



Pack design and in-store | The case for change

How to test every pack, not just the flagship – for up to 80% less

See how shoppers will respond to every design, before it's printed.



Digestibility

86%

Context

Beauty Shelf 1	81%
Beauty Shelf 2	71%
Beauty Shelf 3	88%
Beauty Shelf 4	71%

By shifting budget, you can test not just your flagship products but your whole portfolio maximising your ROI in-store.

Up to 80%

off annual research spend,
modelled at high testing volumes

£ 295,800

freed up by moving
30% of a year's testing to
Dragonfly AI

91%

as accurate as a 40-person
eye-tracking panel



ONE | WHAT IS THE BENEFIT

A better pack is worth up to 5.5% more volume

Leading research firm NielsenIQ BASES, found that a better pack design can lift expected volume by up to 5.5%. In busy categories like confectionery, soft drinks and alcohol, where brands fight for every sale, that is a big gain. And once the pack is printed, the better design costs nothing more per unit to produce.

But most brands only test their flagship products

It isn't that marketing and product teams don't want to test all packs. It's that testing pack designs properly is slow and expensive so priority is given to the flagships. And the rest? Well, often it's left to subjective opinion.

According to NielsenIQ, nine in ten redesigns fail to make a meaningful sales impact.

The problem is that a shopper never meets your flagship product on its own. They meet the whole shelf at once.

Your brand name, main message, pack image, colour and layout all have about a second to capture the shopper's attention – from about a metre away, in direct competition with your rival product next to it.

Attention matters more than anything else, without it, your product simply won't get picked, and yet many brands are still not testing this crucial element across all their products.

Because testing multiple designs properly costs brands millions

Showing one design to a group of shoppers, or tracking where their eyes go, can cost around £22,000 to £37,000 per test and the results take weeks to get back. For a hundred designs, that can add up to an astronomical £2.2m to £3.7m. Brands simply can't afford that time or investment.

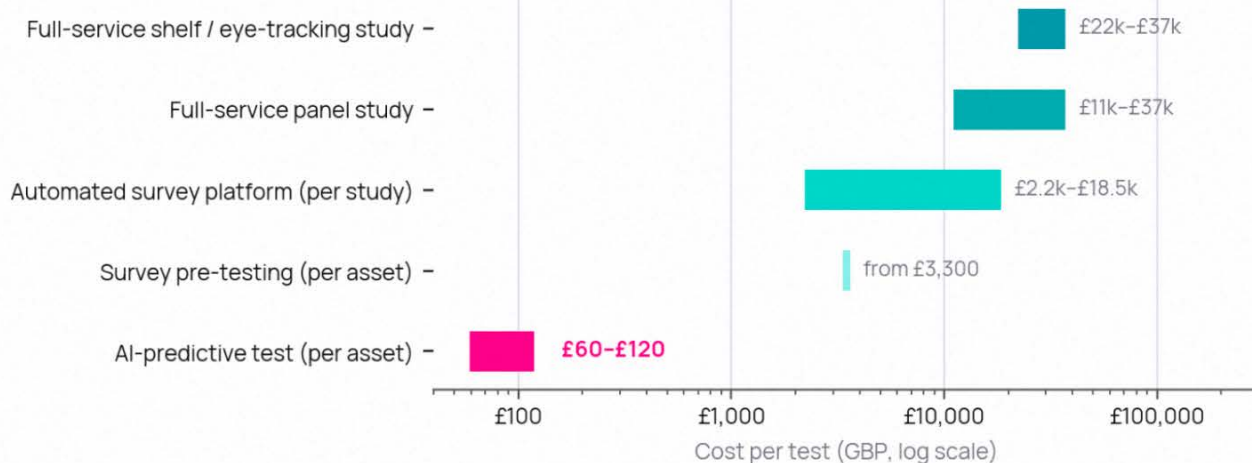
Instead, testing is saved for the big moments – a major relaunch, or the range that next year's growth depends on. The other flavours, variants for other countries, cardboard display trays, in-store posters and seasonal packs all go out on gut feel and experience. Yet every one of them shapes what shoppers think of the brand and whether they pick it up – if they notice it at all.

How can you test more, for less?

We started looking at this in more detail: how can we help brands unlock the ability to test all their creative assets to ensure they maximise ROI without adding any more to their testing budget.

We compared the cost of the traditional testing methods to the average cost of a pack test with Dragonfly AI. When you also consider the time it takes to run a full service shelf or eye-tracking study or a full service panel study versus an AI predictive analysis, it suddenly feels like a substantial saving of both time and money.

This opens the door to testing more broadly across all creative assets. It makes testing not just the flagship products possible, but testing multiple versions, as well as other flavours, variants, seasonal and promotional products. Suddenly the whole line of products across all your brands can be fully optimised for attention performance against your competitors. It's this that shifts those volume increases.



