Diversifi
Compendium 2025

Collaborative intelligence:

Driving business value with AI and Behavioral Science

A compendium by:

cowry · Nudge lab · HumanRisk · SONIA FRIEDRICH CONSULTING · Prismatic Strategy · BeHive Consulting · ROGER DOOLEY

AI's usefulness hinges on understanding its users and its use, or it will be found everywhere but nowhere in the productivity statistics. Ultimately, it is a behavioral challenge, not a technical one.

- *Martine Carlson, Partner & CEO at Nudgelab*

Contents

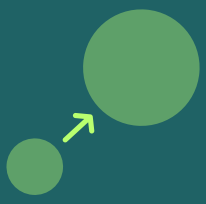
Editors' digest

The strategic advantage of integrating AI with BeSci

In today's rapidly evolving market, the synergy of Artificial Intelligence (AI) and behavioral science offers an opportunity to understand, influence, and predict human behavior and decision-making with greater accuracy than before.

This digest, a collaborative effort by Cowry and Nudgelab on behalf of the Diversifi Global network, captures the ongoing change of behavioral science businesses and how we can actively shape this change in a human-centered way within an AI-driven world. For businesses, the integration of AI and behavioral science can create a competitive advantage for their company and their clients around the world.

The contributions show **three key areas of opportunity for businesses**, which we've used to scaffold this compendium. Read more on the next page.



Augmenting BeSci with AI

Improving AI with BeSci

Nudging for effective AI adoption

We look forward to continuing the discussion on how AI and behavioral science can improve your business and learn how you choose to apply these insights.

Curious regards,

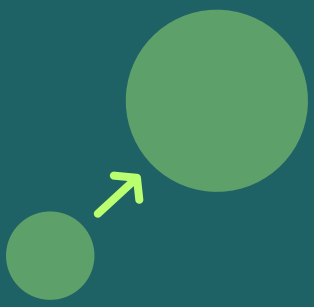
Lisa Bladh
Behavioural Designer
at Cowry

&

Anna Malena Njå
Behavioural Consultant
at Nudgelab

Explore the compendium

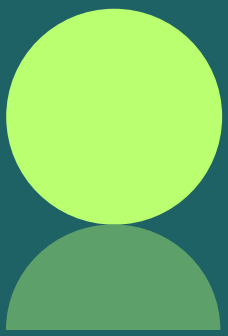
Three stories about AI and BeSci



Augmenting BeSci with AI

Using AI to enhance the speed, scale, and accuracy of behavioral research and interventions. For clients, this means more robust and effective solutions.

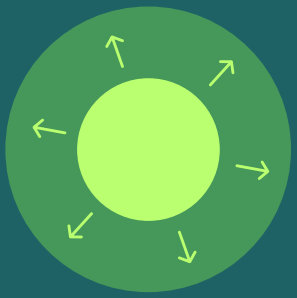
Go there ⟶



Improving AI with BeSci

Applying an understanding of human psychology to build more intuitive, user-centric AI systems and address the non-human irrationality of AI by correcting skewed judgments from biased data. The result is a better customer experience and smarter internal tools that your people actually want to use.

Go there ⟶



Nudging for effective AI adoption

Use behavioral science principles to help organizations integrate AI into existing workflows and adopt new ways of working. This can help overcome natural resistance to change and ensure you realize the return on your investment in technology.

Go there ⟶

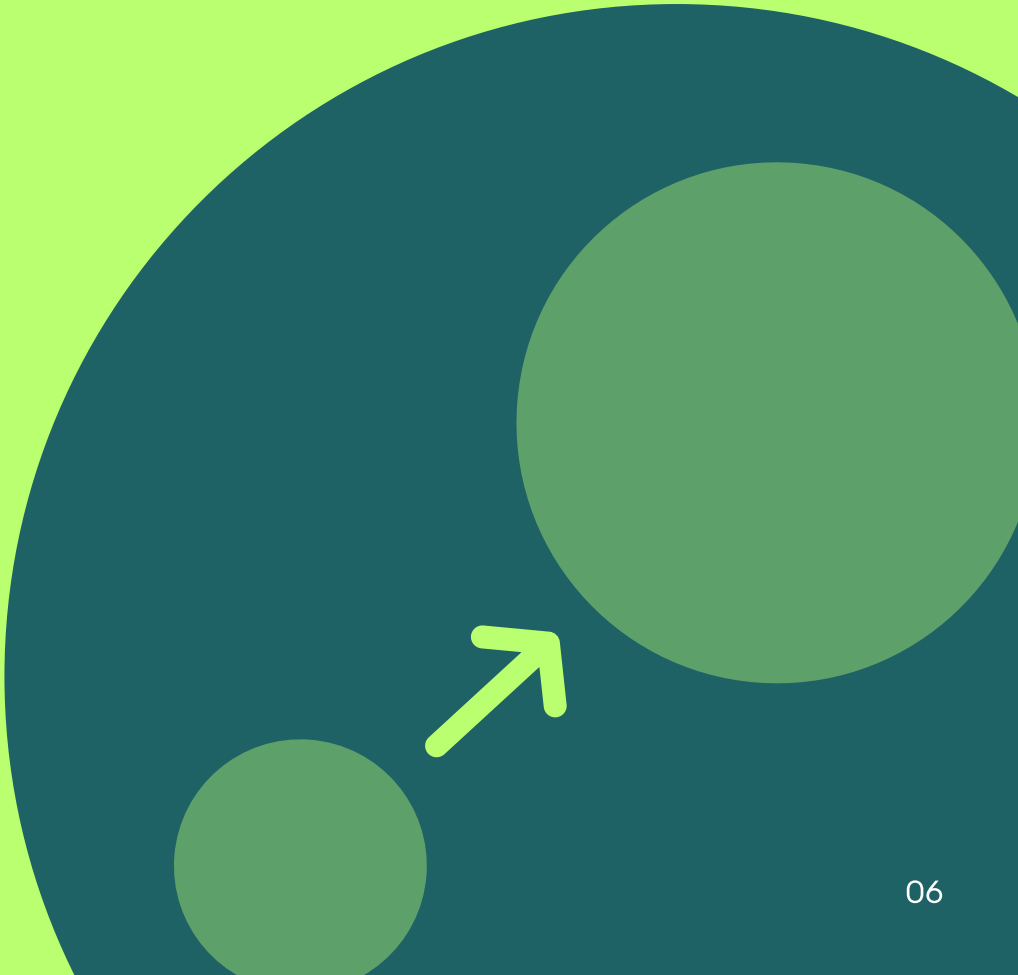
Augmenting BeSci with AI

Using AI to enhance the speed, scale, and accuracy of behavioral research and interventions. For clients, this means more robust and effective solutions.

How AmosNL is solving challenges faced by behavioral science consultancies
Anna Malena Njå - Nudgelab

We're Leading Behavioural Science + AI for eCommerce Globally
Sonia Friedrich - Sonia Friedrich Consulting in partnership with behamics

When Machines Teach Empathy: How AI Amplifies Behavioral Science in Business Communications
Roger Dooley - Dooley Direct LLC



How AmosNL is solving challenges faced by behavioral science consultancies

Predictive AI for behavior change / Scaling interventions

Summary

This piece explores how AI can solve the key challenges of information, speed, and scale that limit behavioral science consultancies. It introduces AmosNL, a tool that integrates AI with human expertise to generate and test interventions on client-specific synthetic personas. It explores human-AI synergy and what future developments could lie on the horizon.

Anna Malena Njå
Behavioural Consultant at Nudgelab

To drive measurable and value-creating behavioral change, behavioral science consultancies must adapt. We believe the key is using artificial intelligence to enhance human expertise. That's why we developed AmosNL, our new tool that provides smarter, evidence-based predictions for effective interventions. AmosNL tailors and tests solutions in client's specific context, with the purpose of unlocking real value.

The following article outlines three key statements that summarize our perception of the current developments of AI usage in behavioral science.

Statement 1: AI offers opportunities to overcome the challenges of behavioral science consultancies

The first statement posits that AI offers significant opportunities to overcome challenges inherent to behavioral science consultancies. These challenges include information, speed and scale. The volume of research makes synthesizing evidence-based, relevant insights for clients across industries a time-consuming task. Traditional processes, like literature reviews and lengthy testing cycles, further limit project speed. This reliance on human-intensive work ultimately constrains the ability of behavioral science consultancies to scale effectively.

AI offers clear opportunities to mitigate these challenges:

- **Enhanced information processing:** AI can summarize large amounts of literature, extract key mechanisms, and identify patterns in specific contexts, allowing for quicker and more precise insights. Tools like Elicit and GeminiResearch are useful for this.

- **Accelerated testing cycles:** AI enables faster A/B testing cycles by facilitating the prediction of effective interventions on synthetic agents within specific contexts, accelerating impact delivery and increasing client confidence in outcome.

- **Scale:** Automating repetitive and time-consuming tasks frees up consultants to focus on high-value strategic thinking, enabling business to scale.

The “PhD quality assurance stamp”

However, despite AI's capabilities, human expertise remains crucial. We call this the "PhD quality assurance stamp." While artificial intelligence is powerful in processing large datasets, pattern recognition, predictive analysis, and optimizing outputs, it lacks human intelligence. Human intelligence refers to humans' nuanced understanding of culture, motivations, norms, social settings, common sense, and ethical reasoning.

An example from Tromsø, Norway, illustrates the risks of trusting AI output without human intelligence. The municipality used AI to write a report that recommended the closure of eight schools and five kindergartens. Researchers at the University of Tromsø later discovered that eleven out of 18 sources were hallucinated.

This case highlights the ongoing need for human reasoning and quality assurance. As my colleague Martine says, "in a time where you don't have to know anything, knowing has become more important".

At Nudgelab, AI serves to augment behavioral scientists by integrating AI's decision intelligence with their human intelligence. This collaboration allows them to concentrate on higher-level strategy, interpreting outcomes, validating insights, and addressing ethical implications.

