The Centaur and the Minotaur

Design, artificial intelligence and the new frontier of creative decision-making.

A Design Council Discussion Paper

Ben Ramalingam, Mat Hunter and Cat Drew

Key messages

Message 1 Design and AI are now deeply intertwined in all areas of practice. But who is in charge in the relationship between human judgment and machine outputs? The evidence suggests that most organisations are not asking this question deliberately, systematically or frequently enough.

Message 2 Two figures from Greek mythology capture the diverging ways of working. In the centaur, the human head leads and the machine body provides power. In the minotaur, the logic is reversed: the machine frames the problem, generates the options, and embeds the assumptions, while humans respond to what it presents. The difference between the two hinges on how design processes and capabilities are structured around AI.

Message 3 Research shows that human-AI design teams work best when humans lead. Zaha Hadid Architects demonstrate this in practice: they select a portion of cent of AI output to take forward, maintaining deliberate human control over what the machine produces.

Message 4 Despite the evidence, in all areas of design there is an observable and largely unmanaged drift toward the minotaur model. Experiments show that when AI is in the lead, it narrows rather than expands the creative solution space, while giving the appearance of expansive exploration. Moreover, the more fluent and comprehensive the AI output, the less critically humans engage with it.

Message 5 The mechanism underlying this drift is automation bias: the tendency to follow AI recommendations even when they contradict available evidence or one's own judgment. Research shows this is self-reinforcing. The more we defer to AI output, the less practised we become at questioning it.

Message 6 The consequences of automation bias are already apparent. Accountability is challenged by design decisions that are becoming impossible to attribute. The capability pipeline is hollowing out as junior designers lose the formative work through which design judgment develops. Homogenisation is spreading as centralised training data displaces the abductive reasoning on which genuinely new design depends. And trust, which arrives on foot and leaves on horseback, is eroding among clients, users, and the profession itself.

Message 7 Four practical actions follow from this analysis. The first is sequencing: human judgment should precede AI output, not respond to it. The second is making the machine's assumptions visible and contestable, so that designers engage critically with what the AI has already decided before they see the results. The third is protecting the capability pipeline through mentorship, structured critique and apprenticeship rather than AI training alone. The fourth, and perhaps the most important, is applying design principles to the decision architecture of the whole process, not just the interface of individual tools.

Message 8 The stakes extend well beyond design organisations. Businesses, institutions and governments are harnessing AI tools in ways that will make challenges of automation bias more prominent. Design, properly applied, is well placed to make AI more intentional, more accountable, and more attentive to the people it affects.

INTRODUCTION

An architecture studio in London. An in-house AI tool has generated two hundred facade variations for a competition entry. Each is optimised against solar gain, structural performance, material efficiency, and visual impact, while responding to planning requirements.

The junior architect's instinct tells her the scheme needs to respond to things the model cannot see or replicate: the way the square empties after school pick-up and fills again at dusk, elderly residents on the adjacent estate who shared their memories of the building it replaces, and the soft sheen on the nearby brickwork after rainfall.

But the project team has already built a shortlisting process around the AI-generated options: comprehensive, polished, structured. The design review is tomorrow. Introducing considerations the model hasn't surfaced would mean going back upstream - to the model's input parameters, constraints and selection criteria - and unwinding work the team believes is already done.

Who is in charge here?

A design research team working on a large-scale public sector service redesign. The AI platform has processed every user interview and distilled the findings into weighted thematic clusters with confidence scores.

The lead researcher has undertaken his own painstaking synthesis of the same data, but the project sponsor has already received the AI summary and is asking why the two accounts diverge. The researcher has two decades of contextual, sector-specific knowledge that no dataset contains - but raising this against the AI output now feels like he is being defensive.

Who is in charge here?

A fashion house in Milan. An AI agent has analysed multiple seasons of sales data, social media sentiment, and runway imagery to propose a capsule collection. The prints are striking, the colour palette is on-trend, and the creative director can see the commercial logic. But when the collection launches, consumers call it 'soulless.' The brand team cannot explain why, because the design process that would have told them - the sketching, the draping, the trade-off between formal evidence and design instinct - was replaced by a series of prompts.

Who is in charge here?

