

Google Trends Analysis

Purpose: To show that visual characteristics are correlated with consumer pairwise search on Google

- Goal is to predict top 2 rivals for each focal make-model
 - Use functional characteristics alone
 - Use visual characteristics alone
 - Using disentangled visual characteristics
 - Using entangled visual characteristics (VAE)
 - Using pretrained ResNet50 model
 - Use functional + visual characteristics
- Show that prediction using any kind of visual characteristics or functional+visual is better than functional characteristics alone
 - Especially for redesigned and crossover cars
- Question: How is this better than Krista Li / Peter Boatwright / ResNet50?

Purpose: How is consumer choice in lab explained by visual and functional characteristics of focal product and competitors?

- Leverage generative ability to come up with ~10 car designs within a car segment
- Randomize prices to be within 30-40K (randomizing prices helps in quantifying the value of visual characteristics)
- 1st Choice Task (Choose 1 out of 10 alternatives whose image, functional characteristics is shown)
- 2nd Choice Task (Repeat the above task after the first choice is taken out)
- Obtain ground truth on substitution matrix
- Is the functional distance matrix or the visual distance matrix or functional+visual distance matrix more correlated with the ground truth substitution matrix?
- Do ablation study to understand which visual aspect is more helpful in obtaining a higher correlation score
- Alternative: Replicate Ilya's paper by estimating demand with 1st choices with a model with disentangled visual char, entangled visual char, pretrained embeddings? CHALLENGE: Cannibalizes 3rd paper

Question: Is this a good simulation of the buying process of high-involvement category like cars?

Consideration Set Study

Purpose: Understand the potentially different roles played by functional and visual characteristics in the sequential dynamic consideration set formation

- Ask respondents to choose 3 out of 20 cars (track the sequence of choosing the cars or ask them to rank them)
- Show the picture of the car, functional features, price?
- Ranking might allow it to obtain data for the choice experiment as well
- Question: Do consumers first consider visual aspects and then narrow things down to functional or vice-versa? Is the variance of the visual char smaller or larger than the functional char within the 3 cars selected?
- Application - Advertising (what aspect of the car should you emphasize basis what stage of the purchase funnel are you at?)
- Others might have studied this but disentanglement will tell you which dimension they care about (grille or boxiness)
- Ask demographics and apply crosstabs to figure out what type of consumers first consider visual and then consider functional (and vice versa)

Developing Big Picture Rationale for chosen Y variables (in blue in earlier slides)

- Reviewers can argue that we should only care about consumer choices, not the details of the process
- So, we need to provide rationale for *WHY* the Y variable matters. Suppose we have data on final consumer choice in the field. Then, is there any advantage of knowing something about 1) pairwise search on Google 2) choice in lab 3) consideration set formation?
 - Need to argue that some intervention would be different based on this knowledge. Would be helpful to specify exactly what intervention, and how it would depend on this knowledge.
-
-

Results

1. **Results from consumer data / Evidence that consumers care about visual**
 - a. (Google Trends) Pairwise search as a function of visual and functional similarity.
 - Measured by raw score and top rival for a product
 - b. (Lab) Choice as function of product characteristics of all products in market
 - c. (Lab) Consideration Set as function of visual and functional similarity
2. **Results from Product data / What insights can we get from visual data?**
 - a. (Validating value of visual) Segment = function (visual, functional) Remove segment M. Binary classification.
 - b. Does form follow function? Some aspects are correlated.
 - c. Can visual explain differentiation across segments?
 - ~~d. How do brands differ in visual signature?~~
 - e. Which product is closest to mine? (Positioning)
 - f. Category crossing with Crossovers: which characteristics did the hybrid match with cars?
Which characteristics with SUVs?
3. **Comparison with Existing Methods / Show why disentanglement is better than alternatives**
 - a. Boatwright and Li papers
 - b. ResNet and other pretrained embeddings. Show how including color messes things up. Red Ferrari / Black Ferrari / Red Toyota

Aug 19, 2025

- Idea (1): no generative aspect; Idea (2,3): generative aspect
- Stated Preference (2,3); Revealed Preference (maybe 1?)
- Followup with Yougov
- Show the value of disentanglement by comparing the disentangled embeddings and pretrained embeddings of cars of same color v diff color
(note: diff color cars should be similar or exactly same)
- Did the vendor get back with a list of user forums?

2025-08-20 Notes

What would we want to show?

Currently, we have some insights from disentanglement, e.g. body shape or grille width. However, there is no way to validate that consumers actually care about these visual characteristics.

So, idea is to get user forum data to see if consumers are talking about these aspects that are picked up by our disentanglement method.

Idea: People talk more about distinct things, rather than the ordinary.

Can we map user discussions about visual design aspects to the visual characteristics discovered by disentanglement? Can use this as a dependent variable. Same independent variables (difference between focal car and competing cars in various characteristics).

Volume of Conversations about V1 (@ mmy) = $f(\text{how distinct is this mmy relative to others ON this characteristic? } V1)$. 8 dependent variables 4 functional and 4 visual.