Statement 2: The future of behavioral science lies in leveraging AI to augment human capabilities and intervention impact

The second statement highlights that the future of behavioral science lies in leveraging AI to augment human capabilities and intervention impact. AmosNL is Nudgelab's answer to this. It is inspired by current technological developments and studies like Shrestha and Krpan's (2025), demonstrating the use of “synthetic participants (...) as a good approximation of human participants for preliminary testing and piloting of policy-relevant views and interventions”.

AmosNL addresses the mentioned challenges of information, speed, and scale by integrating literature reviews, client data, and behavioral insights to generate research-backed suggestions for interventions. Crucially, AmosNL allows us to test these interventions on client-specific synthetic behavioral segments, reducing client risk and increasing potential return of investment. And of course, we guarantee adherence to confidentiality protocols, prioritizing client privacy and data security.

Statement 3: The future of AI is vast and unknown

We recognize that the future of AI is vast and unknown, necessitating continuous testing and iteration of AmosNL. We are continuously exploring best practices for prompt engineering, information volume, and process order within AmosNL. Already identified challenges include how narrow a problem statement needs to be, finding enough relevant research literature for niche projects, and identifying hallucinations.

Looking ahead, we foresee two key developments at the intersection of AI and behavioral science. The first is dynamic hyper-personalization, where AI will analyze individual behaviors and preferences in real-time to create constantly adapting nudges. This means identifying individual biases and contexts to deliver the right nudge to the right person at the right time. The second is autonomous choice architecture, which is using AI and big data to automatically design the environment in which we make decisions, called choice architecture, that steer behavior in a predetermined direction. We also anticipate a change in human-AI collaboration and an increased use of synthetic personas for digital intervention testing, where AI predictions will become increasingly accurate. Furthermore, that behavioral science will play a crucial role in AI development itself, for example, in building humane AI and anthropomorphic agents.

Conclusion

In conclusion, the future of behavioral science consultancies lies in the strategic collaboration of human expertise and AI. By overcoming challenges of information, speed, and scale, our tool, AmosNL, provides smarter, evidence-based predictions for effective interventions. We will continuously iterate AmosNL with ongoing technological advancements, to ensure continued reduced client risk and increased potential return on their investment.

References

BeHive. (2025). Bridging the AI adoption gap — From inaction to action. LinkedIn. https://www.linkedin.com/posts/behive-consulting_bridging-the-ai-adoption-gap-from-inaction-activity-7340728181475160065-KFNx?utm_source=share&utm_medium=member_desktop&rcm=ACoAACzS1lMBb0YltI5H7qgFxEO8thRgArUSTNU

Blythe, P. A. et al. (2025). Comments on "AI and the advent of the cyborg behavioral scientist." Journal of Consumer Psychology. https://doi.org/10.1002/jcpy.1453

Green, M. (2025, June 19). Flawed by Design: Is AI just as biased as we are? And is behavioural science doomed? Claremont. https://claremontcomms.com/2025/06/flawed-by-design-is-ai-just-as-biased-as-we-are-and-is-behavioural-science-doomed/

Katz, L. (2025, February 19). Google "AI Co-Scientist" cracks decade-long research problem in two days. Forbes. https://www.forbes.com/sites/lesliekatz/2025/02/19/google-unveils-ai-co-scientist-to-supercharge-research-breakthroughs/

Kosmyna, N., Uyttersprot, T., Diederich, J., Bense, C., Zogaj, K., Ndiaye, N., & Bense, C. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. arXiv. https://doi.org/10.48550/arXiv.2506.08872

LSE. (2024, October 9). AI and the future of behavioural science | LSE Event [Video]. YouTube. https://www.youtube.com/watch?v=x1hojUoNbFM

Mills, S., & Sætra, H. S. (2022). The autonomous choice architect. AI & Society. https://doi.org/10.1007/s00146-022-01486-z

Momentum Investments. (2023, September 4). Behavioural science and AI: Unlocking better decision-making. YouTube. https://www.youtube.com/watch?v=oFD8eAyKTYI

Shrestha, P., Krpan, D., Koaik, F., Schnider, R., Sayess, D., & Binbaz, M. S. (2025). Beyond WEIRD: Can synthetic survey participants substitute for humans in global policy research? Behavioral Science & Policy. https://doi.org/10.1177/23794607241311793

UN Innovation Network. (2025, April 17). Introduction to Artificial Intelligence and Behavioural Science.. YouTube. https://www.youtube.com/watch?v=nBHgxeHLwxw

Vieira, H., & Vieira, H. (2025, March 4). How agentic AI can be applied to behavioural science. LSE Business Review - Social Sciences for Business, Markets, and Enterprises. https://blogs.lse.ac.uk/businessreview/2025/03/04/how-agentic-ai-can-be-applied-to-behavioural-science/

Wider World. (2023). Behavioral science meets AI a primer. London School of Economics and Political Science. https://www.lse.ac.uk/pbs/assets/documents/Wider-World-Student-Projects/Behavioural-Science-meets-AI-by-Annabel-Gillard-Anna-King-Isabel-Kaldor-Muskaan-Mehlawat.pdf

Want more? Check these reads out.

Thinking clearly about behavioural science and AI: A guide for the perplexed

Beyond the Hype: Putting People at the Centre of AI Adoption

When Machines Teach Empathy: How AI Amplifies Behavioral Science in Business Communications

We're Leading Behavioural Science + AI for eCommerce Globally

Increased profitability in eCommerce / Personalized messaging

Sonia Friedrich
Founder of Sonia Friedrich Consulting, in partnership with behamics

Summary

This piece explores how AI can be practically implemented to drive bottom-line value. It discusses the "Revenue-As-A-Service" platform that Sonia Friedrich Consulting in partnership with behamics deploys intelligent nudges on ecommerce sites and prove success with a live randomised control. Keep reading to learn about their three drivers of success for Behaviourally Intelligent eCommerce - as well as what to avoid.

The emergence of behavioural science + AI is a perfect coupling. It allows for scale. And it increases profitability. At Sonia Friedrich Consulting our focus has always been to drive bottom line impact by measuring outcomes in real time, and make or save client's money. Yes, we're not scared to say it. It's why clients collaborate with us. They want bottom line impact.

To stay ahead of the AI curve, since 2023 we've partnered with well entrenched companies who know this landscape backwards. It's not a fad for them. Why? Because they are leaders in their respective niche areas and they are creating it. This is the same success model we've used over the last decade. Rather than build an Agency, early on we decided we want Consultants and Partners who are the best in the world. We bring them in, as and when clients need them, it reduces overheads, internal costs and gives us and edge with clients too.

For this compendium we'd like to share one of these partnerships. Sonia Friedrich Consulting are a Global Strategic Partner with behamics. A spin-off of St Gallen University, since 2019 they lead in behavioural science + Causal AI for direct revenue uplift for eCommerce.

Let's take a closer look...

We've identified 3 Key Success Drivers for eCommerce:

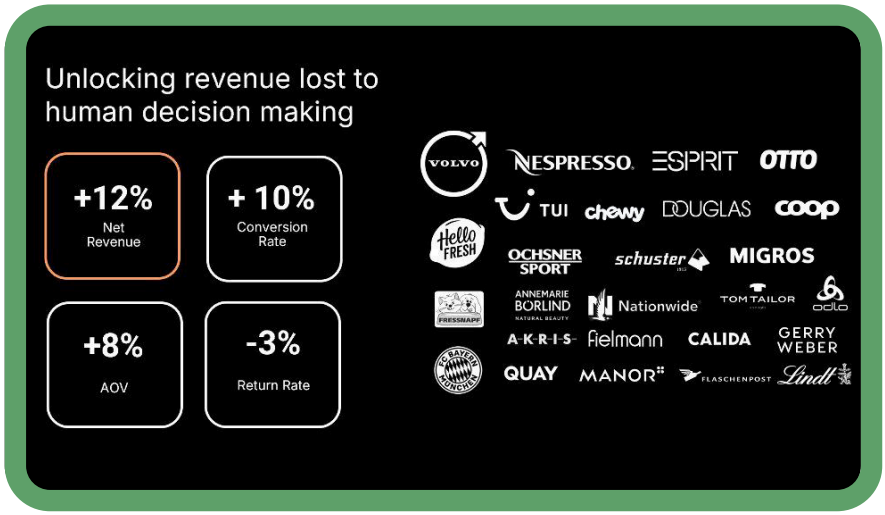
1. Proving Net Revenue Uplift.

Everyone in retail knows there is significant loss and leakage on their eCommerce site. The customer experience is well known. How John or Jane behaves this morning on site can be different to this afternoon, and to tomorrow. Not many have truly optimised this in a dynamic way. We're optimising this by enhancing the customer experience. Reducing cognitive and behavioural friction with intelligent nudges, driven by purpose-built AI, applying machine learning and causal inference modelling.