These are not thought experiments. Each is drawn from real situations encountered over the last three years. Together, they reveal a pattern increasingly visible across all design-intensive sectors: the transfer of creative and strategic agency from humans to machines,[1] often without anyone consciously choosing to do so.

This pattern has roots in a problem that predates generative AI, and will be familiar to anyone working in design: the tendency of design process to become performative. IDEO's Michael Hendrix famously called this the 'theatre of innovation.' [2]

> *"We get a lot of the materials that look like… they make us more creative. That could be anything from getting a bunch of Sharpie markers and Post-its and putting them in rooms for brainstorms... but without some kind of history or strategy to tie them together, and track their progress, track their impact, they end up being a theatrical thing that people can point to and say, 'oh we did that.'"*

[1]Computers in Human Behavior: 'Human vs. AI: The impact of AI assistance on the decision-making process of knowledge workers.' https://www.sciencedirect.com/science/article/pii/S0747563222001303

[2]Fast Company: 'IDEO breaks its silence on design thinking's critics.' https://www.fastcompany.com/90257718/ideo-breaks-its-silence-on-design-thinkings-critics

Research supports Hendrix's claims.[3] Researchers from the QUT School of Design found that many of the positive qualities of design thinking have been lost to a standardised, one-size-fits-all set of methods applied at speed across diverse socio-cultural contexts: the journey maps, the persona cards, the ideation sprints, the carefully assembled decks. These all work to create a sense of human-centredness without asking the hardest questions: Whose needs and interests are actually being centred? What assumptions are we protecting? What would genuinely challenge our thinking?

The 'prolific impact' of Generative AI on creative industries[4] is amplifying this problem rapidly. It can be used to produce a convincing design rationale, a plausible research synthesis, and a set of polished prototypes within hours. But there are downsides. The 2025 State of AI in Design report,[5] based on a survey of over 400 designers, warned that as AI adoption spreads, the market risks flooding with 'AI slop': designs that feel generic and repetitive, where distinctive visual identities fade and brands struggle to stand out.

The defence of such practices stems from commercial software development: launch to learn, iterate through market feedback and let selection pressures do the work. But in software, iteration is cheap and patches are possible. You cannot patch a failed collection or rebuild a soulless building after the fact.

The tension between adoption and quality is captured in Figma's 2025 AI report: of over 2,500 designers surveyed, almost 4 out of every 5 said AI enhanced their efficiency, but less than a third of these said they could actually rely on AI outputs.[6]

To reiterate, this is not a problem created by AI. It is a problem created by weak design processes which AI now exposes and amplifies, faster, more cheaply, and in increasingly convincing ways, and with pernicious effects.

This is also not a call to keep AI at arm's length. For well-defined, optimisation-led tasks, and for much important work around design processes such as bid writing, specification drafting and checking against regulations, AI is already proving incredibly useful. The concern of this paper is narrower and sharper: the creative and strategic decisions that determine what a design is for, whom it serves, and what it means.

THE CENTAUR AND THE MINOTAUR

A useful way to understand this problem draws on two contrasting figures from Greek mythology. In a centaur, a human head and torso sit atop the body of a horse. In a minotaur, the logic is reversed: a bull's head on a human body.

The metaphor captures something precise about how AI enters design decision-making.

In the centaur model, designers frame the problem, generate the options, and use AI for supporting analysis.

In the minotaur model, the machine frames the problem, generates the options, and embeds the assumptions, from which human users then make their choices[7].

---

[3]Design Studies: 'Tokenistic co-optation: Design thinking's capture by management consultancy.' QUT School of Design. https://www.sciencedirect.com/science/article/pii/S0142694X25000729

[4] Carey, H., Giles, L. and Hay, B. (2026) *Creative Industries Skills Audits: Summary Report*. Creative Industries Policy and Evidence Centre (Creative PEC). DOI: 10.5281/zenodo.20072107. Available at: https://pec.ac.uk/wp-content/uploads/2026/05/Creative-Industries-Skills-Audits-Summary-Report.pdf

[5]State of AI in Design report. https://www.stateofaidesign.com/

[6]Figma: 'Figma's 2025 AI report: Perspectives from designers and developers.' https://www.figma.com/blog/figma-2025-ai-report-perspectives/

[7]The minotaur framing was developed independently but shares logic with Cory Doctorow's parallel concept of the "reverse centaur", which describes humans reduced to an appendage of a machine that sets the pace and the terms of work. While Doctorow's concern is chiefly labour and surveillance, the focus here is the design process: how problem framing, option

In practice there is a spectrum between these poles. However, evidence from design - and also from healthcare, law, and public administration[8] – all points in the same direction.

