The Empathy Engine:

Elevating Human Performance with Generative AI

The 'WEIRD' Psychology of Generative AI: A Hidden Default Setting

Before deploying generative AI, schools admissions teams must understand its inherent psychological biases. These systems are not neutral tools; they are overwhelmingly shaped by training data from a narrow slice of humanity: namely, people from Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies. This is a foundational characteristic with significant implications for engaging across cultural boundaries.

A useful framing, recently articulated by Andrej Karpathy, is to think of LLMs not as entities that hold opinions or possess understanding, but as simulators.[1] When prompted, an LLM does not draw on lived experience or cultural knowledge; instead, it activates statistical patterns from its training data. The critical question is not "what does the AI think?" but "what is the AI's default simulation calibrated to?"

Research from Harvard University ("Which Humans?") has shown that when technologists claim AI achieves "human-level performance," they largely overlook the vast psychological diversity of humanity. WEIRD populations are global outliers, characterised by a distinct set of psychological traits that most of the world does not share. These traits include a strong sense of individualism (viewing the self as an independent, autonomous entity), a preference for analytic thinking (analysing objects in isolation from their context), and a reliance on impersonal prosociality (extending trust to strangers based on rules and contracts rather than established relationships). When a Large Language Model (LLM) is trained on internet data, it inevitably absorbs this peculiar psychological profile.

Evidence of AI's WEIRD Bias

A systematic comparison of ChatGPT's responses to large-scale, cross-cultural human data provides clear, quantifiable evidence of this bias:

- **Clustering with Western Nations:** In a hierarchical cluster analysis using data from the World Values Survey, GPT's responses were psychologically closest to those from the United States, Canada, the UK, and other WEIRD nations. Conversely, its responses were most dissimilar to Asian cultures.
- **Correlation with Cultural Distance:** Researchers found a substantial inverse correlation between a nation's cultural distance from the U.S. and the similarity of its population's survey responses to GPT's. In short, the less a culture resembles the United States, the less the AI's "thinking" resembles that of its people.
- **Cognitive Style Bias:** In a cognitive "triad task," where participants group two out of three related items, GPT exhibited a strong preference for analytic thinking (e.g., pairing "hair" and "beard" because they are in the same abstract category). This contrasts sharply with the

[1]Karpathy is a prominent computer scientist, who was a founding member of OpenAI and former Senior Director of AI at Tesla.

holistic thinking common in many non-WEIRD populations, who would be more likely to pair “hair” and “shampoo” based on their functional relationship.

In Karpathy’s terms, the model’s “statistical centre of mass” sits squarely within WEIRD psychological norms. It is not that the model lacks knowledge of non-Western communication patterns, which is encoded in its parameters from multilingual and multicultural training data. The problem is one of activation: without deliberate intervention, the model’s default simulation reliably produces WEIRD-calibrated output. The culturally diverse knowledge is latent but under-activated.

This phenomenon can be summed up as “WEIRD in, WEIRD out.” For international schools, this poses a significant but unrecognised risk: using a generic LLM means operating with a WEIRD cultural default, encoding a specific worldview into the parent communications process, which has real financial consequences.

The High Cost of the Empathy Gap

In the increasingly competitive world of international school admissions, a mismatch between a school’s communications and a prospective family’s cultural expectations can lead to an “empathy gap”: a critical friction where Western-leaning, low-context communication paradigms alienate the very families that constitute their primary growth market.

At an international school with tuition fees in the region of $20,000, a single family with one or two children typically represents over $200,000 in lifetime revenue.

The $200k Problem

Consider that a single, templated email response, optimised for Western efficiency, can inadvertently fail across different high-context cultures, resulting in lost lifetime revenue of over $200,000 per family. For example, this seemingly innocuous message: “Thank you for your interest! We’d love to schedule a campus tour at your earliest convenience.”