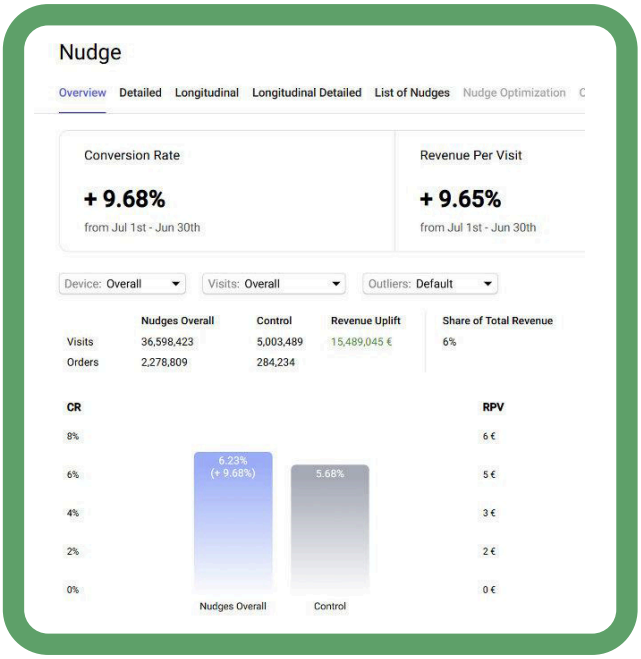
Across 15 countries and replicated across multiple verticals, what started to reduce the issue of product returns in fashion across Europe is now proven in sports, ticketing, insurance, marketplaces and more. What we measure may change. Specifically, we are looking to increase client conversions, no matter how you do this and revenue optimisation. With Volvo which is not a traditional online store the goal was to increase test drives. Replication and longevity of results are critical to success. Behamics (born in science) have developed a Revenue-As-A-Service platform with a live control group on site 100% of the time. It sets a new standard. It's the same rigour that is used for clinical trials. A live randomised control trial sits alongside the nudge group starting with a 50:50 split. Showing results of the nudge v control in:

- Conversion rate
- Revenue per visit
- Average order value
- Add to cart and
- Overall net revenue
- Or whichever metric the client wishes to track. We can modify the dashboard for them.

We're consistently increasing revenue up to 12%. For some clients it has reached 20% and we're reducing product returns.



We focus on the key challenges for the clients and apply nudges to optimise this for them. Interesting to note is that for some clients when product returns are a problem, our nudges can decrease AOV because customers make better choices, yet conversion rate increases, still driving net revenue growth. We're applying machine learning that has behavioural intelligence to act in real-time, with 90%+ confidence.



When the sample size is large enough, we can incrementally move the nudge v control group allocation to a 60:40, 70:30, 80:20 a 90:10 split while we continue to uphold statistical significance. The aim of course is to optimise maximum revenue for the client. For our longer-term clients some are so confident they now want to be 100% with nudge and no control.

Outside of what clients are already doing ie. A:B tests, promotions, new products, or for a marketplace onboarding new sellers, we're proving incremental effectiveness. It's the cherry on top. We're reducing the leaks and losses. Reducing friction.

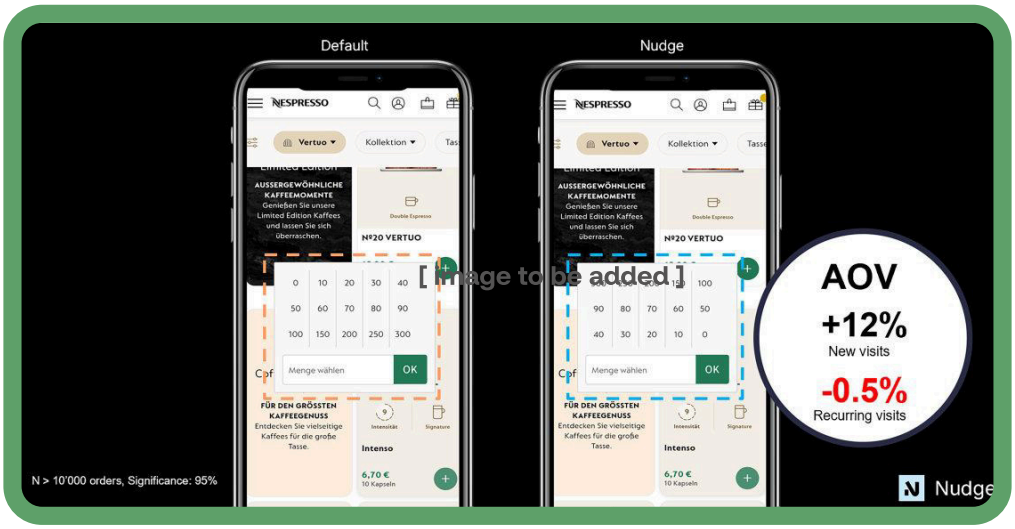
Client feedback that solves real time problems has driven new products. With other products like Diagnostic, we can attribute revenue loss to on-site issues and show exactly where they occur.

This has led to the creation of a suite of Automation Agents to fix these problems where we can. For example, we can fix the image size problems easily, among others. We now have a suite of automations that take the hassle out of fixing these problems and reducing FTEs for clients.

2. BeSci + AI = Behavioural Intelligence

We have a library of more than 500 nudges, some with more than 30 variants of each. Don't worry, a customer doesn't see 500 nudges. The last thing you need is overwhelm and overload. We know this only too well. A brand may have 50 to 100 nudges in their Primary Nudge Set. Yet a customer may see only one or two. Some won't see any at all.

What's interesting is that one nudge that works in sports might backfire in fashion. Or what works for a new customer may backfire for a returning customer. Understanding backfire effects (unintended consequences) is critical. A compelling reason why not to hard code nudges into a site. This is a real limitation of the A:B split testing norm that exists today. We predict this will change drastically and are at the forefront of creating this new norm. You need a dynamic approach that adapts to individual customer behaviour. It is not about running a test and picking a 'winner'.



Intelligent Nudge Example: Increasing AOV with Anchoring

What we are doing at Behamics is using Causal Inference Modelling. Unlike traditional predictive modeling, which only learns associations (correlations). Causal inference is about identifying true cause-and-effect relationships. Plus, we focus this on the highest value pages in the customer journey: The PLP, PDP and Cart. Where customers face friction, a lot of choice and decision fatigue.

Intelligent Nudge Example: Reducing Product Returns with Loss Aversion

A loss aversion nudge can be triggered if a customer orders multiple sizes of the same product. The intervention focuses on the amount of time a customer will lose (32 minutes) by trying them on and having to return what doesn't fit versus the time it takes to look at the size guide (3 minutes). This nudge alone helped reduce return rates by 5% for a client across 15 countries. We now have a library of more than 30 nudges for product returns alone.

behamics in collaboration with another behavioural science agency, academic institutions and clients, has run the largest live product returns experiment globally. With more than 200,000 customers, 100,000 orders across 8 countries. There are insights anyone in eCommerce needs to know.

Intelligent Nudge Example: Reduce Discounting

Discounting, offers and promotions saturate the eCommerce customer experience. The amount of money spent for on-site promotions is excessive. We've learned and can prove not everyone needs a coupon. We can predict who needs a coupon, who doesn't, and even who will come back without one! This saves on the client's promotional budget immediately. We are helping clients reduce their reliance on discounting and building their brand value again with customers.

Live in Minutes

Making it easy is mandatory. A key focus is to make introducing Nudge capacity light for clients. How do we make it as low risk as possible? eCommerce teams are busy enough. We also know most clients don't have a behavioural science team. From the beginning the aim was, where possible to 'do it for them'. This has stood true and continues to be a driving force for new product development. Here's how we do it to make it as easy as possible for clients:

- **Our behavioural scientists create the client Primary Nudge Set** (with client input and approval to ensure it fits within brand and company guidelines)
- **Clients have behavioural science support.** We don't have customer success people. We assign a behavioural scientist to you.
- **It's one line of code.** Seriously. Every tech team laughs and then they see its real. Unheard of. Many of our clients roll their eyes and think we've heard this before and it took months. You can literally go live in minutes. Feel free to challenge us to prove it.

  `<script src="https://cdn.behamics.com/behamics.js"></script>`

Is BeSci + AI a bed of roses?

No. We want to share some learnings when you start your AI journey so you don't fall into these three traps.

- **Accuracy vs. Persuasion:** AI is compelling because it sounds authoritative, but it can be confidently wrong. When we're designing nudges, an inaccurate output is worse than no output — it can nudge in the wrong direction. Always validate before you persuade.
- **Overload vs. Simplicity:** People are already overwhelmed by AI talk. Cognitive overload makes messages less effective. Nudges must cut complexity down, not add to it — otherwise, you risk feeding into fear or hype instead of driving real behaviour change.
- **Personalisation vs. Manipulation:** AI makes hyper-personal nudging possible. But with great precision comes ethical risk. If the nudge feels like a push instead of a guide, trust erodes. Keep autonomy and transparency front and centre.

behamics was the first of our AI partnerships for Sonia Friedrich Consulting back in 2023. We have another with IZZYAGENTS who monetise chatbots. Again, they are not novices. They have led this space since 2018. We will continue to partner with companies that lead in AI in their respective niche areas and stay ahead of the curve for our clients. Critical for us is to provide solutions with proven replication and longevity. Why? We want to stay focussed and not get distracted. Bottom line impact, well today that matters more than ever.

FOR MORE:

BeSci Agencies: We are now partnering with BeSci Agencies because it is a natural fit. . BeSci Agencies and Consultants who want to add AI into their offer yet don't need to become the experts.
eCommerce Brands: Ask for a demo of behamics to see more.
You: Ask for a copy of the WHITE PAPER 'The Psychology of the Return'.

Contact sonia@soniafriedrich.com today.
Results reproduced with permission from behamics

Want more? Check these reads out.

Computed irrationality: A new frontier for Behavioural Science and AI

How AmosNL is solving challenges faced by behavioral science consultancies

Thinking clearly about behavioural science and AI: A guide for the perplexed

When Machines Teach Empathy: How AI Amplifies Behavioral Science in Business Communications

Learning to understand AI / Personalized messaging

Roger Dooley
Author, Keynote Speaker, Neuromarketer

Summary

This piece explores how empathy failures in corporate communications lead to costly brand damage. It discusses how modern AI can amplify behavioral science, showcasing a real-world example how an AI accurately predicted customer backlash to a tone-deaf message and drafted an effective alternative. It also introduces the 'Empathy Audit Protocol' for using AI to improve customer relations.

The most expensive sentence in business might be, "We apologize for any inconvenience." Not because apologies are costly, but because this phrase often signals a fundamental failure of emotional intelligence. Today, AI models can use behavioral science and empathy principles to help organizations predict customer reactions, avoid relationship-damaging mistakes, and build stronger connections.

