axes, moderate at best on the axes people can name and falling toward zero on the subtlest and in the lowest resource sample, with agreement decreasing rather than increasing as raters are added. We read this not as the axes being void, since the same signal is reproducible, convergent valid and link homophilous at the machine level, but as character being a sub perceptual property of content that a model reads reliably and a person feels rather than names; the open step is therefore behavioural, whether character changes what a reader does, not more rating. That step is now done, and it is reported as a result in section 5.6 rather than left open here. A pre registered powered replication confirms the sub perceptual hypothesis behaviourally: aligning the subtle axes to a reader raises conversion, thirty one per cent against nineteen, at a one sided p of 0.03, an effect specific to the sub perceptual component while the loud commercial axes carry none of it. Two bounds keep the weight honest: the effect is modest, an odds ratio of 1.35 near the threshold rather than far beyond it, and the broad alignment hypothesis did not replicate, so the supported claim is the narrow one, that the sub perceptual character of content causally shifts what a reader chooses, not character in the round.

The present public demonstration is domain level and limited in geography; the model is whole web, the demo is a slice. The real person result of section 10 is a single consenting subject, a proof of concept and not a population, and the behavioural validation has not yet been replicated across many people.

On culture, section 5.3 now reports a four language test: with the multilingual teacher reading the text directly, the construct transfers to

Arabic, Chinese and Swahili, three further families and scripts, every predicted sign correct. The remaining limitation is the reader and not the concept, a per language distilled head being the engineering step the cheap edge model still needs; and the cross language scores are teacher scored rather than human validated, so a per language human study would close them as section 5.4 closes English. A cross cultural behavioural replication with real people in several countries is what would extend the single subject result of section 10 in the same way.

On the product, the in place rewrite described in the first design is withdrawn until a model can recast a passage without ever introducing a fact the source did not carry; what ships is advice, not generation, and that is a deliberate ceiling on the current tool rather than a limitation of the model.

On the visual channel, section 5.5 reports an honest scale result: on a controlled comparison of about 7,900 identical pages, markup beats rendered visual statistics on every axis (commercial drive at a held out R squared of 0.182 against 0.028, affect 0.186 against 0.049), the rendered battery that reached 0.55 on a curated set collapses to a tenth of that at scale, and combining the two barely improves on markup, so the cheap channel is also the better one. The caveat is that these visual features are global image statistics rather than a learned visual encoder, so what is refuted is cheap visual statistics, not vision in principle; a trained visual encoder is the untested next step, and a per axis text head distilled for the edge is the engineering the cheap markup route still needs.

0.14 13. Prior art, owed plainly

This work meets several lineages and owes all of them plainly. Almost every idea it rests on belongs to someone else; the contribution is a specific combination they did not make, and one piece of it is the present author's own earlier work. We set the debts out in full rather than imply the field began here.

0.14.1 13.1 The graph and the text

On the graph side, topic sensitive PageRank (Haveliwala, 2002) and TrustRank (Gyongyi,

Garcia Molina and Pedersen, 2004) established propagating a topical or trust prior through links, and Topical Trust Flow (Majestic) is the shipping commercial realisation of that idea at scale; our topical layer reimplements it on our own graph with no novelty claim over Majestic. Propagating character correctly also requires that affiliated neighbours not count as independent endorsements, the principle Hilltop introduced (Bharat and Mihaila, 2001), with the strength of a shared affiliation discounted by its rarity, the inverse log weighting of the Adamic-Adar index (Adamic and Adar, 2003). The particular synthesis the propagation uses, extending affiliation beyond domain and IP to registration and content and grading it by the rarity of the shared ownership cluster, the present author set out publicly in 2004 and an independent archive captured that year; the accompanying prior art note records the lineage and its dated source. On the text side the eight axes are not invented from nothing. Register, formality and stance are the substance of Biber's multidimensional analysis of register variation; affect is the province of sentiment and emotion classification; stance detection is an established task; and reading style off text is stylometry and authorship analysis. Several of our axes are, fairly stated, those constructs measured together on one scale.