Ideally, we want:

- High overlap is set of models between Edmunds and UK Car market.
- More redesigns

Identify a segment (maybe segment D) that best satisfies these conditions. Then collect data from user forum for cars in this segment.

<https://pubsonline.informs.org/doi/abs/10.1287/mksc.1110.0633?journalCode=mksc>

<https://www.overleaf.com/project/67f688c9f5f2ea881c0a2861>

2025-08-20 Notes

Impact of Redesign on User Conversations (Part 2)

Can do a difference-in-differences (before and after redesign).

Hypotheses (can we confirm or disconfirm?):

- 1) When there is a visual redesign, users talk more about visual stuff
- 2) Users talk more about visual characteristics that are identified as having changed more by the disentanglement method.

2025-08-22 – Connecting Disentanglement Insights with Consumer Choice

Ankit, I've written the idea below to get professional reviews from newspapers, auto magazines, websites etc. for car models each year.

Can you detail the implementation after discussing with the data vendor?

We can then check what aspects they are talking more about functional or visual. And within visual, if what they are talking about maps with our finding of the 4 visual characteristics. Can also do the same for functional.

This will connect the insights from disentanglement to something that would impact consumers (reviews).

2025-08-22 – Connecting Disentanglement Insights with Consumer Choice

I looked at several sites - AutoCar, AutoExpress, WhatCar, CarBuyer

Example: <https://www.autocar.co.uk/car-review/audi/a4-2008-2014>

Typically, each review has clear subsections: Introduction, Design & Styling, Interior, Engine & Performance, Ride & Handling, MPG & Running Costs, Verdict

So, very broadly, according to them these are the attributes for which consumers have preferences over (obv preferences are heterogenous and not the same for everyone)

Question Is: How do I connect expert reviews of 'Design & Styling' to disentangled visual characteristics?

V: My sense is there will be some trial and error. Can even ask an LLM, and ask it to show reasoning why the connection. So long as it is done in a neutral way, without cherry picking, it should be fine.

2025-08-22 – Connecting Disentanglement Insights with Consumer Choice

Example: Audi A4 2008-14 (**ChatGPT says the text in bold refers to boxiness**)

Virtually every component in the Audi A4 is new apart from the engines, although passers-by might struggle to notice this given the generic nature of mainstream Audis' styling. If you look solely at its **silhouette and proportion, you can trace this car's look directly to the 'aerodynamic' Audi 80 of 1986**. Such evolutionary styling, and a mechanical confection whose advances in large part lie on the option list, leave this Audi feeling slightly less novel than an all-new platform implies it might. That said, it seems many buyers like the Audi look of one basic design in different sizes. It also has the back-handed advantage that an Audi A4 won't date too quickly. The A4 seems to get more attractive with each generation, as the current generation testifies. It has a **shorter front overhang** than before, allowing Audi to extend the wheelbase without significant impact on the overall dimensions. The positioning of the differential and the clutch (or torque converter) has been swapped around, allowing the front axle to be moved forwards so the engine can sit further behind the front wheels.

2025-08-22 – Connecting Disentanglement Insights with Consumer Choice

Example: BMW 1 Series 2011-15 (**ChatGPT says the text in bold refers to boxiness, and the highlighted text may be referring to change in grille proportions**)

The BMW 1 Series is a car that looks as if it has been steered, albeit carefully, in the right direction. BMW talks less about dynamic rear-driven handling, and at reassuring length about enhanced comfort and accommodation, improved efficiency and new-to-the-class technology. Which is a good start. Longer than the old 1 Series by 85mm, the current car has a wheelbase that has been enlarged by 30mm, with 21mm of that extra inter-axle length gone to additional rear legroom. Both tracks have been widened, too, by a gnat's over 40mm at the front axle and over 60mm at the rear. Although it has grown, the new car is 30kg lighter than the old one. It would have been 60kg, but climate control now comes as standard. And a thorough structural redesign means the car's body-in-white is now more than 30 percent more torsionally rigid across the front bulkhead. That's good news for ride and handling, too. Adrian van Hooydonk's styling update hasn't cured the ungainly proportions of BMW's smallest model, but the net effect is a clear improvement. The new car looks **lean and more aggressive** than the last. The biggest aesthetic bugbear remains the car's profile, though. **Short, tall and backward leaning**, it still looks **awkward: like a gangly pup** that has had the carpet pulled out from under its paws. Developed in tandem with the new BMW 3 Series, the new 1 Series, like the last, has all-independent suspension – MacPherson struts up front and a five-link rear end. The car is sold exclusively with turbocharged engines, including big-selling 1.6-litre petrol and 2.0-litre common-rail diesel units. Recent additions to the range include the 125i, 125d diesel and M135i, a flagship that forms part of BMW's new M Performance range. Every 1 Series model can now be had with M Sport trim, offering welcome styling upgrades but with lowered suspension that does nothing for ride quality.