The minotaur configuration is far more likely to produce *automation bias*: the tendency to follow AI recommendations even when they contradict available evidence or one's own informed judgement.

A systematic review of 35 studies found that automation bias has become a critical challenge across high-stakes fields. The more AI is used, the more we see a process of cognitive offloading: the gradual reduction of mental effort as analytical tasks are delegated to machines. This is reinforced by a growing trend in professional cultures to treat AI outputs as authoritative, making human oversight and judgement feel unnecessary.

The centaur configuration, by contrast, positions the human as the primary decision-maker and AI as a secondary supporting system. Research from Harvard's Data Science Review has shown that centaur models outperform both the best algorithm and the best human experts.[9]

A comparative study of human-AI design co-creation across different levels of experienced designers found that when AI was embedded within a designer-driven process - rather than leading it - creative performance improved significantly.[10]

Experienced designers maintained control, using AI selectively for refinement rather than ideation. The question, as one participant put it, was not whether AI could generate design solutions, but *whether the designer remained in charge of asking the right questions*.

THE CENTAUR IN PRACTICE: ZAHA HADID ARCHITECTS

Zaha Hadid Architects (ZHA) were among the earliest architecture studios to integrate generative AI deliberately into their practice. At the outset, the studio made a deliberate choice about the division of labour.[11] ZHA architects define the brief, set the parameters and constraints, evaluate the output, and then develop the selected concepts through traditional modelling methods. They use a range of AI tools - DALL-E 2, Midjourney, and Stable Diffusion - to generate concept images, with prompts combining 'Zaha Hadid' with programmatic and spatial attributes. The architects apply considerations such as better transit, more daylight, and lower energy use to shape the design direction.

But the studio selects only around 10 to 15 per cent of the AI output to take forward to the 3D modelling phase. The rest is discarded.

ZHA has reported that AI has increased productivity by up to 50 per cent during mid-stage design, and that in competitive bidding the firm can now produce double or triple the number of high-quality proposals as previously.[12]

---

generation and assumption-setting drift to the machine, and what this does to creative judgement. See Doctorow, C. (2026) The Reverse Centaur's Guide to Life After AI: How to Think About Artificial Intelligence Before It's Too Late. London: Verso.

[8]AI & Society: 'Automation bias in human-AI decision making: A systematic review.' https://link.springer.com/article/10.1007/s00146-025-02422-7

[9]Harvard Data Science Review: 'Centaur models: Human-algorithm collaboration for decision making under uncertainty.' https://hdsr.mitpress.mit.edu/pub/3rvlzjtw

[10]Frontiers in Computer Science: 'Human-AI co-creation across design experience levels.' https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2025.1672735/full

[11]Dezeen: 'Zaha Hadid Architects uses AI tools including DALL-E 2 and Midjourney.' https://www.dezeen.com/2023/04/26/zaha-hadid-architects-patrik-schumacher-ai-dalle-midjourney/

[12]Design Asia Magazine: 'Zaha Hadid Architects builds winner proposals with AI.' https://designasiamagazine.com/zaha-hadid-architects-builds-winner-proposals-with-ai/

As Shajay Bhooshan who leads ZHA's computational design group explains, the AI models offer 'an insight into how we can make decisions'[13].

Humans set direction and scope; AI extends reach at a considerably accelerated pace.

This is the centaur in action: the human head leads, the machine body provides power the humans could not achieve alone. The division of labour is explicit, deliberate, and maintained through conscious discipline - including transparency with clients about how AI is used.

THE MINOTAUR IN PRACTICE: DESIGN FIXATION

Compelling evidence for the minotaur in design comes from a controlled experiment presented in 2024 at ACM's CHI - the leading conference in human-computer interaction - that directly tested the claim that generative AI augments human creativity.[14]

The conclusion is that it does not.