Kantar LINK and full-service study pricing as cited in Dragonfly AI company materials; Zappi per-study pricing (G2, 2026); eye-tracking and shelf study pricing from industry research, July 2026. The AI figure is our own estimate from real customer usage.



THREE | IS IT TRULY COMPARABLE?

Am I choosing quantity testing over quality?

The obvious next question is whether you're jeopardising the results by using a more cost-effective solution. We actually recommend using a combination of both. But let's compare.

The most expensive part of traditional research is not the analysis. It is the people: finding the right mix of shoppers, paying them, running the study and waiting for the results.

Dragonfly AI looks at how shoppers will respond from the pack design itself, so none of that is needed. On attention – whether a pack gets seen, and where the eye lands – our predictions are 91% as accurate as putting 40 people in a room with eye-tracking equipment. On how memorable a design is, it is about as good as asking 34 people. On emotional response, it is about as good as asking 36.

The Attention technology was developed at Queen Mary University of London and is protected by patents in the UK, US and Europe, which the university holds.

What you give up is what only a person can tell you: whether the idea connects with them, what they think it means, and why they would or wouldn't buy it. That is why the best approach is to use both.



Patented in the EU and USA



Birds Eye's new packs were 45% more likely to win attention than rival brands

+65%

Visibility

+1.9%

Volume growth

+0.7%

Value growth

Birds Eye refreshed their packaging across potatoes, fish and peas – their first big redesign in over a decade. Before rolling it out, they photographed and filmed real supermarket shelves and tested the new designs against their old packs and rival brands. In the potato aisle, the new designs were **+14.5% better** at catching shoppers' attention than the packs they replaced, and **45% more likely to win attention than competitors**.



One Durex image redesigned for Amazon resulted in £325k - £370k increase in sales

+23bps

Conversion improvement

+£325k-£370k

market value sales

A pack that works on a shelf won't automatically work everywhere else. On Amazon in Germany, where competitors were gaining ground, the existing Durex multipack image was taking just 5.7% of a shopper's attention on the page. It wasn't a bad pack – it had simply never been designed for that setting.

Reworking it for online shopping lifted its share of attention to around 6.6% and nudged sales up by a fraction of a percentage point, worth a forecast £325,000 to £370,000. Reckitt ran their own A/B test on Amazon to check, and it confirmed the result.



Shift 30% of the tests you already run and free up £295,800 a year

We modelled what would happen if you ran the same number of tests per year, except you shifted a percentage of them from traditional testing to Dragonfly AI. The bars show what you could get back over a year by shifting 10%, 20% or 30%. The colours tell you where that saving comes from.



What you free up by shifting 10%, 20% or 30% of your tests across to Dragonfly AI.

These results are based on what our larger customers actually tested over 9 months to July 2026, priced against industry benchmarks for traditional research. The three categories are additive - you don't need all three to benefit.

Shelf and in-store are 74% of that saving

Shelf tests are the most expensive type of test – so packs and displays alone capture most of the value. A 20% shift covers the software inside the same year; 30% pays back in roughly eight months.

Where these numbers come from

Our platform data for volumes, published research credited inline, and models applying industry pricing to real volumes – reasoned estimates, not quotes.

The 80% headline is modelled on a 150-study programme; the range is 76–87%. Full sources overleaf.

SEVEN | OVER TO YOU

Put us to the test

Everything in this paper comes from published research, our own platform data or a model based on industry pricing. None of it proves anything about your brand. Only you can find that out – and it's completely free.

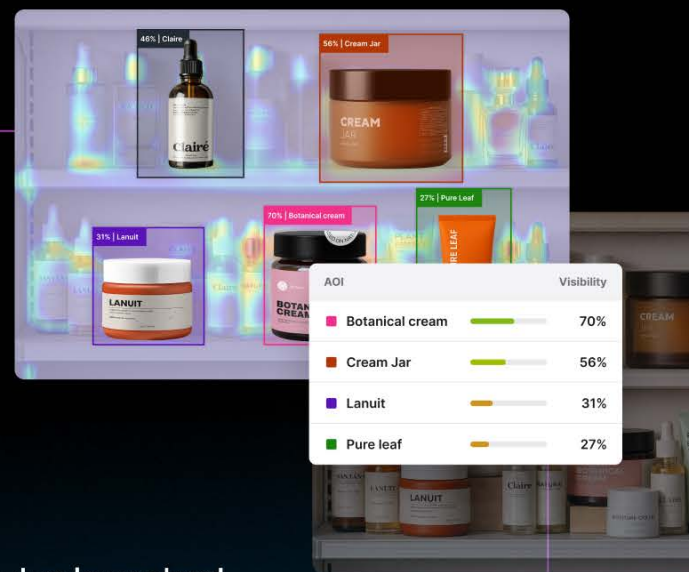
THE OFFER

Send us your
packs.
We'll score them.