- **For a Chinese family:** This response contains no data on university outcomes or academic rigor. The family’s internal reaction is, “Are they serious?” The school may be immediately perceived as unserious because it fails to address the deep-seated anxiety around measurable success, an echo of the hyper-competitive Gaokao system.
- **For a Japanese family:** This response may be perceived as overly direct and pressuring. The family’s reaction is, “Are they rushing us?” It violates the cultural norm of slow, consensus-based decision-making (Ringi system), creating discomfort and mistrust.
- **For a Thai family:** This response may be interpreted as cold and transactional, lacking the warmth and deference required to build a relationship. The family may think it violates the principle of Kreng Jai (consideration) by attempting to force a quick outcome rather than investing in social harmony.

In all cases, the probable outcome is the same: no response. The family silently disengages to preserve face, and the school loses a high-value enrolment without realising the mistake. The failure lies in the embedded psychological assumptions of the communication.

WEIRD Psychological Assumptions	Non-WEIRD Family Frameworks
Individualism: The primary goal is to maximise the individual child's happiness and autonomy.	**The Interdependent Self:** The child's success is a reflection of family honor and the fulfilment of filial duty.
Analytic Communication: The goal is a direct, linear, and efficient exchange of information to complete a task.	**Holistic Thinking:** A school is not just a service; it is a node in a complex social network. The context is as important as the content.
Impersonal Trust: The relationship is transactional. Trust is based on the contract and the quality of the service provided.	**Relationship-Based Trust:** Trust is not given; it is earned over time through demonstrated respect, warmth, and the recognition of hierarchy.

In sum, cultural fluency is not a "soft skill" but a hard asset: the empathy gap is a direct cause of revenue loss. But if generic AI creates cross-cultural friction by embedding WEIRD communication styles at scale, can a different application of AI be used to solve it?

From Automation to Augmentation: A New Model for Improved Human Performance

A landmark study of a generative AI deployment in a customer support setting provides an instructive counter-narrative. The research, detailed in the NBER paper "Generative AI at Work," demonstrates how an AI framework, when carefully designed to augment human judgment, can drive valuable improvements in quality and consistency across a team.

The study, which tracked over 5,000 customer support employees at a Fortune 500 company, found that giving them access to an AI-based assistant increased their performance, measured by issues resolved per hour, by 14% on average. This gain was a composite of several factors:

- A significant decline in the average time required to handle a single chat.
- An increase in the number of concurrent chats a worker could effectively manage.
- A measurable increase in the percentage of problems that were successfully resolved.

The most critical finding, however, was the uneven impact of the AI tool. The study found a 34% performance improvement for newer employees, whereas the impact on the most experienced employees was lower.

This AI tool effectively reshaped the employee experience curve. Analysis from the study shows that employees with access to the AI tool needed only two months of experience to perform as well as those who had over six months without the tool. The AI accelerated the path to competence, compressing the time it took for new hires to become proficient, productive members of the team.

While customer support and school admissions appear different, they do share some critical characteristics: both are high-volume, relationship-based interactions where expertise is unevenly distributed and difficult to codify. So the underlying mechanism reveals the real potential of generative AI for admissions teams' performance in an international school setting.

If the problem is that LLMs are simulators whose default calibration is Western-biased, and the evidence shows that AI copilots drive performance by disseminating tacit expert knowledge, then the solution is not to replace the simulator, but to re-calibrate it. The model already encodes culturally diverse knowledge in its parameters; it simply needs to be moved from its statistical centre of mass to the right cultural territory, then guided with structured guardrails and examples. This is the principle that underpins Eplex's context assembly line: providing structured context that activates the model's latent cultural knowledge at the point of generation.

The Mechanism of Success: How Generative AI Scales Tacit Knowledge

The AI's success in the customer support study was not a result of simple automation. Instead, it stemmed from the system's unique ability to distil and disseminate implicit, hard-to-articulate skills. In international schools in Asia, cultural fluency exemplifies this kind of high-value tacit knowledge.

Such a tool would work by encoding best practices inside the workflow, providing effective on-the-job training that ensures conversational patterns, content, and culturally respectful language through comprehensive guardrails and examples. In short, best practices would be delivered in-workflow at the point of use.