ADD PIC FROM DATASET



2025-08-22 – Connecting Disentanglement Insights with Consumer Choice

One approach is to see the language survey respondents used when they were shown disentanglement traversal, and see if it can be connected with expert mentioning certain aspects.

-> Text that corresponds to boxiness when humans saw the traversals -> We need to connect that text with professional reviews. Just ask about volume and valence of things that correspond to everything that humans have said.

-> How do professional reviews react to outliers v prototypical designs?

2025-08-24 – Data Vendor

Concerns:

For panel dataset, can the scraper tell us if he can find older reviews for other years. For each make-model-year, can we get a review. Maybe using Wayback M/C the vendor can find more.

2025-08-24 – Data Vendor

-> Volume: Functional; Visual / Aesthetics / Appearance; Overall

-> Valence: Functional; Visual; Overall

We will be classify proportion of review -> functional ; visual ; etc.

-> Temperature | Version # -> It should be done as a script (as an API call).

-> Gather the review response, query ChatGPT, and then generate volume/valence of overall, functional, visual, v1-v4 separately.

Provide v1-v4.csv and list of make-model-years.

Ask him to come up with independent sources (plus search myself) -> Give a union then.

2025-08-25 – email for data vendor

Hi Ajay,

I hope you are doing well, I need your help with the following assignment, and would appreciate your inputs on whether your team can execute it. If it is possible, then you can share an estimate of the time and money it will cost. This is time sensitive.

1. Step 1: Find professional car review websites in the UK (some examples are AutoCar, AutoExpress, WhatCar, CarBuyer)
2. Step 2: Search for professional car reviews for each make-model-year present in the attached CSV. Note, in some websites, for some make-models you may only be able to see reviews for certain 'years'. For these make-models, you may have to rely on Wayback Machine or something similar to find professional reviews for all make-model-years. The goal is to have at least 10-20 professional reviews for each make-model-year.
3. Step 3: Using ChatGPT API, ask what is the volume of the overall review (# of sentences, words etc.), review related to the functional aspects of the car, review related to the visual aspects of the car. Also, ask the valence of the overall review (is it a +ve or -ve review), valence of the functional review (is it +ve or -ve on the functional aspects), valence of the visual / design / aesthetics aspect of the car (is it +ve or -ve).
4. Step 4: I will provide you four different text files - each called v1.txt, v2.txt, v3.txt, v4.txt - using ChatGPT API, ask what volume of the review corresponds to the text in each of the v1-v4.txt (you would obtain 4 different numbers, one each for v1 to v4). Similarly, ask what is the valence of the part of the review that corresponds to the text in each of the v1-v4.txt.

Note: Please write a script to use ChatGPT API in which you clearly define the version and temperature etc.

Please ensure the code is replicable so that if it is run a year later, we will obtain the same result.

2025-09-06 – Prompt Engineering Learning

ChatGPT told me there are two ways to prompt engineer:

- seeded (supervised)
- blind (unsupervised)

It suggested that I should first build a lexicon for ~100 responses each for v1 to v4, and then put that shorter lexicon (will also help reduce number of tokens in API calls) in every prompt for a full text review. This is what LLM refers to as the supervised approach.

2025-09-06 – Prompt #1

Think silently and set Temperature =0.

In the below review:

1. Tell me the percentage of the review discussing Functional aspects of the car.
2. Tell me the percentage of the review discussing Visual (aesthetics or form) aspects of the car.
3. Tell me the percentage of the review that is Neither functional nor visual.
4. Of the Visual part, tell me whether they are directly, indirectly, or not at all discussing aspects related to:
 - a) body shape of the car (being hatchback-like or sedan-like)
 - b) grille width (being small or large)
 - c) grille height (being small or large)
 - d) whether the car is very boxy or very sleek.
5. For each aspect, also provide 2–4 short verbatim examples (≤25 words) from the review that support the coding.
6. Provide valence scores (-3 to +3) for Functional, Visual, Neither, and Overall.

Please output in JSON:

```
{  
  "functional_percentage": 0.00,  
  "visual_percentage": 0.00,  
  "neither_percentage": 0.00,  
  "valence": {  
    "overall": 0.00,  
    "functional": 0.00,  
    "visual": 0.00,  
    "neither": 0.00  
  },  
  "visual_aspects_breakdown": {  
    "body_shape": {  
      "present": true,  
      "evidence_type": "direct/indirect/not",  
      "examples": []  
    },  
    "grille_width": {  
      "present": true,  
      "evidence_type": "direct/indirect/not",  
      "examples": []  
    },  
    "grille_height": {  
      "present": false,  
      "evidence_type": "not",  
      "examples": []  
    },  
    "boxy_vs_sleek": {  
      "present": true,  
      "evidence_type": "direct/indirect/not",  
      "examples": []  
    }  
  },  
  "[REVIEW_TEXT]  
<<< paste review here >>>
```