Specifically, they found that designers who used AI-generated images during ideation led to higher fixation on an initial example.

Participants who used AI produced fewer ideas, with less variety and lower originality compared to a non-AI design baseline. The AI did not expand the solution space. It contracted it - while creating the appearance of expansive exploration.

The researchers identified two mechanisms through which this happened.

First, the act of creating prompts anchored participants on particular features of the initial example, narrowing subsequent thinking.

Second, the AI's outputs themselves became sources of fixation: participants treated the machine's suggestions as the boundaries of the possible, rather than as one input among many.

The study describes this as *fixation displacement* - the focus of fixation shifts from the original example onto the AI's outputs. The frame of what could be explored had contracted around what was immediately presented.

This is the minotaur in empirical form. The machine frames the problem, generates the options, and embeds the assumptions. The designer responds to what the AI presents. The process looks generative - dozens of variations, rapid iteration, visual richness – but all the while, the conceptual space is actually narrowing.

Other evidence confirms that generative AI exhibits design fixation as an inherent property of the models. The tendency toward homogenised design perspectives is driven by centralised training data, and is therefore a systemic problem, not an incidental one.[15]

The consequences are already visible in various areas of commercial practice. Fashion brands such as Collina Strada and BAGGU faced consumer backlash after releasing AI-generated prints that audiences called 'soulless.'[16]

[13]Wallpaper*: 'Zaha Hadid Architects and Nvidia team up to explore AI in architecture.' https://www.wallpaper.com/architecture/zaha-hadid-architects-nvidia-ai-in-architecture

[14]Proceedings of CHI '24: 'Beyond the Creativity Barrier: How Generative AI Constrains Creative Exploration.' University of Melbourne. https://dl.acm.org/doi/full/10.1145/3613904.3642919

[15]arXiv: 'Design Fixation in Generative AI.' https://arxiv.org/html/2502.05870v1

[16] Business of Fashion: 'Fashion's AI Dilemma Is Getting Worse' https://www.businessoffashion.com/articles/technology/fashion-ai-dilemma-is-getting-worse/

A 2023 study pre-empted the drama, finding that consumers generally do not see AI-generated fashion as authentic or high quality.[17]

However, these perceptions shifted when consumers were given the option to feed into the AI design process. A subsequent study across a range of settings found that human-AI co-creation in luxury fashion generated consistently more positive consumer responses than AI-only design. This effect is explained by the customers perception that human involvement signals genuine creative effort.[18]

In other words, the minotaur does not only narrow the design space internally, but these limited outputs are visible to the people design is supposed to serve – and they are voting with their feet.

A comparative review across designers of different experience levels gives more detail.[19] Experienced designers naturally - and perhaps predictably - maintain centaur configurations: they use AI selectively, discard outputs that interfered with their decision-making, and treated it as an instrument rather than an authority. As one participant put it, 'I actually discarded a lot of the AI-generated content because it was overly ornate - it interfered with my later-stage decision-making.'

Novice designers behaved quite differently. They were more inclined to accept AI outputs during the early stages where problem framing happens - precisely the stages where the fixation study showed AI narrows rather than expands the solution space. This may be as much a generational shift as an experiential one, among designers who have grown up in a faster, more digital environment where process and judgement have developed differently.

Perhaps most discomforting is that the designers most likely to defer to machines are those who most need the depth of understanding that comes from doing the work themselves.

THE DRIFT TOWARD AUTOMATION BIAS

Perhaps what is most concerning from all of this is the current direction of travel. Centaurs are achievable - and the evidence shows they work - but they require deliberate design, sustained discipline, and resistance to the pull of the machine. Short of this conscious effort, what we are seeing in many design organisations is a drift towards automation bias in human-AI interactions. A systematic review of AI-enabled design tools published last year identified a recurring pattern of creeping overreliance on automation and technology-first development[20].

This drift - from centaur to minotaur - rarely happens through explicit choices.

It happens through small acts of convenience and deference, each one shifting the structure of decision-making.

Why might this be the case?

Consider the design research team described in the opening scenario. AI was used to process and synthesise user voices across geographies and programme areas - a genuine analytical achievement. The AI could identify what was happening and where: patterns across regions, shifts in language, changes in emphasis over time. But it could not explain *why* these things were happening, how significant they were, or what they meant for design choices. Those questions required contextual judgement that only experienced practitioners could provide.