Consider Southwest Airlines' recent decision to charge for checked bags, ending their decades-old "bags fly free" policy. This wasn't merely adding a fee or offering “options” as their CEO euphemistically suggested. Rather, it violated a core brand promise and triggered widespread loss aversion. Customers didn't perceive they were gaining new pricing flexibility and more choice. They experienced losing something they psychologically owned. The predictable result: customer anger, brand damage, and defection to competitors. Both American and Delta Airlines immediately offered more generous status matches to poach Southwest's most loyal flyers (1).

Here's a fact that will surprise some: An AI language model could have predicted customer reactions with greater accuracy than the executives who made the decision. This isn't speculation. It’s readily demonstrated by systematic testing of AI's ability to apply behavioral science principles to real business scenarios.

The Empathy Paradox: Why Algorithms Understand What Executives Miss

Recent research from the Universities of Geneva and Bern revealed something that challenges our fundamental assumptions. When tested on standardized emotional intelligence assessments, AI models averaged 81% accuracy, far higher than humans at 56%. GPT-4 scored in the 89th percentile of human performance (2). While AI doesn't "feel" emotions, it excels at cognitive empathy, i.e., recognizing and predicting emotional responses based on patterns in human behavior.

This capability transforms how we apply behavioral science in business contexts. I tested Claude (Sonnet 4 model) with a real cruise line announcement that had infuriated customers. Unlike the sweeping Southwest policy announcements, this one was a private communication affecting a small group of customers. Mid-cruise, six-star luxury line Silversea forced 700 luxury guests to delay their departure from the ship by four hours for a marketing photoshoot. Guests had to rebook flights, ground transportation, and other arrangements during the busiest travel period of the year (3). This change was communicated in a sterile letter that failed to acknowledge the inconvenience, even suggesting guests celebrate on the pool deck during the photoshoot.

Claude rated the letter as "poor" for empathy, noting the tone was corporate, detached, euphemistic, and promotional. The AI pointed out that "asking guests to wake up at 6:15am to celebrate and toast something that is inconveniencing them" was "tone-deaf." Claude also identified a variety of behavioral science principles that the communication violated, noting that guests would experience loss aversion and reactance by being forced to change travel arrangements they had carefully constructed months before. Their social identity and status would be threatened when even elite loyalty members were treated like mass-market customers instead of getting the expected "six-star" treatment. Claude mentioned attribution theory, noting that guests would attribute the inconvenience to controllable corporate choices, not forces beyond the company's control. (The guests wouldn't be wrong.)

More importantly, the AI correctly predicted specific customer reactions: rebooking stress during peak travel season, social media backlash, and damage to brand loyalty among the company's highest-value customers. These effects were indeed visible in guest comments in cruise forums and discussion groups. The AI drafted a far better letter to guests that acknowledged the disruption, offered concrete remedies, and demonstrated understanding of the emotional impact—elements entirely absent from the original communication.

When I asked if the CEO should approve the photoshoot, Claude recommended rejecting the idea due to potential brand damage and loss of future revenue from high-spending guests. (When I've used this example in a conference presentation, by far the most common response from perplexed audience members is, "Couldn't they have used Photoshop? Or AI?")

The Behavioral Science Amplification Effect

AI doesn't replace behavioral science expertise; it scales it:

1. **Scale Without Sacrifice:** Organizations can now audit thousands of customer touchpoints for psychological friction and emotional triggers. Anything from important policy announcements to routine order acknowledgments can be fine-tuned.
2. **Pre-emptive Pattern Recognition:** AI can identify potential behavioral land mines before they explode. By analyzing communications through the lens of cognitive biases (Kahneman/Tversky), social proof and reciprocity (Cialdini), loss aversion, etc. AI models can help organizations avoid empathy failures that cause customer defection and employee turnover.
3. **Cultural and Contextual Calibration:** Modern AI models can adjust for cultural differences in emotional expression and expectation. Even leaders with strong emotional intelligence need help in avoiding culture-driven mistakes.

From Theory to Practice: The Empathy Audit Protocol

The integration of AI and behavioral science follows a systematic protocol that any organization can implement:

First, provide your preferred AI model (e.g., Claude, ChatGPT, Gemini, etc.) with context about the audience and situation, including demographic and psychographic profiles, relationship history, and current emotional state. For instance, airline passengers facing flight changes are already stressed; loyal customers expect recognition of their status; price-sensitive segments respond differently to fee increases than premium buyers.

Second, have the AI analyze the communication for behavioral triggers. Does it invoke loss aversion, arguably the most influential bias in Kahneman and Tversky's prospect theory? Create unnecessary cognitive load? Violate reciprocity norms? Ignore the peak-end rule in service recovery? And, ask the AI to identify empathy failures.

Third, request specific revisions that maintain business objectives while addressing identified psychological friction points. Ask the AI to explain the reason for each revision.
Finally, validate the AI's recommendations against human expertise and contextual factors the AI might miss, such as regulatory requirements, financial constraints or competitive dynamics.

The Competitive Necessity of AI Empathy

Organizations that fail to combine AI's pattern recognition with behavioral science and empathy insights will find themselves at an increasing disadvantage. Their competitors will predict customer reactions more accurately, communicate more persuasively, and recover from service failures more effectively.

But perhaps the greatest value lies in prevention. Every corporate communication disaster, from tone-deaf layoff announcements to customer-alienating policy changes, represents a failure to anticipate human psychological response. AI, guided by behavioral science, transforms this from guesswork to science.

The tools exist. The behavioral frameworks are proven. The combination of AI's computational power with behavioral science's human insights offers unprecedented capability to communicate with genuine empathy, even when that empathy is algorithmically augmented. **In an era where a single tone-deaf message can destroy decades of brand equity, the integration of AI and behavioral science isn't just valuable, it's what keeps brands alive.**

References

1. Dooley, R. (2025, May 29). Southwest Airlines just made a costly mistake in consumer psychology. Forbes. https://www.forbes.com/sites/rogerdooley/2025/05/29/southwest-airlines-just-made-a-costly-mistake-in-consumer-psychology/

2. Elyoseph, Z., Levkovich, I., & Shinan-Altman, S. (2024). "Assessing the capacity of large language models to exhibit cognitive empathy." University of Geneva and University of Bern Study on AI Emotional Intelligence.

3. Dooley, R. (2024, March 6). Royal Caribbean cruise line photoshoot sparks passenger chaos. Forbes. https://www.forbes.com/sites/rogerdooley/2024/03/06/royal-caribbean-cruise-line-photoshoot-sparks-passenger-chaos/

Want more? Check these reads out.

AI for Perspective, Not Accuracy: Bringing Behavioural Science to Compliance

We're Leading Behavioural Science + AI for eCommerce Globally

Computed irrationality: A new frontier for Behavioural Science and AI

Improving AI with BeSci

Applying an understanding of human psychology to build more intuitive, user-centric AI systems and address the non-human irrationality of AI by correcting skewed judgments from biased data. The result is a better customer experience and smarter internal tools that your people actually want to use.

Computed irrationality: A new frontier for Behavioural Science and AI
Lisa Bladh - Cowry

Thinking clearly about behavioural science and AI: A guide for the perplexed
Elina Halonen - Prismatic Strategy

Computed irrationality: A new frontier for Behavioural Science and AI

Learning to understand AI / Improving AI-systems

Summary

The article introduces Machine Psychology, a new field that tries to understand the non-human irrationality of AI models that risks large-scale hiccups for businesses. It argues that recognizing and addressing its' flawed judgements stemming from biased data and non-human reasoning is a crucial new frontier for behavioral science consultancies to explore.

Lisa Bladh
Behavioural Designer at Cowry

AI models have quickly become trusted companions in almost every aspect of our lives. These thinking-partners, presented to us as pixels on a screen rather than humans of flesh and blood, are relied on as powerful tools for both analysis and decision-making with unparalleled ease. In other words, these models are seen as rational and deliberative - essentially representing what we know as System 2 thinking. What many of us are missing, though, is that AI displays a whole new type of irrationality. These overlooked errors originating in AI models present a new frontier of challenges ripe for BeSci intervention.

As we've grown to appreciate AI's capability, we've also learnt that algorithmic appreciation is easy to overdo. As a society, we're increasingly aware that AI makes flawed judgements all the time. Many of these flaws stem from biases inherited from training data. We're all familiar with the scandals of the past decade, from discriminatory hiring algorithms and face-recognition technology failing to detect non-white faces to overwhelmingly white outputs from image generation platforms. However, this is just one piece of the puzzle. The true challenge lies in a new black box for psychology: the hidden ways AI algorithms arrive at conclusions. And just like in the 60s, when the behaviouralists were uncovering the inner workings of the human mind, we're now looking out on a new frontier of our field seeking to understand the inner workings of a technology we made with our own hands (and brains). It's called Machine Psychology, and it could represent a whole new horizon for Behavioural Science Consultancy.

This discipline acknowledges the bounded rationality of AI models trained on data that is coloured by our own heuristics and biases. It uses the principles of experimental psychology to understand the psychology of machines, analysing the patterns between model in- and outputs. Human-made training material is steeped with biases that we fail to acknowledge, leading to irrational leaps in judgment reaching beyond the more talked-about discriminatory judgments emerging from AI technology over the past years.

Over machine psychology's infancy, various similarities between human reasoning and AI reasoning has been uncovered. When tasked with Kahneman and Tversky's classic riddles, AI models tended to mimic human predictable irrationality, for example by neglecting base rates, ignoring the influence of sample sizes and falling for anchoring effects. Naturally, however, as models are updated and corrected, many of these biases have started to fizzle out. If you try and trip ChatGPT up with a classic Linda the librarian task, it's already three steps ahead of you.