2025-09-06 – Output #1

```
{  
    "functional_percentage": 66.3,  
    "visual_percentage": 17.4,  
    "neither_percentage": 16.3,  
    "valence": {  
        "overall": 1.4,  
        "functional": 1.5,  
        "visual": 1.3,  
        "neither": 1.0  
    },  
    "visual_aspects_breakdown": {  
        "body_shape": {  
            "present": true,  
            "evidence_type": "direct",  
            "examples": [  
                "A relatively compact 2+2 coupé has become a rarity.",  
                "That would make this 2 Series a BMW 1 Series coupé.",  
                "The BMW 4 Series is a BMW 3 Series coupé.",  
                "Even numbers are coupés or otherwise swoopier versions of the odd numbers."  
            ]  
        },  
        "grille_width": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "It has the typical BMW hallmarks – a long bonnet, rear-wheel drive, subtle kidney grilles.",  
                "The 2 Series Coupé has had a bit of work done to it lately. Cosmetically, this consists of a new front end.",  
                "The grilles get lower and narrower."  
            ]  
        },  
        "grille_height": {  
            "present": false,  
            "evidence_type": "not",  
            "examples": []  
        },  
        "boxy_vs_sleek": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "Looking for an everyday coupé that offers more style than the usual SUV...",  
                "Even numbers are coupés or otherwise swoopier versions of the odd numbers.",  
                "The BMW X6 is a BMW X5 with a sloping roofline.",  
                "The whole point of the car is the chassis balance afforded by the native rear-wheel drive."  
            ]  
        }  
    }  
}
```

2025-09-06 – Prompt #2 (SHORTER)

Think silently. Temperature=0.

From the review, output JSON with:

- % Functional, % Visual, % Neither (sum ≈100, word-based).

- Valence scores (-3 to +3) for Functional, Visual, Neither, Overall.

- For Visual: body_shape, grille_width, grille_height, boxy_vs_sleek → mark present true/false, evidence_type (direct/indirect/not), and 2–4 short verbatim examples (≤25 words).

Schema:

```
{  
    "functional_percentage": 0.00,  
    "visual_percentage": 0.00,  
    "neither_percentage": 0.00,  
    "valence": {  
        "overall": 0.00,  
        "functional": 0.00,  
        "visual": 0.00,  
        "neither": 0.00  
    },  
    "visual_aspects_breakdown": {  
        "body_shape": {"present": true, "evidence_type": "", "examples": []},  
        "grille_width": {"present": true, "evidence_type": "", "examples": []},  
        "grille_height": {"present": false, "evidence_type": "", "examples": []},  
        "boxy_vs_sleek": {"present": true, "evidence_type": "", "examples": []}  
    },  
    [REVIEW_TEXT]  
}
```

2025-09-06 – Output #2

```
{  
    "functional_percentage": 66.87,  
    "visual_percentage": 19.52,  
    "neither_percentage": 13.61,  
    "valence": {  
        "overall": 1.22,  
        "functional": 1.30,  
        "visual": 1.10,  
        "neither": 0.50  
    },  
    "visual_aspects_breakdown": {  
        "body_shape": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "a relatively compact 2+2 coupé has become a rarity",  
                "it rides on a short version of the CLAR platform",  
                "every 2 Series Coupé gets the excellent eight-speed ZF automatic gearbox",  
                "the car's low centre of gravity and the option of an ultra-low seating position"  
            ]  
        },  
        "grille_width": {  
            "present": true,  
            "evidence_type": "direct",  
            "examples": [  
                "subtle kidney grilles",  
                "BMW hallmarks – a long bonnet, rear-wheel drive, subtle kidney grilles",  
                "a new front end",  
                "updated wheel and paint options"  
            ]  
        },  
        "grille_height": {  
            "present": false,  
            "evidence_type": "not",  
            "examples": []  
        },  
        "boxy_vs_sleek": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "even numbers are coupés or otherwise swoopier versions of the odd numbers",  
                "the BMW X6 is a BMW X5 with a sloping roofline",  
                "offers more style than the usual SUV",  
                "more driving engagement than a front-wheel-drive hot hatch"  
            ]  
        }  
    }  
}
```

2025-09-06 – Prompt #3 (SHORTER WITH VALENCE)

Think silently. Temperature=0.

From the review, output JSON with:

- % Functional, % Visual, % Neither (sum ≈100, word-based).

- Valence scores (-3 to +3) for Functional, Visual, Neither, Overall.

- For Visual: body_shape, grille_width, grille_height, boxy_vs_sleek → mark present true/false, evidence_type (direct/indirect/not), 2–4 short verbatim examples (≤25 words), and valence (-3 to +3).