The problem emerged when AI outputs were presented to sponsors separately from the human analysis. Challenging the machine's framing required more energy and justification than accepting it. The burden of proof had

[17] Lee, G. & Kim, H.Y., 'Human vs. AI: The Battle for Authenticity in Fashion Design and Consumer Response', *Journal of Retailing and Consumer Services*, vol. 77, 2024. https://www.sciencedirect.com/science/article/abs/pii/S0969698923004411

[18] Wang et al., 'How Does Human-AI Co-Creation Design Help Luxury Brands Win Consumers' Favor?', *Journal of Consumer Behaviour*, 2025. https://onlinelibrary.wiley.com/doi/10.1002/cb.70065

[19] Frontiers in Computer Science (ibid)

[20] Luo, Y. et al., 'Designing With AI: A Systematic Literature Review on the Use, Development, and Perception of AI-Enabled UX Design Tools', *Advances in Human-Computer Interaction*, 2025. https://onlinelibrary.wiley.com/doi/abs/10.1155/ahci/3869207

shifted. Experienced researchers found themselves arguing against a fluent, comprehensive analysis - and the conversations felt like defensiveness rather than rigour.

In this case, the drift was caught and corrected: the process was re-sequenced so that human judgement preceded AI synthesis rather than responded to it. But the correction required conscious effort. The machine's outputs were so clear, comprehensive, and professionally assembled that they exerted a kind of authority that overrode any explicit judgement about whether they were actually correct.

The drift toward minotaur configurations is increasingly well-documented - and the research suggests it may be self-reinforcing. A controlled experiment in interface design tested six different ways of interrupting automatic deference to AI models - from textual explanations to confidence scores to AI-generated reflective questions[21]. Remarkably, when the AI was wrong, over half of participants followed the incorrect recommendations anyway. But asking them to think more critically about the AI suggestions led to another unexpected problem: the designers became slower and less effective.

A Harvard field experiment found the same pattern at scale. Across over 3,000 screening decisions on early-stage innovations, the more fluent and well-reasoned the AI's recommendations, the less critically evaluators engaged with them. The authors called this the 'human-AI oversight paradox': the very quality that makes AI assistance useful - its ability to process what we find overwhelming - is what makes us stop questioning it.[22] One pith observation attributed to Karl Lagerfielsd is that clear thinking, at the wrong moment, can stifle design talent and inspiration. This may be even more true of clear, un-questioned thinking.

This is why the drift is so difficult to counteract at the level of individual decisions. Creativity, speed, accuracy, and critical judgement cannot all be maximised at once - every intervention that strengthens one weakens another. In professional settings with deadlines and commercial pressure, individual vigilance alone is not enough. What works - as ZHA demonstrate - is designing the process itself. Human creativity and judgement need to be institutionally protected, rather than relying on each person to resist the pull of the machine in the moment.

Nor is this drift limited to the interaction of AI and design. In many sectors, creativity struggles to hold its ground in the face of more confident, more fluent forms of analysis.

WHAT THIS MEANS FOR DESIGN: FOUR CONSEQUENCES

The consequences of leaving the problem of automation bias unaddressed are already visible across design professions.

Accountability. When AI-generated design artefacts pass through multiple rounds of review without anyone having made key conceptual design decisions embedded within them, responsibility becomes distributed to the point of disappearing.

The AI generated it; the designer reviewed it; the stakeholder approved it; but at no point did anyone actually decide what the design should do, for whom, and why.

Figma's data cited earlier captures the tension: 78% of respondents say AI enhances efficiency, but only 32 % say they can rely on the output.

There is also a second, less visible form of accountability drift. The market, it is argued, will decide: users will vote with their feet, poor designs will fail, better ones will survive. This is the invisible hand invoked as a substitute for design judgement, and it is becoming the default governance mechanism for AI-generated work across the profession.