Now what's interesting is that as AI responses get more and more correct in general, as we detect and correct for inherited cognitive biases, we are also finding new ways models' reasoning capabilities are irrational and, importantly, distinct from humans. In a recent study by Macmillan-Scott and Musolesi (2024) exploring the irrationality of seven commonly used AI models, they found that, whilst models are generally getting more and more correct, the proportion of incorrect responses that stem from non-human like errors in reasoning is increasing. This new theme of irrationality rising out of the black box of AI-reasoning poses an unprecedented threat to individuals and organisations: lapses in rationality beyond what we recognise as human psychology.

As human-like mistakes are getting less and less prevalent, we as behavioural consultancies have a unique opportunity to explore, develop and employ Machine Psychology in future client interactions. In a future world where we depend on AI in almost every domain of life, behavioural science will play a key role in building fruitful, seamless interaction - but only if we grasp Machine Psychology. Let us explore three possible venues for a new set of Machine Psychology Principles:

- **Differing semantic structures:** First, machines operate on a fundamentally different semantic structure than humans. While humans build meaning by connecting the dots multimodally, AI models learn by finding statistical patterns between the dots. In other words, most AI models don't yet naturally perform the same generalisations we're used to in human reasoning. This difference can lead to new types of logistical lapses that we wouldn't expect from a human brain.

- **Vulnerability to noise:** Second, machine reasoning can be fragile in ways that human reasoning isn't. The vulnerability to the tiny input changes that enables adversarial attacks to cause such massive changes in model outputs is starkly different from human reasoning, which in many ways is protected from perceptual noise.

- **Causal struggles:** Thirdly, machines struggle with causality in a way humans do not. Human psychology has a natural understanding of cause and effect, built on years of direct interaction with the physical world. We intuitively grasp that throwing a ball causes it to fly, and that rain causes the ground to get wet. AI, of course, does not "experience" the world in this way. It learns from ready-made causal models of the world, lacking the direct, embodied understanding of causation that underpins human psychology.

These new forms of irrationality demand a new approach to our work, one that extends beyond simply debiasing inherited human heuristics and focuses on the unique psychological vulnerabilities of machines. Take the differing semantic structures, for example. Imagine a large retailer firm using an AI model to generate advertising copy for new products. The AI, having learned from a vast corpus of online text, frequently uses statistically common buzzwords like "paradigm-shifting" or "next-generation" in awkward or incorrect contexts. Because of the AI's semantic structure bias, the copy risks sounding meaningless and fails to connect with consumers. A behavioural science consultancy, applying the principles of Machine Psychology, would be uniquely equipped to diagnose, and resolve this error. By leveraging their expertise in how humans build meaning and context, they could help the client understand why their AI model makes these logistical lapses and help develop the model to better meet human needs.

For behavioural science consultants, these developments present a new service horizon. Machine Psychology allows for a theoretical venue through which the same experimental methods we've used for decades to understand human behaviour can be turned towards the systems we create. This could mean a range of projects: from conducting behavioural research to resolve AI biases with a Human-AI interaction lens to proactively identifying AI biases before they cause large-scale issues, to developing prompts and providing targeted code adjustments to correct for lapses in models' thinking. By leveraging our expertise in uncovering, applying and ideating using Behavioural Science methods, we can together start to uncover the black box of AI reasoning and develop the targeted interventions fit for our client's needs as the world grows ever more AI integrated. As we enter into a new era of AI, understanding its unique psychology is no longer just an academic exercise but a business imperative.

Want more? Check these reads out.

How AmosNL is solving challenges faced by behavioral science consultancies

When Machines Teach Empathy: How AI Amplifies Behavioral Science in Business Communications

Beyond the Hype: Putting People at the Centre of AI Adoption

Thinking clearly about behavioural science and AI: A guide for the perplexed

Scaling interventions / Improving AI-systems / Personalized messaging

Summary

A three-dimensional matrix to provide a structured approach to identify the intersection of AI and behavioral science, categorizing the relationship based on Purpose (is AI a tool or a subject of study?), Scale (is the focus on individuals or entire systems?), and Interaction (is the engagement transactional or relational?).

Elina Halonen
Founder and Strategist at Prismatic Strategy

Artificial intelligence is reshaping how decisions are made, communicated, and acted upon. For organisations, this creates opportunities but it also creates uncertainty about where and how behavioural science can add value. The term “AI + behavioural science” can refer to very different activities:

- Using AI tools to accelerate research, review evidence, or scale interventions.
- Analysing how algorithms influence behaviour through personalisation, choice architectures or hypernudges.
- Applying behavioural insights to shape the design, alignment, or governance of AI systems themselves.

Each of these is valid, but without clarity, the label risks becoming so broad that it is hard for clients to act on, and hard for consultants to explain where they add value. To clarify these differences and make this emerging space easier to navigate, I developed a matrix. It maps the roles behavioural science can play with AI, the modes of engagement, and the kinds of systems involved.

For clients, the matrix can help:

- Scope and focus projects by identifying where behavioural science can add the most value
- Select the right expertise by ensuring the right skills are matched to the right kind of challenge
- Reduce the risk of blind spots by accounting for the full behavioural context of AI systems

The Matrix

To make sense of the many ways behavioural science intersects with AI, the matrix maps three core dimensions. Each one captures a meaningful difference in how we engage with AI systems, what kind of behavioural expertise is needed, and where the work is focused.
Before we get to the dimensions, there's one important distinction to keep in mind: the term AI can mean many things. In the past few years, AI has become synonymous with Generative AI, even though it is only a part of a bigger field. For simplicity, I will group everything else into "non-generative AI" that operates through classification, prediction, or personalisation. These includes e.g. recommender algorithms, credit scoring tools, or dynamic choice architectures – all of which have been around for much longer.

The matrix combines three intersecting axes:

- **Purpose:** Is AI being used as a tool to support behavioural science work, or is behavioural science being used as a lens to interrogate or guide AI systems?
- **Scale:** Is the focus on individual-level decisions and experiences, or on the broader population- or system-level dynamics that AI creates or reinforces?
- **Interaction:** Are humans engaging with AI in a one-off, transactional way, or in an ongoing, adaptive, relational dynamic?

Each axis gives a different view of the relationship between behavioural science and AI. Used together, they create a multidimensional map of where contributions can be made and where different kinds of expertise might be needed. In reality, projects often span multiple areas so these dimensions are best understood as continuums.

1. Purpose: AI as a tool vs. behavioural science as a lens

This axis helps clarify whether AI is the instrument being used, or the subject being examined: is AI supporting behavioural science work, or is behavioural science being used to understand or shape AI systems? At one end, AI is used as a tool to extend or enhance what behavioural scientists already do. This includes work where AI enables faster analysis, broader reach, or greater precision.

AI as a tool for behavioural science

Here, AI is used to support behavioural science workflows or goals, such as:

- Reviewing literature at scale
- Analysing behavioural data
- Designing or delivering personalised interventions
- Running simulations or virtual experiments

At the other end, behavioural science acts as a lens for interrogating how AI systems influence human behaviour—whether by shaping decisions, altering incentives, or introducing new sources of bias and uncertainty.

Behavioural science as a lens on AI

Here, behavioural science is used to examine, critique, or guide the development of AI systems:

- Identifying behavioural risks, biases, or blind spots
- Understanding how systems shape cognition, trust, or agency
- Aligning AI design with human needs, norms, and decision environments

Most real-world projects fall somewhere between these poles. For example, a team using an LLM to streamline intervention design might also need to assess how the model's outputs reflect its training data and embedded assumptions.

2. Scale: From individual experiences to system-level effects

This axis helps clarify the **level at which behavioural dynamics are being addressed**: is the work focused on **individual experiences and decisions**, or on **system-wide effects and patterns**?

At one end, behavioural science is applied to micro-level interactions, where AI influences how individuals perceive, decide, and act. This is where BeSci traditionally operates: modelling cognitive processes, shaping decision environments, or designing friction and feedback loops.

Micro-level focus

Here, behavioural insights are used to shape or evaluate how AI systems interact with individual users:

- Personalising recommendations or messages
- Reducing friction in decision pathways
- Modelling user attention, motivation, or trust
- Designing choice architectures in AI-enabled interfaces

At the other end, behavioural science is used to understand macro-level impacts—how AI systems affect populations, institutions, and social structures over time. This includes evaluating not just what an AI system does to one user, but how it scales, who it benefits or excludes, and how it reshapes norms and expectations.

Macro-level focus

This involves a systems-level lens on how AI influences behaviour at scale:

- Assessing institutional adoption and policy implications
- Investigating long-term effects on incentives, norms, or public trust
- Analysing disparities in access, influence, or outcomes
- Supporting governance, oversight, and accountability mechanisms

This is a particularly important space for behavioural science to contribute because our perspective helps link individual experiences with potential unanticipated, adverse consequences. For example:

- **Recommender systems:** Optimising for engagement can increase short-term relevance and user satisfaction, but also amplifies filter bubbles, reinforces confirmation bias, and contributes to social polarisation. These effects may go unnoticed unless behavioural science perspectives are applied at the system level.
- **Generative models used in decision support (e.g. LLM copilots):** While they increase speed and confidence in completing tasks, they may also introduce subtle distortions—users can over-trust outputs, anchor on misleading suggestions, or unknowingly internalise the model's assumptions. Behavioural science can help anticipate these risks by modelling how humans interact with AI-generated content over time.

3. Interaction: Transactional vs. relational dynamics

This axis highlights the nature of human–AI interaction. Are people engaging with the system in a one-off, outcome-focused way—or building ongoing, adaptive relationships over time?