Schema:

```
{  
    "functional_percentage": 0.00,  
    "visual_percentage": 0.00,  
    "neither_percentage": 0.00,  
    "valence": {  
        "overall": 0.00,  
        "functional": 0.00,  
        "visual": 0.00,  
        "neither": 0.00  
    },  
    "visual_aspects_breakdown": {  
        "body_shape": {  
            "present": true,  
            "evidence_type": "",  
            "examples": [],  
            "valence": 0.00  
        },  
        "grille_width": {  
            "present": true,  
            "evidence_type": "",  
            "examples": [],  
            "valence": 0.00  
        },  
        "grille_height": {  
            "present": false,  
            "evidence_type": "",  
            "examples": [],  
            "valence": 0.00  
        },  
        "boxy_vs_sleek": {  
            "present": true,  
            "evidence_type": "",  
            "examples": [],  
            "valence": 0.00  
        }  
    }  
}
```

[REVIEW_TEXT]

<<< paste review here >>>

<https://www.overleaf.com/project/67f688c9f5f2ea881c0a2861>

2025-09-06 – Output #3 BMW2SERIES

```
{  
    "functional_percentage": 63.5,  
    "visual_percentage": 23.0,  
    "neither_percentage": 13.5,  
    "valence": {  
        "overall": 1.2,  
        "functional": 1.5,  
        "visual": 0.9,  
        "neither": 0.0  
    },  
    "visual_aspects_breakdown": {  
        "body_shape": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "A relatively compact 2+2 coupé has become a rarity",  
                "This consists of a new front end, side skirts painted in the body colour",  
                "This 2 Series a BMW 1 Series coupé",  
                "BMW X6 is a BMW X5 with a sloping roofline"  
            ],  
            "valence": 1.2  
        },  
        "grille_width": {  
            "present": true,  
            "evidence_type": "direct",  
            "examples": [  
                "subtle kidney grilles",  
                "Cosmetically, this consists of a new front end",  
                "the grilles get lower and narrower"  
            ],  
            "valence": 0.7  
        },  
        "grille_height": {  
            "present": false,  
            "evidence_type": "not",  
            "examples": [],  
            "valence": 0.0  
        },  
        "boxy_vs_sleek": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "The BMW X6 is a BMW X5 with a sloping roofline",  
                "more style than the usual SUV",  
                "offers more driving engagement than a front-wheel-drive hot hatch",  
                "swoopier versions of the odd numbers"  
            ],  
            "valence": 1.0  
        }  
    }  
}
```

2025-09-06 – Output #3 FORDFOCUS

```
{  
    "functional_percentage": 71.30,  
    "visual_percentage": 16.85,  
    "neither_percentage": 11.85,  
    "valence": {  
        "overall": 1.70,  
        "functional": 2.10,  
        "visual": 0.80,  
        "neither": 0.30  
    },  
    "visual_aspects_breakdown": {  
        "body_shape": {  
            "present": true,  
            "evidence_type": "direct",  
            "examples": [  
                "the Focus comes only with five doors. There's no three-door hatch option",  
                "the estate's load bay is also 119mm wider than the hatchback's",  
                "the Focus Estate rides on the same wheelbase as the five door",  
                "rear seat cushions and folded the seat backs (which go almost, if not quite flat)"  
            ],  
            "valence": 1.20  
        },  
        "grille_width": {  
            "present": false,  
            "evidence_type": "not",  
            "examples": [],  
            "valence": 0.00  
        },  
        "grille_height": {  
            "present": false,  
            "evidence_type": "not",  
            "examples": [],  
            "valence": 0.00  
        },  
        "boxy_vs_sleek": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "the car looks like it's moving even when standing still",  
                "It has a sophisticated feel that lifts it above the impression one gains from... the Vauxhall Astra",  
                "we do mourn some of the loss of its old sparkle",  
                "it could quite easily come from the class above"  
            ],  
            "valence": 0.80  
        }  
    }  
}
```

2025-09-06 – Output #3 HONDACIVIC

```
{  
    "functional_percentage": 62.7,  
    "visual_percentage": 25.1,  
    "neither_percentage": 12.2,  
    "valence": {  
        "overall": 1.05,  
        "functional": 1.22,  
        "visual": 0.45,  
        "neither": 0.30  
    },  
    "visual_aspects_breakdown": {  
        "body_shape": {  
            "present": true,  
            "evidence_type": "direct",  
            "examples": [  
                "its profile lower, wider and more aerodynamic",  
                "the tenth generation Civic hopes its radical, if slightly more conventional look and bigger size will close the gap",  
                "naturally longer and wider than its predecessor, with a longer wheelbase and 20mm lower too"  
            ],  
            "valence": 0.8  
        },  
        "grille_width": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "the Civic's styling, such as its front wings, bodyside surfaces and rear bumper treatment",  
                "front bumper finish is made piano black",  
                "aggressive bodykit, adaptive dampers, low profile tyres"  
            ],  
            "valence": 0.2  
        },  
        "grille_height": {  
            "present": false,  
            "evidence_type": "not",  
            "examples": [],  
            "valence": 0.0  
        },  
        "boxy_vs_sleek": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "exterior styling is more athletic and elegant",  
                "profile lower, wider and more aerodynamic",  
                "design improvements have been made"  
            ],  
            "valence": 0.3  
        }  
    }  
}
```

2025-09-06 – Prompt #4 (WITH ONLY JSON)

Think silently. Temperature=0.