---

[21]arXiv: 'Investigating human-AI collaboration on interface design: Interrupting automation bias through different mechanisms.' https://arxiv.org/abs/2501.16627

[22]Harvard Business School Working Paper: 'The Human-AI Oversight Paradox.' https://www.hbs.edu/ris/Publication%20Files/25-001_8ebbe0cb-2a19-453c-9014-1e301e8dd2fb.pdf

Design professionals know the stakes are rising and that practice is struggling to keep pace. The ethical questions these issues raise, about who is responsible for what a design does and to whom, remain largely unaddressed across the profession.

Capability. A 2025 study by Microsoft Research and Carnegie Mellon University found that knowledge workers using generative AI became more confident in their outputs while progressively abandoning critical problem-solving to the machine.[23]

When organisations automate the work traditionally done by junior designers -the research synthesis, the sketching, the first drafts, the painstaking assembly of evidence into design rationale - they hollow out the learning pipeline through which design judgement develops. As Jony Ive observed, opinions are not ideas[24]. Mistaking the confidence of AI-generated rationale for the development of design judgement risks losing something the profession may struggle to recover.

The effects are already visible in design, with Nielsen Norman Group's State of UX 2026 report painting a stark picture. Entry-level design positions are increasingly scarce. Available roles demand judgement rather than artefact production. Responsibilities once spread across multiple specialists are being compressed into single roles.

The UK Creative Industries Skills Audits identifies this same pattern of AI displacing entry-level roles and choking off the talent pipeline for mid-level positions across numerous sectors.

Strengthening the design knowledge that develops through tacit experience means more learning by doing: mentorship in design discernment and judgement, social learning and apprenticeships. The Skills Audit also points to a parallel need to build AI capability across the creative workforce, from everyday AI literacy to the legal and ethical knowledge to handle IP and copyright, and the specialised skills to deploy AI tools well. Much of this is currently acquired informally rather than through structured training.

Homogenisation. If every design team uses the same AI tools and defers to them in the same way, the result is convergent design. Research on AI-generated outputs finds that repeated reliance on the same models produces a homogenisation effect: outputs become more formulaic and less distinctive over time.[25]

The models' centralised, data-driven training produces a tendency toward homogenised design perspectives. A systematic review of AI in graphic design identifies stylistic homogenisation as an emerging concern across creative fields.[26]

When every studio's concept images, every brand's collections, and every firm's facade studies emerge from the same generative models, the diversity of thinking on which good design depends is lost.

What is also being lost is a specific kind of thinking, referred to by Lucy Kimbell as abductive reasoning[27]: the creative leap that generates genuinely new possibilities rather than extrapolating from what already exists. It is precisely what centralised training data cannot produce.

Trust. The old adage that trust arrives on foot and leaves on horseback is especially apparent in the rapidly evolving relationship between designers, clients, and AI.

Nielsen Norman Group's 2026 analysis identifies trust as a major design problem for AI experiences, noting that users who have encountered unreliable AI-designed features are more hesitant to adopt new ones. This is likely being amplified by the growing deployment of AI agents to generate hypothetical user responses as a substitute for

[23]Microsoft Research / Carnegie Mellon University: 'The Impact of Generative AI on Critical Thinking.' https://www.microsoft.com/en-us/research/wp-content/uploads/2025/01/lee_2025_ai_critical_thinking_survey.pdf

[24] Ive, J. (2021) Commencement Address, California College of the Arts, May 2021.
[25]OECD: 'Artificial Intelligence and Competitive Dynamics in Downstream Markets.' https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/11/artificial-intelligence-and-competitive-dynamics-in-downstream-markets_c6e81d0e/ccf0624a-en.pdf

[26] Premier Science: 'Artificial Intelligence in Graphic Design: Challenges and Opportunities.' https://premierscience.com/pjs-25-1147/

[27] Kimbell, L (2015) *Applying Design Approaches to Policy Making: Discovering Policy Lab*. University of Brighton.

research. While simulating users lowers costs, there is a difference between being modelled and being listened to, and users are all too aware of it.

Trust issues also manifest within professional cultures. The RIBA AI Report 2025 found that while nearly 60 per cent of UK architectural firms now use AI, up from 41 per cent just a year earlier, almost 70% of architects believe AI increases the risk of work being imitated, and nearly half fear that unqualified use could result in unsafe or inadequate buildings[28]. More than nine in ten reject the notion that AI could substitute for professional decision-making. The overall conclusion was that the profession is adopting AI faster than it is developing the governance, professional behaviours and ethics to use it well.