Technically, this works because the LLM is a simulator, which does not need to "understand" Thai Kreng Jai or Japanese Ringi. Instead, it needs to be given the right context so that its simulation is calibrated to those cultural frames rather than its Western default: reference scaffolds and cultural rules serve as the structured context that shifts the model's output, activating knowledge the model already possesses but would not surface unprompted. The result is output that reads as culturally fluent, not because the AI "knows" the culture, but because it has been positioned in the right region of its vast (but only vaguely remembered) knowledge base and given expert examples to follow.

The study also provides compelling evidence of lasting skill convergence and knowledge transfer:

- **Durable Learning:** when the AI tool was unavailable, employees still performed better than their pre-AI baseline. This suggests they had acquired genuine, enduring skills, not just a crutch.
- **Communication Pattern Convergence:** An analysis of chat logs revealed that after the AI was deployed, the communication patterns of less experienced employees changed significantly. Their language, phrasing, and conversational structure became measurably more similar to those of greater experience.
- **Improved Customer Interactions:** The quality of service improved demonstrably from the customer's perspective. AI assistance led to a marked improvement in customer sentiment scores.

In essence, the AI assistant became a real-time coach for every employee, raising the baseline of quality and consistency across the entire team. This provides a blueprint for ensuring the quality of cross-cultural communications in admissions.

Why Not Just Prompt ChatGPT?

A reasonable objection follows from the argument above: if the model's latent cultural knowledge can be activated by providing the right context, why not simply tell ChatGPT to "respond as if writing to a

Thai family"? In fact, recent research (Tao et al., 2024) shows that cultural prompting does improve alignment for 71–81% of countries.

The gap between cultural prompting and a purpose-built system is the gap between a one-off instruction and a reliable process. Ad-hoc prompting breaks down in practice for several reasons: it produces inconsistent results across different staff members, who will each phrase the prompt differently; it offers no structural guardrails to prevent culturally inappropriate phrasing from slipping through; it lacks scenario-level granularity (a fee enquiry from a Chinese family requires fundamentally different handling from a campus visit request); and it suffers from prompt drift over time as staff revert to convenience. A purpose-built system encodes these distinctions once, maintains them consistently, and delivers them at the point of use without requiring each officer to become a prompt engineer. The difference is analogous to the gap between a spreadsheet and a CRM: both store data, but only one enforces a process.

Moving from 'WEIRD In' to 'Expert In'

The two cases presented in this report reveal the two contrasting faces of generative AI. The first, seen in international admissions, demonstrates the danger of generic models that carry a hidden "WEIRD" cultural bias, creating costly empathy gaps with global families. The second, from the customer support study, illuminates the potential of a purpose-built AI framework, aware of contextual nuance, to improve engagement across cultures at scale.

The shift from "WEIRD in" to "Expert in" is, at its core, a shift in what the simulator is given to simulate. A generic LLM, with a typical prompt, simulates a WEIRD personality vector implied by its training data. The same LLM given dynamic cultural scaffolds, institutional knowledge, and brand voice constraints simulates a culturally fluent admissions expert. Eplex's context assembly line is the system that makes this shift reliable and repeatable.

This is why the quality of the context engineering matters more than the choice of model.

The challenge is not how to automate tasks, but how to enable human excellence in a complex and high-stakes situation like international school admissions. And the key is in how we approach AI implementation.

- **Reframe AI's Role from Automation to Augmentation.** The primary value of generative AI is not as a tool for replacement, but as a system for capturing and disseminating valuable tacit knowledge across an existing team.
- **Invest in Specialisation for Quality Control.** Value is created by an additional layer of domain-specific data and understanding that represents specialised "best practices." This process creates a uniquely transparent "Expert-in-the-Loop" system that transforms robotic text generation into a context-aware competitive asset.

In future, AI in admissions will prove most effective when the right model, framework and context combine to achieve genuine personalisation. The prize is the transformation of a templated one-size-fits-all process into an empathy engine that conveys a real sense of belonging.

Research by Eplex (www.eplex.ai). Eplex is a context engineering tool that encodes cross-cultural admissions expertise into the workflow of every school.

Note: These patterns represent cultural tendencies supported by in-depth research. Individual families vary, sometimes significantly. The core insight is that Western-default communication systematically alienates the majority pattern in each cultural group, but your specific understanding of the family should always balance any cultural tendency.