Transactional dynamic

At one end, AI systems operate transactionally by generating outputs like credit scores, fraud alerts, or product recommendations in response to predefined inputs. These decisions often feel one-sided: people can't contest, discuss, or adapt them. Behavioural science can help organisations understand not just how people respond in the moment, but how these systems affect long-term perceptions of fairness, control, and trust. That matters because:

- People are more likely to disengage or push back when they feel excluded from decisions
- Opaque outcomes can damage brand trust, even when technically correct
- Small frictions or perceived unfairness can accumulate into reputational risk

Relational dynamic

This dimension brings the human–AI relationship into focus. Some systems do more than just give answers—they behave in ways that feel responsive, even social. As people interact with these tools over time, they start to form expectations, habits, and patterns of trust. How these relationships develop can shape not only what people do, but how they think and feel. Tools based on large language models often behave less like tools and more like collaborators or conversational partners. They respond to tone, adapt to prompts, and influence how users frame their own thinking. Behavioural science can help guide these interactions to support clarity, autonomy, and appropriate trust:

- Analysing how tone, fluency, or confidence shape user perceptions
- Exploring how trust, reliance, or overreliance build over repeated use
- Designing interactions that set realistic expectations about what the system can and can't do

This is another area where behavioural science offers distinctive value. As AI systems become more embedded in everyday interactions, the line between tool and teammate starts to blur. Understanding how users interpret, adapt to, or rely on these systems especially when the system feels responsive or social requires a behavioural lens because these dynamics involve trust, norms, and meaning making, which behavioural scientist are trained to understand. For example:

- **Opaque scores shape behaviour from a distance:** Risk scores and classification tools assign labels like creditworthiness, fraud risk, hiring potential that influence high-stakes decisions. Even when wrong or contestable, these scores feel definitive. Users adapt their behaviour based on opaque outputs, which can entrench disadvantage, erode trust, or drive workarounds.
- **Repeated interactions create relational expectations:** As users engage with AI assistants, copilots, or companions, they develop expectations that go beyond function. Tone, responsiveness, and memory shape how users interpret the system's intent, reliability, and personality. Over time, people may respond to these systems socially, even emotionally which creates new behavioural and ethical dynamics.

Pulling it together

These three dimensions – purpose, scale, and interaction – offer a structured way to think about the role of behavioural science in AI. The framework is intended to support reflection and strategic clarity so that we can notice what kind of dynamics are in play, what questions are being asked, and which ones might be missing.

The dimensions offer a way to map how behavioural science and AI fit together. In practice, most projects land somewhere in this three-dimensional space: for example, a tool built for individual users might end up shifting organisational norms, or a system designed for speed might evolve into something people rely on and relate to.

Every AI system ends up entangled with human expectations, assumptions, and responses. Even when the technical side is dominant, the behavioural consequences tend to show up: what feels like a data pipeline can turn out to be a trust problem, and what looks like a workflow optimisation can start to shape incentives or shift accountability.

As AI becomes more embedded in products, services, and decisions, behavioural perspectives can help teams ask better questions, see unintended effects earlier, and design with people in mind from the start.

Want more? Check these reads out.

AI for Perspective, Not Accuracy: Bringing Behavioural Science to Compliance

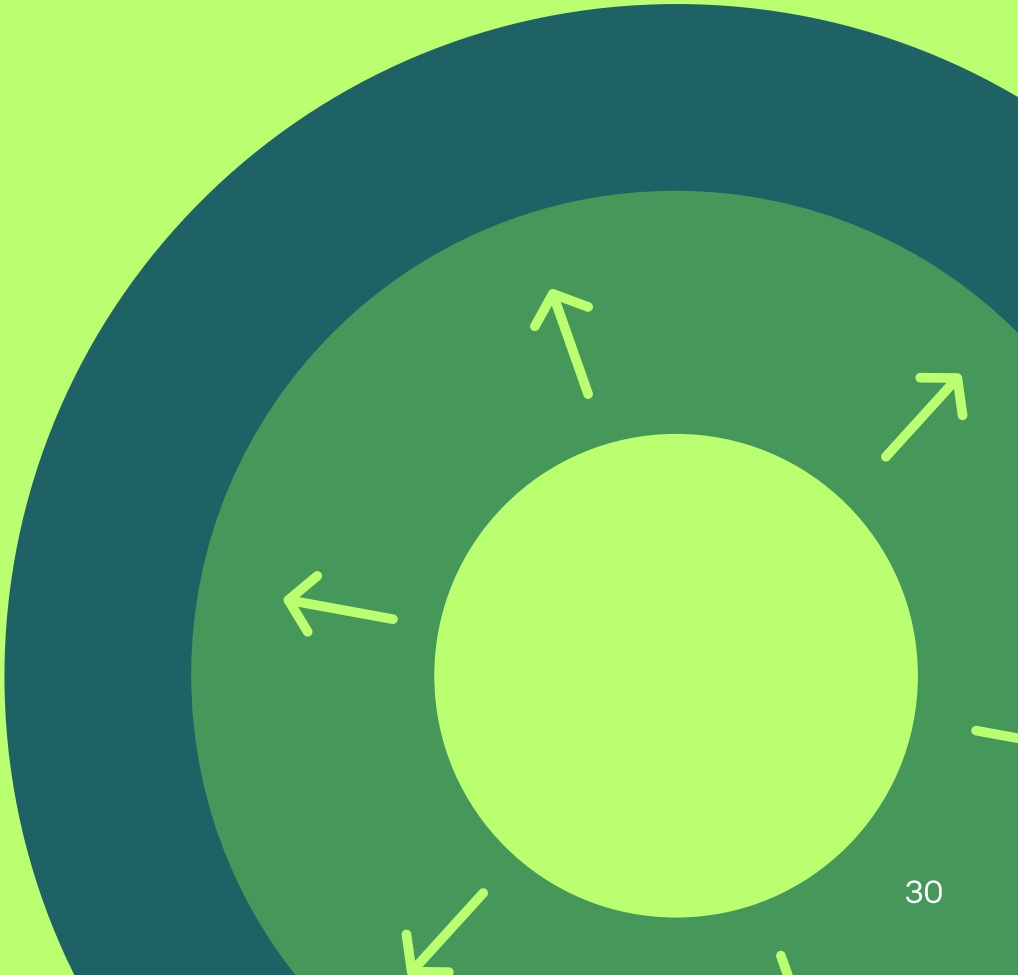
We're Leading Behavioural Science + AI for eCommerce Globally

Nudging for effective AI adoption

Use behavioral science principles to help organizations integrate AI into existing workflows and adopt new ways of working. This can help overcome natural resistance to change and ensure you realize the return on your investment in technology.

Beyond the Hype: Putting People at the Centre of AI Adoption
Samuel Keightley - BeHive

AI for Perspective, Not Accuracy: Bringing Behavioural Science to Compliance
Christian Hunt - Human Risk



Beyond the Hype: Putting People at the Centre of AI Adoption

AI-adoption in organizations / System implementation in organizations

Summary

This piece emphasises the human element in AI adoption. It advocates for an approach rooted in behavioral science to ensure new tools fit the systems they are meant to serve. This is explored through the ADOPT framework, which outlines five key pillars for successful AI integration as well as the essential steps leaders can take to make it work for their business.

Samuel Keightley
Senior Behavioural Science Researcher at BeHive

The Context

AI is set to be one of the most disruptive innovations of the 21st century, with what seems like daily developments in new use cases and innovations, each often leaving us trying to assimilate to what is possible and what may be on our horizon. The constant stream of AI developments has created a widespread, and seemingly unavoidable need to brace for change, often accompanied with a pressing drive to adapt and adopt. Nowhere is this sentiment felt more strongly than in the organisational domain. With the promise of cost cutting efficiency and ability to solve historical problems, a figurative AI gold rush has begun, with companies clambering to implement and innovate with these shiny new tools in any way they can.

With such a strong narrative of the value potential of AI tool implementation, it comes to no surprise that, here in 2025, we see around 78% of organisations attempting to implement the tools into at least one aspect of their business. Yet around 80% of these efforts fail to deliver the intended impact, or worse, create new problems that leave organisations paying the price (McKinsey, 2025).

The Problem

Distracted by the sheer promise of AI's value potential, decision makers often start by asking questions like **'What can the technology do?'**

While valid, this question becomes problematic when technological capability becomes the focal point of adoption. Often driven by competitive pressure, a tech-first, top-down mindset risks overlooking how tools will fit into the human systems they are meant to serve.

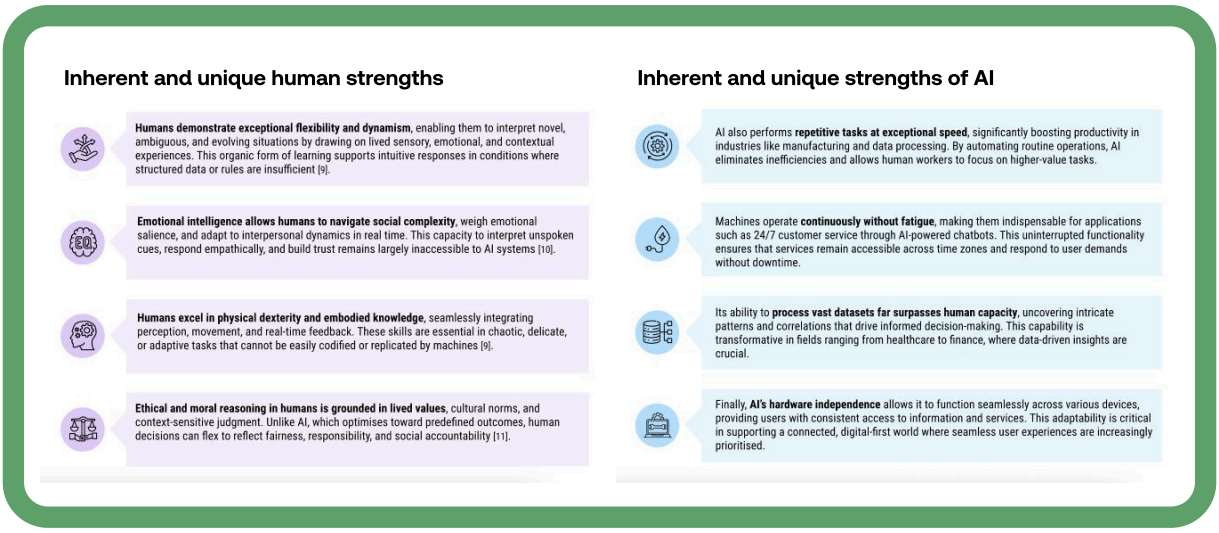
These dynamics create frictions that block adoption, where employees may feel uncertain about their role or threatened by automation, leaders underestimate how much cultural attitudes shape uptake, and rigid processes make integration feel disruptive rather than supportive. In practice this means that what begins as enthusiasm at the top often meets resistance at the frontline, with people reverting to familiar routines, bypassing tools they do not trust, or building manual workarounds that undermine intended impact.