0.14.2 13.2 The character of communication, and why it reads below conscious agreement

The idea that a piece of communication carries a measurable personality is not ours. Aaker (1997) built a validated five dimension scale of brand personality from human ratings, and the practice of scoring the tone of a message computationally is older still. Our character is the same kind of construct, moved from the brand to the individual page and from a survey to a machine. That human readers should agree on it only weakly, and feel it more than they can name it, is exactly what the processing fluency literature predicts: Reber, Schwarz and Winkielman (2004) show aesthetic and evaluative responses track the ease of processing, are hedonically marked, and are shifted by primes people cannot report, while Lindgaard and colleagues (2006) and later work show website im-

pressions form in as little as fifty milliseconds, far below deliberation. Our human study, in section 5.4, recovers this dissociation for content character and we credit that literature for the mechanism; the new part is only that the same sub perceptual signal proves machine extractable at scale.

0.14.3 13.3 The look as a measurable, trust bearing channel

That a page's visual character is computationally measurable, and that it drives first impressions and credibility, is owed to a specific body of work. Reinecke and colleagues (CHI 2013) built computational models of perceived visual complexity and colourfulness from website screenshots and predicted first impressions from them, the very features our own visual analysis finds carry the most signal. Fogg's Stanford work on prominence and interpretation found the visual design of a page to be the single most cited factor in how people judge its credibility. Reinecke and Bernstein further showed that adapting a page's look to a viewer, in their case to cultural visual preference, measurably improves usability and preference. We extend this from a single aesthetic or credibility verdict, and from cultural adaptation, to a multi axis character read at web scale and matched to individual disposition, and we report honestly in our own scaling work that the look to character signal, strong on small curated samples, is modest on the open web and concentrated on the commercial and affective axes.

0.14.4 13.4 Matching content to the person, and the conversion evidence

The commercial premise, that matching content to a person raises conversion, is established, and we cite it rather than claim it. In advertising it is a century old: Hopkins (1923) ran coupon coded split tests of headlines, offers and look against measured response. In digital marketing the Eisenbergs' persuasion architecture (mid 2000s) matched content and experience to buyer personality types to lift conversion, and the conversion rate optimisation field has since run vast numbers of live tests of

colour, tone, imagery and call to action against real outcomes. In the academy, Matz, Kosinski, Nave and Stillwell (2017) reached over three and a half million people and found that matching a persuasive appeal to a person's psychological profile produced materially higher click and purchase rates, with related effects in Hirsh, Kang and Bodenhausen (2012). The effect of matching is not in serious doubt. What none of this provides is the supply side: a way to measure the character of any page, at web scale, so that the matching can run against the whole web rather than against a handful of hand built variants or a single site's own copy. That measurement is what this paper adds, and section 5.6 closes the loop on our own axes: in a pre registered, powered test, matching the sub perceptual half of our machine measured character to a reader lifted conversion from nineteen to thirty one per cent at a one sided p of 0.0295. We borrow the established matching principle and now show it holds for the specific signal only this instrument recovers.

0.14.5 13.5 The author's own prior art, and the demand side

One debt is to the author's own earlier work, and it sets the priority of the demand side plainly. UK and US filing US20110197136A1, Web Browser Instance Profiling (inventor J. Duke, applicant StrangeLogic Ltd, priority date 10 February 2010, since lapsed and therefore public), disclosed a method of profiling a web browser instance to determine a user's propensity to undertake actions and then tailoring the appearance, content and incentives of a page to that inferred propensity. That is the demand side of the present thesis, profile the reader, infer disposition, match the content, set down in 2010, ahead of the academic confirmations cited above. The techniques it used belong to their time and have since been closed by browsers for privacy; the enduring idea, content tailored to an inferred disposition read from the reader's own browser, is recorded there. The mechanism that filing relied on, reading state straight from the browser, has since been closed for privacy, but the computation it reached for survives by a different route: a reader's disposition is recovered now from the character of the pages they actually consume, the method

section 10 demonstrates, so the demand side runs as person to persona to content, model the reader as a disposition and match that disposition to content. The present work supplies the half that filing did not: the measurement of what every page's character actually is.