Send us the packs and in-store materials you've launched in the past twelve months, along with your main competitors' packs. We'll score them all.

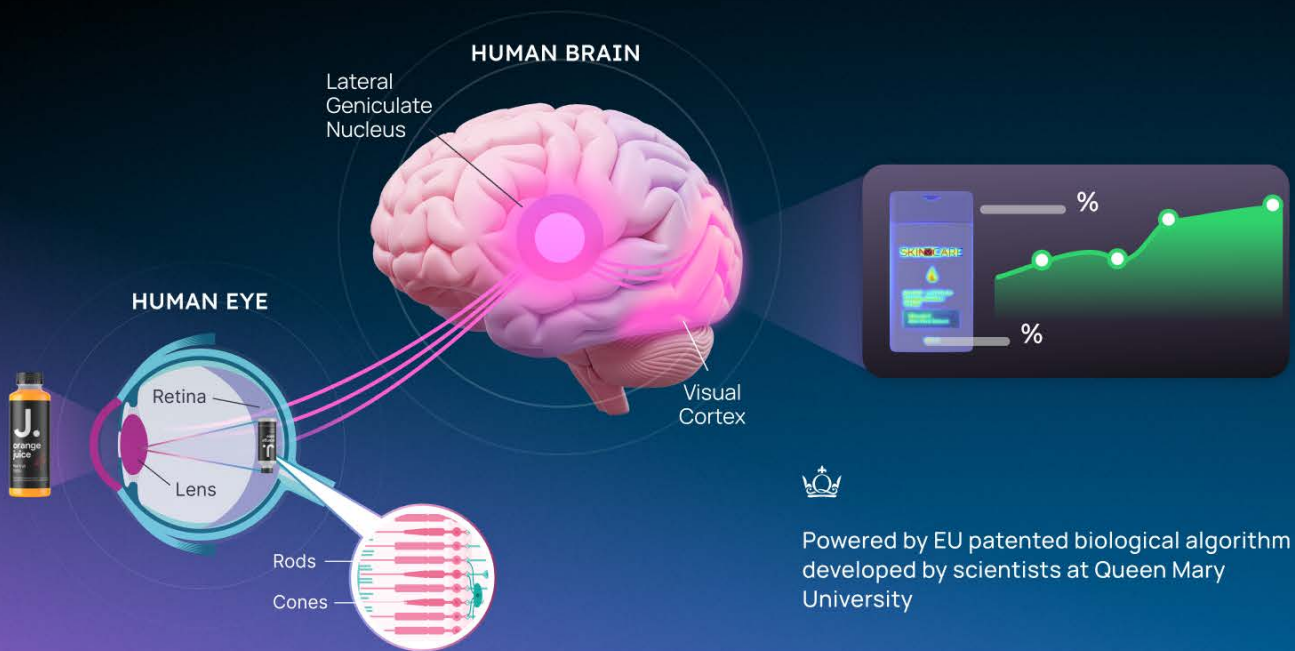
Then compare the results with what you already know: which packs performed well, which disappointed, and which ones your team debated for three rounds. If the results don't tell you anything you recognise, you've lost an afternoon. If they do, you'll have a benchmark for your whole range and a figure that is yours, not ours.

TEST SHELF IMPACT BEFORE ROLLOUT



Your results could feature in the next paper like this one.

Dragonfly AI grew out of neuroscience research at Queen Mary University of London and now works with **more than 50 companies worldwide**, including Estée Lauder, Bacardi, Pernod Ricard, Reckitt, Tangle Teezer, Co-op, Aldi and Subway.



SOURCES AND ASSUMPTIONS

NielsenIQ BASES: optimised packaging can affect forecast volume by up to 5.5%; nine in ten redesigns fail to make a meaningful sales impact.

Testing costs: Kantar LINK from £3,300 per design; full-service studies £11,000-£37,000; Zappi £2,200-£18,500 (G2, 2026);

eye-tracking and shelf studies ~£22,000 - £37,000 (industry research, July 2026). Benchmarks, not quotes.

Savings and payback: our platform data for the nine months to July 2026 priced against those benchmarks; payback against list price.

The 80% headline is modelled on a 150-study, 30,000-interview programme - both sides of that comparison are estimates.

Customer results: Birds Eye, via shelf photography and video. Reckitt's Durex multipack on Amazon Germany, validated by Reckitt's own A/B test.

Accuracy: attention 91% relative to a 40-person eye-tracking panel; memory and emotion equivalent to 34 and 36 people;

validated across 22,620 campaigns. Patents held by Queen Mary University of London.

Dragonfly AI predicts how people respond to what you make — packaging, shelves, adverts, online listings and more.

DRAGONFLYAI.CO

BOOK A 30-MINUTE SESSION WITH DRAGONFLY AI >>