While introducing new and innovative tools may initially create a surge of interest, it is often the systemic hurdles that prevent the conditions necessary for workers to meaningfully integrate these tools into their daily routines, ultimately feeding a cycle of repeated adoption attempts with consistently low success rates.

The Solution

By taking on a grounded and problem focussed approach, the true value of AI can be purposefully placed and productively integrated. This often starts with asking questions such as: 'What problems do we have that these new technologies may be able to meaningfully and ethically solve?' and 'Where does this tech fit into our existing human-led system?'. By guiding implementation strategies with bottom up perspectives, we can ensure that the technology introduced is matched to genuine needs, and also considers the people, processes, and relationships it will eventually impact.

Successfully integrating new tech tools first requires an understanding of the unique strengths and weaknesses of your workforce as well as the prospective AI tools you feasibly have available to you. This knowledge can direct an intentional and strategic effort to achieve tech-stakeholder collaboration and avoid unnecessary disruption and conflict.



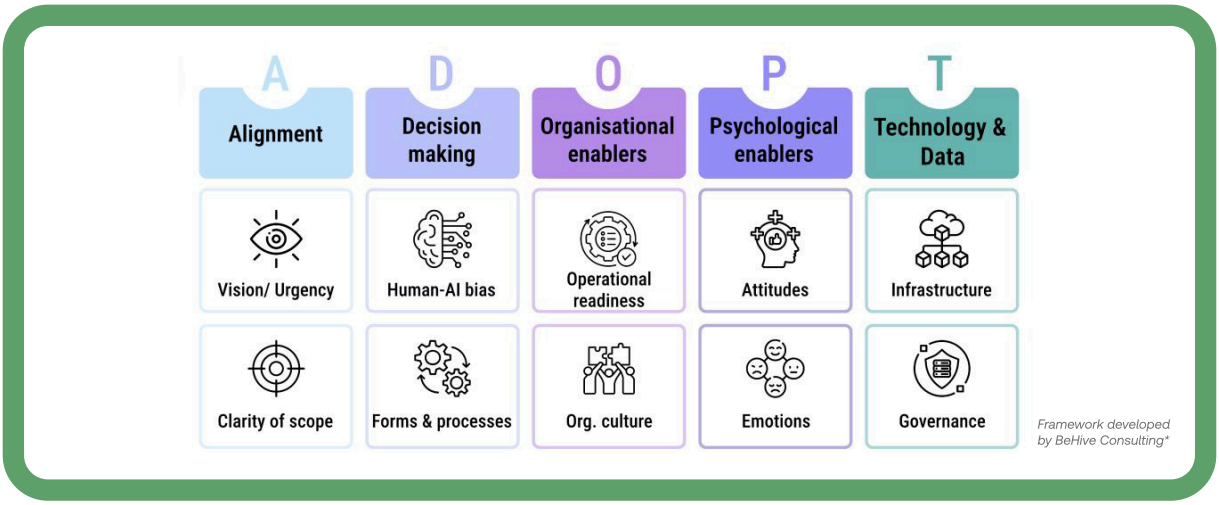
Seeing adoption as a human challenge is a vital shift, moving our attention from features to fit, prompting enquiry into how people will engage with the tool, what support they will need, and what obstacles they might face. It means planning for friction as carefully as function, and recognising that success is not only about making the tool work, but about ensuring people want to work with it.

Where Behavioural Science comes in

Since AI adoption at its core is a human challenge, it requires more than technical planning. The introduction of any new tool into an organisation sets off a chain of effects across complex social systems, influencing how people work, interact, make decisions, and respond to change. A psycho-social-behavioural approach is necessary to anticipate these ripple effects, understand the motivations and barriers that will shape adoption, and to design strategies that build user confidence, foster engagement, and embed adoption into daily practices.

To help organisations fully consider all the critical areas for successful and optimal tool adoption, BeHive has developed the AI adoption framework (ADOPT). Drawing on recent empirical research reviews and hands-on experience, ADOPT maps out the key psychological, social, behavioural, and organisational factors that determine whether AI will be implemented effectively and sustained over time (Murire, 2024; Bankins et al., 2024; Chatterjee et al., 2021; Kelly et al., 2023; Neumann et al., 2024). It highlights what may have been missed, where gaps exist, and how ready a team or company really is to integrate AI in a way that works.

Introducing the ADOPT framework



Alignment: How well AI adoption anchors in a clear strategy and shared direction, connecting its purpose to broader business goals.

- **Vision & urgency:** Shared understanding of why AI is adopted, its link to organisational goals and the rationale for targeted action.
- **Clarity of scope:** Understanding of what AI does, where it applies, and how success will be evaluated, setting feasibility and expectation boundaries.

Decision making: The dynamic interface between human reasoning and AI input, including formal structures and real-time interaction.

- **Human-AI bias:** Recognising distinct human/AI strengths and vulnerabilities in order to prevent bias amplification and promote complementarity.
- **Forms & processes:** Formal mechanisms for AI decision-making, defining rights, review, accountability, and escalation.

Organisational enablers: Internal capacities and conditions allowing AI initiatives to move from intention to action.

- **Operational readiness:** Organisation's ability to effectively absorb and deliver AI-driven change, encompassing capabilities, resources, and leadership support.
- **Organisational culture:** Prevailing norms, values, and beliefs shaping the approach to change and technology, influencing ongoing AI perceptions.

Psychological enablers: Individual beliefs, emotions, and mental models that shape engagement with AI (e.g., trust, confidence, relevance).

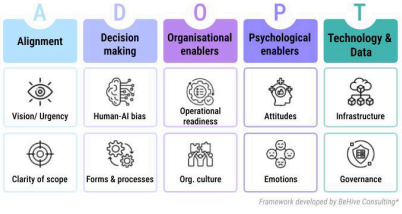
- **Attitudes:** An individual's beliefs, expectations, and evaluations of AI (e.g., usefulness, trustworthiness, threat).
- **Emotions:** Emotional responses to AI, such as trust, anxiety, or confidence, influencing daily tool engagement.

Technology & data: Technical infrastructure, system design, and data foundations that enable AI to scale and function responsibly.

- **Infrastructure:** Technical backbone supporting AI development and deployment, including data pipelines, computing, and system integration.
- **Governance:** Technical and operational controls ensuring responsible AI system building and maintenance, managing risks and ensuring compliance.

Why is this important?

If you take into account all elements...



...you can move towards:

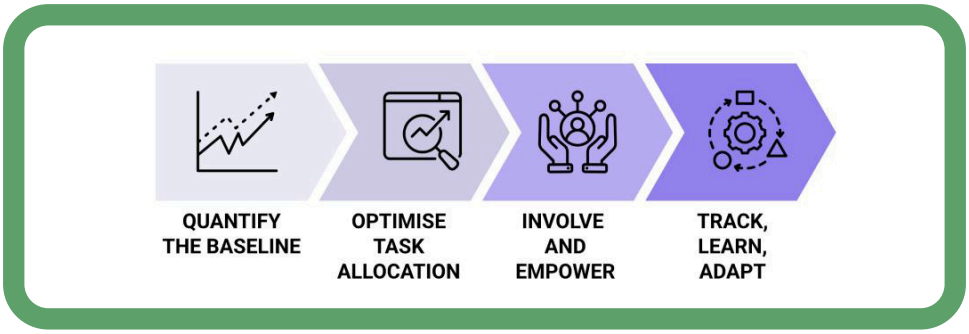
Precision Implementation — Move from generic AI rollouts to **tailored adoption strategies** based on user needs, fears, and motivations.

Behavioural Design — Create environments and processes that make AI **adoption intuitive, rewarding, and natural**.

System-Level Adoption — **Align** AI with **team dynamics and organizational culture** to drive lasting engagement.

Human-Centred AI — Design AI experiences that **enhance** workflows, reduce friction, and support sustainable, healthy **work practices**.

So what essential steps can be taken to ensure a human centric adoption journey?



Leaders have a critical role to play in making AI work for their organisations. That starts by treating adoption not as a technical deployment, but as a behavioural challenge. It's not enough to install the tool. You have to install the conditions for people to use it well.

In order to lay the foundations for a dynamic and meaningful (and ethical) implementation of AI, there are essential steps leaders can take to support both the organisation and its people in solving the right problems, in the right way, with the right technology:

- **Quantify the baseline:** Map how work currently gets done. Understand the starting point across roles, routines, needs, and gaps.
- **Optimise task allocation.** Decide where AI should support, not displace. Be clear on which tasks need a human touch and which do not.
- **Involve and empower.** Engage employees early. Let them shape the rollout, test solutions, and see how their input matters.
- **Track, learn, adapt.** Treat adoption as an ongoing process. Build feedback loops, measure what matters, and adjust when things don't land.

So how do we responsibly shape adoption that lasts?