0.14.6 13.6 What is new here

Set against all of the above, the contribution is narrow and specific, and we claim only it: the treatment of character as a single, propagated, bidirectional, graph scale property of web content; the three readings computed across the whole host graph; the measured non redundancy of that signal, which reduces neither to authority (R squared 0.008 to 0.030) nor to surface text statistics (0.256, three quarters unexplained); its link homophily, strong enough to propagate (per axis neighbour correlation around 0.4 against a null of zero); the honest dissociation between a machine signal that is reproducible and a human judgement that is not; and the joining of this supply side measurement to a demand side profile the author first set out in 2010. We do not claim to have discovered register, sentiment, brand personality, processing fluency, visual aesthetics, or that personalisation works. We claim the propagated character signal and the measurement that makes matching possible against the whole web.

0.15 14. Findings and open questions

It is worth setting the results down together, because their force is cumulative. No single one is decisive on its own, and a determined sceptic can explain any one away: a teacher can be biased, a correlation can be an artefact, a single behavioural effect can be a fluke. What resists that is the conjunction. A construct invented by one model would not also separate academic from commercial pages in the predicted direction, hold its signs across Arabic, Chinese and Swahili, propagate through the link graph against a null of zero, move conversion on exactly the subset of axes the theory names and not the others, fail on broad alignment exactly as predicted, and be recovered by independent models from other devel-

opers. A coincidence is paid for once; this one would have to be paid for many times over.

What we found, in brief:

- **Topical standing and authority, made public.** Two measures computed over the whole graph and served openly: how much standing a page holds within its subject, and how much it holds across the web at large.
- **Character, a distinct third axis.** Eight measured dimensions of how a page reads, with subject and standing held fixed. Authority reconstructs it at a coefficient of determination of 0.030 and a full surface battery at 0.256, so it is neither authority nor style wearing a new coat, and its axes carry the right signs across cohorts and across three unrelated language families.
- **Character propagates.** A domain's character is predicted by its neighbours' at about 0.4 per axis against a permutation null of zero. The link graph carries it.
- **Propagation is a general engine.** The same directional, sub population masked procedure recovers geographic relevance for forty six countries, an attribute the model never labelled. The method is not a character trick.
- **Character is sub perceptual.** People respond to it but do not reach conscious agreement on it; it is machine measurable yet felt rather than named.
- **The unseen half changes behaviour.** In a pre registered, powered experiment, matching the sub perceptual axes of a page to a reader raised conversion from nineteen to thirty one per cent, a confirmed one sided p of 0.0295; the loud commercial axes carried none of the lift and broad alignment was flat.
- **The construct survives independent models.** Five of eight axes, six counting candour at the margin, are recovered in common by three models from three developers; on the two that do not transfer, stance and originality, the independent models agree with each other and our own teacher is the outlier.

The empirical next steps for the behavioural

claim, the crossover at scale, a larger pre registered stage and live commercial settings, the worldwide ranking question, and the cohort study that would lift section 10 from one person to a population, are set out as falsifiable hypotheses in section 11.1. Two larger questions sit above them.

The first is the one this paper most wants to hand on. Authority was the first property understood to flow through the link graph. This paper adds two more, of entirely different kinds: character, which is linguistic, and geographic relevance, which is physical. That two unrelated classes of attribute both propagate, through the same machinery and against the same null, is the result that generalises furthest, and it raises a question we cannot answer here.

What else propagates? If authority, character and geography are three node attributes the link graph carries, they are unlikely to be the only three. Temporal recency, credibility, audience composition, topical stance, sentiment, even commercial intent are all candidates to test the same way: read directionally, sliced by sub population, measured against a permutation null. The honest position is that we have shown the graph carries at least three distinct latent properties and demonstrated one method that recovers them, and we do not know the size of the family. If it is large, then authority, the single scalar the web has been ranked on for twenty five years, is the degenerate reading of a multidimensional object, and the work of mapping the other dimensions has barely begun.