[28]RIBA / Beale & Company: 'RIBA AI Report 2025.' https://beale-law.com/article/riba-ai-report-2025-reveals-how-artificial-intelligence-is-transforming-architecture-and-construction/

A DIAGNOSTIC FRAMEWORK FOR HUMAN-AI DESIGN PRACTICE

The critical issue is how design processes are structured around AI: who frames the problem, who generates the options, who owns the assumptions, and who can be held responsible when things go wrong.

The six dimensions in Table 1 distil the argument of this paper, and Centaur and Minotaur configurations are described for each. It can be used to audit where you are, design where you want to be, and spot where the drift toward automation bias might already be happening.

Table 1: The Centaur-Minotaur Model Applied to Design Practice

Dimension	Centaur (Human First)	Minotaur (AI First)
Who frames the problem?	Designers do the diagnostic work first - identifying user needs, site conditions, material constraints, and the critical questions -before AI is engaged. The design brief is defined by human judgement.	The design problem is shaped by what the AI has been trained on and what its prompts make visible. The frame feels analytically grounded but has not been consciously chosen.
Where do options come from?	Designers develop a guiding direction based on research and judgement. AI generates variations that are evaluated, filtered, and developed further.	AI generates options that appear comprehensive - layouts, prototypes, facade studies, collections. Designers evaluate and select, leaving the generative work to the machine.
Who owns the assumptions?	Designers surface the assumptions embedded in both their own synthesis and the machine's outputs. This includes what the AI treats as good design, whose aesthetic it has learned, and which contexts it may not see. Where AI and domain expertise conflict, designers make the call explicitly. Where AI and domain expertise	The AI's assumptions - about what 'good' looks like, whose aesthetic it has learned, which contexts are absent - are not made visible or open to challenge.
What can the machine not see?	Designers actively consider absent perspectives: underrepresented users, emotions. cultural contexts, memories, material qualities, and factors outside training data.	The analysis feels comprehensive because of its scale. What is absent from training data is not registered as a gap.
How is work sequenced?	Human synthesis and design direction come first. AI is then used to interrogate, extend, and pressure-test - not to generate the design logic itself.	AI output is presented first and treated as a shared reference point. Human reasoning develops in response to the machine's framing, shifting the burden of proof.
Who bears responsibility?	Designers make the calls. They can explain the rationale with an auditable trail of intentions and be held accountable for outcomes. The designer remains the author of the process, not merely the editor of the output.	Responsibility is distributed across those who prompted, reviewed, and approved. Where key decisions were actually made is hard to pinpoint.

WHAT DESIGN LEADERS SHOULD DO: FOUR PRACTICAL ACTIONS

None of the above is an argument against AI in design. That would be tantamount to King Canute urging the tide to retreat.

The critical question is how to design the processes around AI so that human judgement remains central, and the design process - the thinking, testing, iterating and refining - is not displaced by automated generation.

For well-defined optimisation-based tasks AI-first approaches may be entirely appropriate. But for the decisions that define what a design is *for*, whose needs it serves, and what it means, the centaur configuration is not optional. This is where creative judgement, ethical responsibility, and design thinking must remain human-led. It is the newly emerging frontier of creative decision making.

With this in mind, there are four things that design leaders should consider as practical actions.

1. Sequence human judgement first, before AI output. The most impactful change most design organisations can make: ensure that designers form and articulate their own judgements before AI synthesis is presented. Once a fluent, comprehensive AI analysis is on the table, the burden of proof shifts. Pre-committing to a human frame - before opening the AI results - keeps the centaur's head where it belongs. Stanford's research on complementary AI confirms that the most effective human-AI systems are those designed around how humans will interact with them, not those that optimise for predictive accuracy.[29]

2. Make the machine's assumptions visible and contestable. Build explicit checkpoints at which teams ask: what has the AI been unable to see? Which users, contexts, geographies, or traditions are absent from its training? What would a dissenting view look like? In architecture, this means asking what civic, spatial, cultural, or environmental dimensions lie outside the model's parameters. In fashion, it means asking whose aesthetic the model has learned and whose it has not. The CHI design fixation study found that the effectiveness of human-AI co-ideation rests on the strategies designers use to respond to AI suggestions - not on the quality of the AI's output.