With the mass uptake of AI tools exposing both their potential for positive innovation and their risks of negative disruption, it is clear that human centricity holds the key to successful and ethical integration. Decisions to adopt must weigh not only whether AI delivers efficiencies but also how it reshapes people's daily work and wellbeing. That requires a clear-eyed understanding of the challenges adoption brings, the risks it may create, and the responsibility to safeguard those who interact with these emerging tools. This phase of progress demands interdisciplinary expertise that can connect the world of work with an understanding of human psychology and behaviour, positioning behavioural scientists as an invaluable resource to guide and optimise this change. Without this perspective, adoption risks repeating the same cycle of high expectations and failed outcomes that has already left so many organisations paying the price. But with it, AI can move beyond surface-level implementation to become a source of ethical, integrated, and productive progress.

References:

Bankins, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. Journal of Organizational Behavior, 45(2), 159–182. https://doi.org/10.1002/job.2735

Chatterjee, S., Rana, N. P., Dwivedi, Y. K., & Baabdullah, A. M. (2021). Understanding AI adoption in manufacturing and production firms using an integrated TAM-TOE model. Technological Forecasting and Social Change, 170, 120880.

Kelly, S., Kaye, S. A., & Oviedo-Trespalacios, O. (2023). What factors contribute to the acceptance of artificial intelligence? A systematic review. Telematics and informatics, 77, 101925.

McKinsey. (2025, March 12). The state of AI: How organizations are rewiring to capture value. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

Murire, O. T. (2024). Artificial Intelligence and Its Role in Shaping Organizational Work Practices and Culture. Administrative Sciences, 14(12), 316. https://doi.org/10.3390/admsci14120316

Neumann, O., Guirguis, K., & Steiner, R. (2024). Exploring artificial intelligence adoption in public organizations: a comparative case study. Public Management Review, 26(1), 114-141.

Want more? Check these reads out.

When Machines Teach Empathy: How AI Amplifies Behavioral Science in Business Communications

Thinking clearly about behavioural science and AI: A guide for the perplexed

We're Leading Behavioural Science + AI for eCommerce Globally

AI for Perspective, Not Accuracy: Bringing Behavioural Science to Compliance

Deeper insights in compliance / AI-adoption in organizations

Christian Hunt
Founder & CEO at Human Risk

Summary

AI's biggest potential in Compliance is not efficiency or accuracy but in providing perspective that human minds can use as a springboard for deeper insight. By creating custom AI roles, this approach uncovers blind spots and flawed assumptions. It results in a provocation that helps humans overcome biases and provide a richer analysis.

Ask most people about AI in Compliance, and you'll hear about efficiency and accuracy: chatbots, automated reporting, data crunching. Useful, but like the humans it's replacing, it risks being fragile. A chatbot that gives the wrong advice, or a sanctions tool that clears a prohibited transaction, doesn't just fail; it creates risk.

Rather than enter what is already a crowded space, my focus is different. I don't use AI to replace humans or do what they can't. I use it to enhance human capability. Instead of definitive answers, I design behaviourally tuned models that improve decision-making by reducing bias and blind spots. In smaller functions, or those with wide geographic or cultural reach, this also provides something hard to replicate: diverse perspectives.

In my work, I start from a simple principle: Compliance is the business of influencing human decision-making. Programmes often fail because they are designed for homo compliantcus — how we wish people behaved — not how they really do. Too often, they ignore behavioural science, falling into the curse of knowledge (assuming others understand as much as we do) and the curse of passion (assuming they care as much as we do). This spawned the idea of using AI not for accuracy, but for perspective: helping Compliance Officers recognise flaws in their control frameworks, surfacing how non-compliance might occur — wilful or accidental — and exposing assumptions they didn't know they were making.

Methodology

I feed a scenario into AI. This might be a new policy, a sales incentive, or a description of something that went wrong, and I ask it to play it through from different perspectives. Borrowing from Edward de Bono's Thinking Hats, I create a series of custom GPTs that play a cast of characters:

- **Rule-Maker** — creates or interprets rules as a regulator would.
- **Rule-Breaker** — imagines loopholes and rationalisations.
- **Faker** — games the rules without technically breaking them.

These roles are illustrative; others can be defined depending on the context. The key is separation. If one model plays all roles, its later answers are shaped by its earlier ones — effectively marking its own homework. By keeping roles distinct, each brings a genuinely different perspective, creating cognitive diversity. The value lies in how you configure the process: selecting data, designing prompts, and knowing which kinds of reasoning to encourage. The outputs are not predictions of what will happen, but provocations of what might happen. Even absurd scenarios are useful because they stimulate debate and force reflection.

Applications

I've deployed this approach with clients in several areas:

- **Conflict of Interest policy.** When relaunching a policy, the Rule-Breaker surfaced rationalisations such as: "My spouse's business isn't formally in their name, so this isn't caught." Wrong in law, but plausible enough to tighten wording and strengthen training. This helps overcome the curse of knowledge — assuming we've covered everything, while missing scenarios obvious to others.
- **Learning from failed interventions.** Where a breach has already occurred, it's easy to fix the immediate loophole and stop there. AI pushes further, surfacing what else could have gone wrong in similar circumstances. This broadens lessons learned and avoids the bias of focusing too narrowly on a single failure.
- **Sales incentives.** Role-playing as salespeople, the models suggested over-promising benefits or outsourcing "grey area" activity. This showed how incentives could be gamed — with ethical and regulatory consequences. It challenges the assumption that a policy will hold, even against imaginative employees.
- **Third-party onboarding.** Acting as suppliers, the AI tested the process as both a 'large' and 'small' supplier, revealing where processes were duplicative or overly bureaucratic, and helping insiders to see things from a supplier's point of view.
- **Pre-mortems.** As part of war-gaming, AI generated "future headlines" describing compliance breakdowns, then mapped backwards how they might occur. Even unlikely headlines surfaced blind spots and forced leaders to challenge assumptions. This helps overcome the bias of unwillingness to think the unthinkable.

- **Induction training.** Induction is often designed from the organisation's perspective, not the employee's. AI flagged risks like information overload, unnecessary statements (e.g. "we are a very ethical organisation"), or missed employee concerns. This helps stage training more effectively and tackles the bias of assuming newcomers want what the organisation wants to say.

In every case, the AI doesn't provide "the answer." It provides provocations that triggers richer human analysis; that's where I think the real value lies.

Adoption

I sometimes describe this as using behavioural science to sell behavioural science. Compliance Officers adopt it readily because it makes their jobs more engaging: instead of box-ticking, they explore loopholes, test culture, and sharpen judgment.

For organisations, the appeal lies in speed and safety. Pilots can be run quickly with existing policies and data. And because the goal is perspective, not accuracy, the risk lies in how humans use the insights, not in the machine. Even unlikely scenarios surface assumptions and point towards Donald Rumsfeld's "unknown unknowns", possibilities that might otherwise never be considered.

Beyond Compliance

Although rooted in compliance, the same method applies to operational risk, ethics, and wider human risk management, anywhere behaviour interacts with rules and incentives.
AI provides the provocations.

Behavioural science makes them humanly plausible. Human expertise decides what to do with them. That's why, for me, the future of AI in compliance isn't accuracy; it's perspective.

Want more? Check these reads out.

Beyond the Hype: Putting People at the Centre of AI Adoption

Computed irrationality: A new frontier for Behavioural Science and AI

How AmosNL is solving challenges faced by behavioral science consultancies

Introducing Diversifi

Diversifi is a global, cross-cultural organisation that brings together world leaders in applied behavioural science to help solve some of the world's biggest problems. By bringing together a wide spectrum of applied behavioural science capabilities and skill sets, integrated teams with Diversifi create end to end solutions for organisations and communities anywhere in the world.

More details about the team and Diversifi can be found at: **diversifiglobal.com**. If you would like to connect with the team or be involved in any future endeavours, please contact us at: **jezgroom@cowryconsulting.com** or **prakash@1001stories.in**

Our contributors

To further discover the synergetic effects of integrating behavioural science with AI, please feel free to get in touch with our contributors.

Anna Malena Njå Behavioural Consultant at **Nudgelab**
anna@thenudgelab.no

Christian Hunt Founder of **Human Risk**
christian@human-risk.com

Elina Halonen Founder and Strategist at **Prismatic Strategy**
elina@squarepeginsight.com

Jez Groom Founder & International CEO at **Cowry**
jezgroom@cowryconsulting.com

Lisa Bladh Behavioural Designer at **Cowry**
lisabladh@cowryconsulting.com

Samuel Keightley Senior Behavioural Science Researcher at **BeHive**
samuel.keightley@behive.consulting

Sonia Friedrich Founder of **Sonia Friedrich Consulting**
sonia@soniafriedrich.com

Roger Dooley Author, Keynote Speaker, Neuromarketer
roger@rogerdooley.com

Explore with keywords

Explore more writing on topics that peak your interest.

Learning to understand AI

When Machines Teach Empathy: How AI Amplifies Behavioral Science in Business Communications

Computed irrationality: A new frontier for Behavioural Science and AI

AI-adoption in organizations

Beyond the Hype: Putting People at the Centre of AI Adoption

AI for Perspective, Not Accuracy: Bringing Behavioural Science to Compliance

Increased profitability in eCommerce

We're Leading Behavioural Science + AI for eCommerce Globally

Improving AI-systems

Computed irrationality: A new frontier for Behavioural Science and AI

Thinking clearly about behavioural science and AI: A guide for the perplexed

Scaling interventions

Thinking clearly about behavioural science and AI: A guide for the perplexed

How AmosNL is solving challenges faced by behavioral science consultancies

Predictive AI for behavior change

How AmosNL is solving challenges faced by behavioral science consultancies

Deeper insights in compliance

AI for Perspective, Not Accuracy: Bringing Behavioural Science to Compliance

Personalized messaging

We're Leading Behavioural Science + AI for eCommerce Globally

Thinking clearly about behavioural science and AI: A guide for the perplexed

When Machines Teach Empathy: How AI Amplifies Behavioral Science in Business Communications

Diversifi Compendium 2025

Edited by Anna Malena Njå & Lisa Bladh