Fill this JSON with values derived from the review:

```
{  
    "functional_word_count": 0,  
    "visual_word_count": 0,  
    "other_word_count": 0,  
    "valence": {  
        "overall": 0.00,  
        "functional": 0.00,  
        "visual": 0.00,  
        "other": 0.00  
    },  
    "visual_aspects": {  
        "body_shape": {  
            "present": true,  
            "evidence_type": "direct/indirect/not",  
            "examples": [],  
            "valence": 0.00  
        },  
        "grille_width": {  
            "present": true,  
            "evidence_type": "direct/indirect/not",  
            "examples": [],  
            "valence": 0.00  
        },  
        "grille_height": {  
            "present": true,  
            "evidence_type": "direct/indirect/not",  
            "examples": [],  
            "valence": 0.00  
        },  
        "boxy_vs_sleek": {  
            "present": true,  
            "evidence_type": "direct/indirect/not",  
            "examples": [],  
            "valence": 0.00  
        }  
    }  
}
```

[REVIEW_TEXT]
<<< paste review here >>>

2025-09-06 – Output #4 BMW2SERIES

```
{  
    "functional_word_count": 1345,  
    "visual_word_count": 246,  
    "other_word_count": 597,  
    "valence": {  
        "overall": 1.35,  
        "functional": 1.40,  
        "visual": 1.10,  
        "other": 1.25  
    },  
    "visual_aspects": {  
        "body_shape": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "a relatively compact 2+2 coupé has become a rarity",  
                "the BMW X6 is a BMW X5 with a sloping roofline",  
                "that would make this 2 Series a BMW 1 Series coupé",  
                "a coupé that offers more style than the usual SUV"  
            ],  
            "valence": 1.20  
        },  
        "grille_width": {  
            "present": true,  
            "evidence_type": "direct",  
            "examples": [  
                "subtle kidney grilles",  
                "cosmetically, this consists of a new front end",  
                "the grilles get lower and narrower"  
            ],  
            "valence": 0.90  
        },  
        "grille_height": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "new front end",  
                "lower and narrower grilles",  
                "grilles become much less visible"  
            ],  
            "valence": 0.80  
        },  
        "boxy_vs_sleek": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "offers more style than the usual SUV",  
                "a sloping roofline",  
                "more driving engagement than a front-wheel-drive hot hatch",  
                "the shape of the car rounder somehow"  
            ]  
        }  
    }  
}
```

2025-09-06 – Output #4 FORDFOCUS

```
{  
    "functional_word_count": 1290,  
    "visual_word_count": 182,  
    "other_word_count": 672,  
    "valence": {  
        "overall": 1.10,  
        "functional": 1.20,  
        "visual": 0.80,  
        "other": 0.50  
    },  
    "visual_aspects": {  
        "body_shape": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "it's moving even when standing still, thanks to a series of creases and lines",  
                "the Focus comes only with five doors. There's no three-door hatch option",  
                "the Focus Estate rides on the same wheelbase as the five door",  
                "it brought with it a styling theme that Ford dubbed 'new edge'"  
            ],  
            "valence": 0.80  
        },  
        "grille_width": {  
            "present": false,  
            "evidence_type": "not",  
            "examples": [],  
            "valence": 0.00  
        },  
        "grille_height": {  
            "present": false,  
            "evidence_type": "not",  
            "examples": [],  
            "valence": 0.00  
        },  
        "boxy_vs_sleek": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "the idea is that the car looks like it's moving even when standing still",  
                "so quiet at a cruise and stable at motorway speeds",  
                "it has a sophisticated feel that lifts it above the impression one gains from the Astra",  
                "the Focus is not soft and baggy in the manner of the Citroen C4"  
            ],  
            "valence": 0.80  
        }  
    }  
}
```

2025-09-06 – Output #4 HONDACIVIC

```
{  
    "functional_word_count": 1580,  
    "visual_word_count": 245,  
    "other_word_count": 770,  
    "valence": {  
        "overall": 1.10,  
        "functional": 1.30,  
        "visual": 0.60,  
        "other": 0.40  
    },  
    "visual_aspects": {  
        "body_shape": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "its profile lower, wider and more aerodynamic",  
                "naturally longer and wider than its predecessor, with a longer wheelbase and 20mm lower too",  
                "five-door – the third to be assembled at Swindon",  
                "fold down totally flat but also have squabs that flip upright"  
            ],  
            "valence": 0.70  
        },  
        "grille_width": {  
            "present": false,  
            "evidence_type": "not",  
            "examples": [],  
            "valence": 0.00  
        },  
        "grille_height": {  
            "present": false,  
            "evidence_type": "not",  
            "examples": [],  
            "valence": 0.00  
        },  
        "boxy_vs_sleek": {  
            "present": true,  
            "evidence_type": "indirect",  
            "examples": [  
                "a thoughtful evolution... more athletic and elegant",  
                "profile lower, wider and more aerodynamic",  
                "refinement of its ride... will delight high-mileage drivers",  
                "glides over expansion joints and manhole covers with a hushed absorbency"  
            ],  
            "valence": 0.70  
        }  
    }  
}
```

2025-09-06 – Prompt #5 (WITH NO JSON)

Think silently. Temperature=0.