3. Protect the capability pipeline. If AI is automating the analytical and creative work that used to develop design judgement among junior staff, organisations must create deliberate alternative learning pathways. The State of AI in Design survey found that 96 per cent of designers have learned AI through self-teaching; formal training remains the exception.

The comparative creativity research shows that novice designers are the most susceptible to AI-induced fixation - and the most likely to benefit from structured human-first processes. Organisations need to institutionalise regular structured reflection on AI outputs, AI-free analytical processes, and mentorship that rebuilds the practices the machines are displacing. A profession that stops investing in the development of human judgement may gain efficiency but will lose the capacity to do the work that matters.

The Creative Industries Skills Audits reached a similar conclusion, finding that as AI becomes more pervasive, “it would give greater weight to workers having strong, distinctly human skills, such as critical thinking, imagination, creativity and interpersonal skills."

4. Design the decision architecture, not just the interface. The field of human-AI interaction has focused heavily on interface-level solutions that are built into AI tools themselves: confidence scores, transparency mechanisms, and the like. These matter, but they only address the surface of the problem. The deeper challenge is organisational: who presents what to whom, in what order, with what authority?

Research on human-AI collaborative decision-making frames this correctly as an organisation design problem:[30] the configurations of division of labour and mutual learning between human and AI are design choices that must be made deliberately, not allowed to emerge by default. Figma's 'pilot not copilot' philosophy is an example of

[29]Stanford News: 'Research shows how to structure AI-human collaboration.' https://news.stanford.edu/stories/2025/05/research-ai-human-collaboration

[30]Journal of Organization Design: 'Human-AI collaborative decision making as an organisation design problem.' Puranam, P. https://link.springer.com/article/10.1007/s41469-021-00095-2

deliberate architecture at the level of tools. ZHA's 10-to-15-per-cent selection rate is an example at the practice level. But the principle needs to extend to the entire design process: how research is sequenced, how synthesis is presented, how options are generated, and how decisions are made and owned[31].

This also addresses a widespread concern: that AI moves too fast for any designed process to keep up. With new tools emerging weekly, a process built around human framing, sequencing and ownership can hold steady through the churn. The soundest course is to maintain well-established processes and use AI as a tool within them, rather than rebuilding processes around each new model.

WHO IS IN CHARGE HERE?

The drift from centaur to minotaur is a critical risk of this current moment, and it is one that very few across the design professions are managing deliberately. Design organisations are adopting AI at speed. They are not, for the most part, asking whose head is on whose body.

Beyond design, the businesses, institutions and government bodies that shape the built and digital world are harnessing AI tools in ways that will make this challenge even more prominent.

The diagnostic framework offered here is a starting point - a way to examine design processes across disciplines and ask honestly whether humans are still leading, or whether the unconsidered biases and unstated architecture of decisions have transferred agency to machines.

It is worth being specific about why this matters so much. In chess, for roughly a decade, centaurs - human-AI teams - were the strongest players in the world. That period ended: in restricted, rule-bound domains with objective outcomes, humans now add little to AI performance.

But design is not chess. Whether it is an architect reading the civic memory of a site, a service designer hearing what users are saying without words, or a fashion designer knowing when a collection needs to feel unfinished rather than perfect, design is defined by ambiguity, competing values, absent information, empathy, and judgement calls.

The question that really matters, in the end, is the same one the architect, the service design researcher and the fashion designer each faced.

Who is in charge here?

This discussion paper was authored by Ben Ramalingam, Mat Hunter and Cat Drew. Comments and feedback gratefully received from Minnie Moll, William Eccleshare, Elena Marco and Shaun Hazeldine, Paul Monaghan and Tom Gardner.

Suggested citation: Ramalingam, B., Hunter, M. and Drew, C. (2026) The Centaur and the Minotaur: Design, Artificial Intelligence and the New Frontiers of Creative Decision-Making, Design Council: London

[31] See for example how service design was used to develop an ethical framework for data and AI in the UK Government Data Service in Drew, C *Design for data ethics: using service design approaches to operationalize ethical principles on four projects*. *Philos Trans A Math Phys Eng Sci* 13 September 2018; 376 (2128): 20170353