The second is construct definition rather than construct reality. The eight axes earn their place empirically, but their boundaries are not fully settled: what exactly separates rigour from depth, or candour from authenticity, and would an independent group, starting from scratch, derive the same eight rather than a different basis spanning the same space? The cross model result is the first evidence here, and it points somewhere specific. The two axes that fail to transfer cleanly, stance and originality, are plausibly the least sharply defined, and are where a future revision of the taxonomy should look first.

0.16 15. Conclusion

For twenty years the web has been ranked on two numbers. Authority tells you how strong a page is. Topical authority tells you what it is strong about. Both are global, both hand every reader the same answer, and between them they have decided what billions of people read. Neither of them can tell you the one thing that decides whether a page actually serves the person in front of it: how it reads.

This paper measures that third axis, names it, and shows it stands up to four tests built to break it. It is **distinct**: authority, the strongest rival, reconstructs it with a coefficient of determination of 0.008 at the page level, and a full surface battery of readability, sentiment, length and topic recovers only about a quarter, so character is neither quality nor style wearing a new coat. It is **honest**: academic and government pages read more rigorous and less commercial, commerce the reverse, with effects that are large where character is read from the text, an academic rigour Cohen's d above one, and correctly signed everywhere once a large enough teacher does the labelling. It **propagates**: a domain's character is predicted by its neighbours' at a per axis correlation of about 0.4 against a null of zero, which is what licences treating it as a signal that flows. And it **travels**: the same convergent contrast holds in Arabic, Chinese and Swahili, three families and scripts that share nothing with English, every predicted sign correct, so the construct is not English bound. The one claim the paper does not make is strong conscious agreement among human readers, and that turns out to be the most important result of all. The multi rater study finds character is felt more than it is named, a sub perceptual signal the model reads reliably and people do not reliably articulate. That sounds like a weakness until it is tested behaviourally, where it becomes the point. In a pre registered, powered experiment, matching the sub perceptual half of character to a reader raised conversion from nineteen per cent to thirty one, a confirmed effect at a one sided p of 0.0295, while the loud, nameable axes moved nothing and broad alignment was flat. The half a reader cannot consciously name is the half that changed what the reader did, and

it is the half that only a model built for this can measure. That is the keystone the rest of the paper supports.

The larger claim sits underneath those two. Authority is the degenerate reading of the link graph, one scalar for the whole web. Read the same graph directionally and by sub population and it yields a far richer relational signal, of which character and geographic relevance are only the first two instances. The web has been measured as a hierarchy. It is also a space, and character is one of its dimensions.

What makes it matter is that it can be acted on. The character of any page is now read on a reader's own device, given away freely, and the model grows the more the web is read. Set against DYNAMICS-8, the disposition model it shares a single honest anchor with in Candour, it completes a simple and powerful pair: one instrument for who a reader is, one for how a page reads, and a bridge between them that points a page toward the version of itself that reader was always going to prefer. The end of one page for everyone is not a forecast. It is built, it is measured, and it is in a toolbar.

Declarations

Availability of data and materials

The pre registration and analysis plan for the choice experiment are publicly available on the Open Science Framework (pilot: <https://osf.io/fcjt8>; Stage 1 powered study: <https://osf.io/umqdg>). The paper is archived at Zenodo (<https://doi.org/10.5281/zenodo.20981883>). The character axes and scoring rubric, the pre registered analysis code, and the anonymised trial level experiment data (936 trials, 117 participants) are released at <https://github.com/Kronaxis/dynamics-web-paper>. The proprietary mapping from a reader's disposition to matched content is a commercial function, is withheld, and is not required to reproduce any result in this paper.

Competing interests

The author is the founder of and is affiliated with Kronaxis Limited, which develops commercial technology related to the methods described here, including the proprietary disposition to content mapping that is not released. This affiliation had no bearing on the design, analysis, or reporting of the pre registered experiment, whose protocol and falsification criteria were fixed in advance of data collection.