From the review, output a structured JSON with:

- # of words: Overall, Functional, Visual, Other
- Valence scores (-3 to +3): Overall, Functional, Visual, Other
- For front facing visual aspects: bodyshape_hatchback_to_sedan, grille_width, grille_height, boxy_vs_sleek → mark present true/false, evidence_type (direct/indirect/not), 2–4 short verbatim examples (≤25 words), and valence (-3 to +3).

2025-09-06 – Output #3 BMW2SERIES

```
{  
  "word_counts": {  
    "overall": 1270,  
    "functional": 720,  
    "visual": 210,  
    "other": 340  
  },  
  "valence": {  
    "overall": 1.4,  
    "functional": 1.3,  
    "visual": 1.2,  
    "other": 1.0  
  },  
  "visual_aspects_breakdown": {  
    "body_shape": {  
      "present": true,  
      "evidence_type": "direct",  
      "examples": [  
        "It doesn't get much more 'traditional BMW' than the BMW 2 Series Coupé.",  
        "A relatively compact 2+2 coupé has become a rarity, especially with rear-wheel drive."  
      ],  
      "valence": 1.2  
    },  
    "grille_width": {  
      "present": true,  
      "evidence_type": "direct",  
      "examples": [  
        "subtle kidney grilles",  
        "Cosmetically, this consists of a new front end"  
      ],  
      "valence": 1.0  
    },  
    "grille_height": {  
      "present": false,  
      "evidence_type": "not",  
      "examples": [],  
      "valence": 0.0  
    },  
    "boxy_vs_sleek": {  
      "present": true,  
      "evidence_type": "indirect",  
      "examples": [  
        "even numbers are coupés or otherwise swoopier versions of the odd numbers",  
        "sloping roofline"  
      ],  
      "valence": 1.3  
    }  
}
```

2025-09-06 – Output #3 FORDFOCUS

```
{  
  "word_counts": {  
    "overall": 1600,  
    "functional": 1040,  
    "visual": 260,  
    "other": 300  
  },  
  "valence": {  
    "overall": 1.3,  
    "functional": 1.4,  
    "visual": 1.1,  
    "other": 0.8  
  },  
  "visual_aspects_breakdown": {  
    "body_shape": {  
      "present": true,  
      "evidence_type": "direct",  
      "examples": [  
        "the Focus comes only with five doors. There's no three-door hatch option",  
        "An estate version of the Focus is also offered",  
        "This model is a quiet, supple and practical way of sticking to the speed limit"  
      ],  
      "valence": 1.0  
    },  
    "grille_width": {  
      "present": false,  
      "evidence_type": "not",  
      "examples": [],  
      "valence": 0.0  
    },  
    "grille_height": {  
      "present": false,  
      "evidence_type": "not",  
      "examples": [],  
      "valence": 0.0  
    },  
    "boxy_vs_sleek": {  
      "present": true,  
      "evidence_type": "indirect",  
      "examples": [  
        "the Focus follows the latest Ford Fiesta and Ford C-Max in featuring its 'kinetic' design theme",  
        "the car looks like it's moving even when standing still, thanks to a series of creases and lines"  
      ],  
      "valence": 1.3  
    }  
}
```

2025-09-06 – Output #3 HONDACIVIC

```
{  
  "word_counts": {  
    "overall": 2200,  
    "functional": 1320,  
    "visual": 420,  
    "other": 460  
  },  
  "valence": {  
    "overall": 1.2,  
    "functional": 1.4,  
    "visual": 0.6,  
    "other": 0.8  
  },  
  "visual_aspects_breakdown": {  
    "body_shape": {  
      "present": true,  
      "evidence_type": "direct",  
      "examples": [  
        "its profile lower, wider and more aerodynamic",  
        "naturally longer and wider than its predecessor, with a longer wheelbase and 20mm lower too",  
        "retaining the old car's basic mechanical platform and fixed references"  
      ],  
      "valence": 1.0  
    },  
    "grille_width": {  
      "present": true,  
      "evidence_type": "indirect",  
      "examples": [  
        "shortage of design freedom that implies might explain some of the less visually successful parts of the Civic's styling, such as its front wings",  
        "front bumper finish is made piano black, instead of anthracite grey"  
      ],  
      "valence": 0.3  
    },  
    "grille_height": {  
      "present": false,  
      "evidence_type": "not",  
      "examples": [],  
      "valence": 0.0  
    },  
    "boxy_vs_sleek": {  
      "present": true,  
      "evidence_type": "direct",  
      "examples": [  
        "the exterior styling is more athletic and elegant than that of the last car",  
        "this new version described as a thoughtful evolution",  
        "radical, if slightly more conventional look"  
      ],  
      "valence": 0.8  
    }  
}
```

2025-09-06 – POST CHATGPT API

Go back to V1-V4 descriptions that respondents have given us. Can we create 4 clusters to visualize the consumer comments in a 2d space.

Take all the words that CHATGPT says corresponds to a particular visual characteristic. Can we show that ChatGPT text overlaps with consumer text.

2025-09-10 – THOUGHTS

1. Big Picture Idea: First, we show that functional + visual helps better predict the substitution matrix obtained from co-mentions of products using Google Trend data. The visual characteristics could be disentangled or coming from a pretrained embedding like ResNet. Then, write that the accuracy does not drop much if we use disentangled embedding, but we gain the benefit of interpretability.
2. Also, show that if we use ChatGPT, then we obtain the same substitution matrix. The basic point being that Google Trends allows us to construct such a matrix using real consumer deterministic data in a particular time frame in a particular geography. We can't say the same about the data from an LLM.
3. Revised LLM Prompt (line in italics is new) to get review text classified as functional, visual, other by LLM (validate it using human survey respondents later)

Think silently. Temperature=0.

From the review, output a structured JSON with:

- # of words: Overall, Functional, Visual, Other
- Valence scores (-3 to +3): Overall, Functional, Visual, Other
- *For each of Functional, Visual, Other: provide 2–4 short verbatim examples (≤25 words) that the model used to classify text into that category.*
- For front-facing visual aspects: bodyshape_hatchback_to_sedan, grille_width, grille_height, boxy_vs_sleek → mark present true/false, evidence_type (direct/indirect/not), 2–4 short verbatim examples (≤25 words), and valence (-3 to +3).

4. We have LLM's valence output. We will need human survey respondents to validate the valence as well. See the correlation matrix of valence below:

```
> print(round(cor_matrix,2))
      functional visual other overall
functional      1.00  0.07  0.46  0.61
visual         0.07  1.00  0.18  0.46
other          0.46  0.18  1.00  0.75
overall        0.61  0.46  0.75  1.00
```

5. We already have the review text corresponding to each visual characteristic as classified by the LLM. We would need human respondents to validate whether the LLM classification is reasonable.
6. Consideration #1: Re-do the disentanglement by a better dataset, better pre-processing, different supervisory signals etc.
7. Consideration #2: Redo the interpretability survey to obtain more clear response from humans.
8. Gather co-mention of make-models in professional reviews.

2025-10-14 – THOUGHTS

Include more math in the presentations.

Obtain the embedding of Red Toyota, Red Ferrari, Black Toyota, Black Ferrari, Red Apple and then create a 5by5 matrix in each cell contains the cosine similarity. [**DONE**] Visualize the embedding similar to McAuley's papers.

Repeat this with Disentanglement Embedding. (not including color)

Can we do the 1st and 2nd choice experiment. I am not sure how we plan to do this, if not with an experiment. If we do it, then cite Ilya, and then whoever first came up with the 2nd choice idea.

Co-Occurrences using Google Trends: Predict top and top 2 rivals.

Co-Occurrences using Google Trends: Look at sparse matrix prediction literature.

Issue 1: Disentanglement quality is not good because multiple things are changing. Long-term, we need to re-do disentanglement.

Issue 2: Benchmark with ChatGPT. Do a apple-to-apple comparison by giving ChatGPT the same input data and ask it to come up quantified interpretable visual characteristics.

Issue 3: Include other benchmarks (VAE, AE, unsupervised etc.) back in the paper

2025-12-8 – THOUGHTS

Comments from Audience

1. There are small elements in a watch that can also be very important. How can you show that preference for big elements is more important than small elements?
2. Using disentanglement, we can answer what does sporty mean. What visual space in terms of these disentangled visual characteristics does a sporty car occupy?
3. Also, since this is a generative method, one can generate a car in the visual space corresponding to sporty car design. You had mentioned it being compositional.
4. Are there any universal design principles across categories?

Sanity Check: Color and Object Class

I added the sanity check using a red apple and its comparison with Ferrari and Toyota cars in Appendix C.

2025-12-8 – THOUGHTS

Top 1 Rival Prediction Task

I had already added this in Overleaf. The tables are in the Appendix B, and I reference it in Section 4.8.

Additionally, below were some thoughts I had shared about it a while back:

"Also, I added the robustness with top-1 rival prediction. I had earlier done top-2 rival prediction using Google Score. It is in Overleaf now. Random Forest results look good and robust. We might need to see how to interpret the results with XGBoost.

For RF, visual features consistently beat functional in both Top-2 and Top-1, which is great.

For XGBoost, the story is less clean. In Top-2, combined features consistently outperform both visual and functional. But in Top-1, this pattern breaks down: sometimes visual wins, sometimes functional wins, and combined is inconsistent across the different visual representations (Disentangled, VAE, Color, ResNet).

We should discuss whether to: (1) downplay Top-1 entirely and focus on Top-2 as our main result, or (2) keep it but acknowledge the XGBoost inconsistency."