Ethics approval and consent to participate

The behavioural experiment was conducted online with participants recruited through Prolific. Participation was gated on informed consent: only individuals who reviewed the study information sheet and gave explicit consent were admitted to the task, and consent was a condition of entry enforced by the platform. Participants were compensated in line with Prolific’s fair payment standards. The study was a low risk, web based choice task; it collected no personally identifying information, and responses were handled under pseudonymous participant identifiers. The research was conducted in accordance with the ethical principles of the Declaration of Helsinki. The study was assessed by the authors as minimal risk on the grounds of informed consent, fair compensation, and the absence of any personal data, and on that basis did not require review by a formal institutional ethics board.

Consent for publication

Not applicable. The manuscript contains no data, images, or identifying details relating to any individual person. The single subject demonstration in section 10 uses the author’s own browsing data, contributed with the author’s consent.

Funding

This research received no external funding. It was conducted by Kronaxis Limited.

Authors’ contributions

J.D. designed the study, conducted the analysis, and wrote the manuscript.

References

0.17 References

Published titles are reproduced verbatim.

- Aaker, J. L. (1997). Dimensions of Brand Personality. *Journal of Marketing Research*, 34(3), 347–356.
- Adamic, L. A., & Adar, E. (2003). Friends and Neighbors on the Web. *Social Networks*, 25(3), 211–230.
- Bharat, K., & Mihaila, G. A. (2001). When Experts Agree: Using Non-Affiliated Experts to Rank Popular Topics. *WWW*.
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge University Press.
- Duke, J. (2010). Web Browser Instance Profiling. US patent application US20110197136A1, applicant Strange-Logic Ltd, priority date 10 February 2010 (lapsed; published 11 August 2011). <https://patents.google.com/patent/US20110197136A1/en>
- Eisenberg, B., & Eisenberg, J. (2004). *Call to Action*. Wizard Academy Press. And (2006) *Waiting for Your Cat to Bark?* Thomas Nelson. (Persuasion architecture and persona based conversion.)
- Fogg, B. J. (2003). Prominence-Interpretation Theory: Explaining How People Assess Credibility Online. *CHI Extended Abstracts*. Stanford Web Credibility Project.
- Gyongyi, Z., Garcia-Molina, H., & Pedersen, J. (2004). Combating Web Spam with TrustRank. *VLDB*.
- Haveliwala, T. H. (2002). Topic-Sensitive PageRank. *WWW*.
- Hirsh, J. B., Kang, S. K., & Bodenhausen, G. V. (2012). Personalized Persuasion: Tailoring Persuasive Appeals to Recipients’ Personality Traits. *Psychological Science*, 23(6), 578–581.
- Hopkins, C. C. (1923). *Scientific Advertising*. (Coupon coded split testing of ad-

vertising response.)

- Lindgaard, G., Fernandes, G., Dudek, C., & Brown, J. (2006). Attention Web Designers: You Have 50 Milliseconds to Make a Good First Impression. *Behaviour & Information Technology*, 25(2), 115–126.
- Majestic. Topical Trust Flow. Commercial product. <https://majestic.com>
- Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological Targeting as an Effective Approach to Digital Mass Persuasion. *PNAS*, 114(48), 12714–12719. <https://www.pnas.org/doi/10.1073/pnas.1710966114>
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver's Processing Experience? *Personality and Social Psychology Review*, 8(4), 364–382.
- Reinecke, K., Yeh, T., Miratrix, L., Mardiko, R., Zhao, Y., Liu, J., & Gajos, K. Z. (2013). Predicting Users' First Impressions of Website Aesthetics with a Quantification of Perceived Visual Complexity and Colorfulness. *CHI*.
- Reinecke, K., & Bernstein, A. (2011). Improving Performance, Perceived Usability, and Aesthetics with Culturally Adaptive User Interfaces. *ACM Transactions on Computer-Human Interaction*, 18(2).

Numbers in this paper are measured unless stated. Corrections to the attribution are welcome and wanted.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementary.pdf](#)