The Forest for the Trees: Global vs. Local Advice in Human-AI Interaction

Orsolya Szőcs[1] (orsoszocs@gmail.com), Hanqi Zhou[1,2] & Charley M. Wu[1,2]
[1] Human and Machine Cognition Lab, University of Tübingen, Tübingen, Germany
[2] Max Planck Institute for Biological Cybernetics, Tübingen, Germany

Abstract

Artificial intelligence (AI) can enhance human decision-making by providing assistance at different levels of abstraction. This study investigates whether AI should offer broad, high-level guidance (global AI) or focused, low-level assistance (local AI) to optimise performance and learning. Using a hierarchical multi-armed bandit task where both AI types provide equally valuable recommendations, we evaluate how participants leverage AI support in making sequential decisions. Findings reveal that while participants benefited from both types of AI suggestions, global AI led to significantly greater performance improvements. These results contribute to our understanding of human-AI interaction in hierarchical problem-solving, highlighting the importance of designing AI systems that effectively support human cognitive processes.

Keywords: human-AI interaction; hierarchical decision-making; cognitive assistance; complementary AI

Introduction

As the use of artificial intelligence (AI) in decision-making grows, determining how AI can best complement human cognitive processes becomes an important and relatively under-explored topic (Bastani, Bastani, & Sinchaisri, 2021; Elmalech, Sarne, Rosenfeld, & Erez, 2015; Steyvers, Tejeda, Kerrigan, & Smyth, 2022). A key question is how to structure the division of labour between humans and AI systems (Roth, Sushereba, Militello, Diiulio, & Ernst, 2019). Building on a classification framework by Song, Zhu, and Luo (2024), we ask whether it is more effective for AI to offer global (i.e., high-level and general) or local advice (i.e., low-level and specialised).

Human decision-making frequently involves hierarchical problems, where global objectives depend on the interaction between local elements (Consul, Heindrich, Stojcheski, & Lieder, 2022; Gavetti, Levinthal, & Rivkin, 2005; Piriyajitakonkij, Itthipuripat, Ballard, & Pappas, 2024). For example, designing a home requires deciding the function of each room at a global level, while also arranging each room's furniture at a local level (Fig. 1A). Similarly, travellers make global decisions by scheduling an overall route and also local decisions at each traffic light or restaurant menu. By attending to patterns at both levels, people build structured knowledge (Karuza, Thompson-Schill, & Bassett, 2016; Liu et al., 2023), and reduce cognitive load by decomposing complex tasks into manageable subtasks (Correa, Ho, Callaway, Daw, & Griffiths, 2023; Ho et al., 2022; Huys et al., 2015; Rubino, Hamidi, Dayan, & Wu, 2023).

In this view, AI can complement human decision-making across an entire spectrum of specificity, from global strategy to local execution. While current AI navigation tools typically give low-level, turn-by-turn instructions, they could instead suggest useful subgoals along the way (e.g., reach the river, then continue toward the mountains), which may boost situational awareness and spatial learning (Ishikawa, Fujiwara, Imai, & Okabe, 2008; Zhang & Li, 2025). Another example comes from the medical field (Midyett, 2023; Sbitan, Alzraikat, Tanous, Saad, & Odeh, 2025), where AI recommendations are transitioning from a high-level, one-size-fits-all approach (e.g., patients with hypertension should exercise regularly) to low-level, personalised medicine (e.g., this patient should engage in moderate cycling for at least 30 minutes 5 times a week).

Even large language models (LLMs), though flexible, do not automatically switch between high-level and low-level guidance but require explicit prompting from the user to operate at the desired level of detail (Mu, Bai, Bontcheva, & Song, 2024; Shi, Su, Yang, Yang, & Cai, 2023). For instance, when using LLMs for a complex task such as vacation planning or event coordination, a user might have to first ask for an itinerary outline at a global level, and then separately request detailed daily schedules or checklists at a local level. In summary, most current AI tools offer help at a fixed level of granularity, and users must adapt to that level. This observation motivates our investigation into which level of AI assistance (global or local) is more effective for human problem-solving.

Goal and scope. We evaluate whether global or local AI assistance better complements human decision-making in a hierarchical contextual bandit task (Fig. 1B-G). In a between-subjects design, we paired study participants with AI advisors offering either high-level (*global AI*; Fig. 1D) or low-level (*local AI*; Fig. 1E) guidance. While both AI types improved decision-making, global AI assistance led to significantly better performance. When global AI users adhered to the AI recommendation, they managed to improve upon it further and receive higher rewards than local AI users. Moreover, participants showed a subjective preference for the global AI as they rated it more useful, and reported higher adherence to AI suggestions, even when objectively that wasn't the case.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

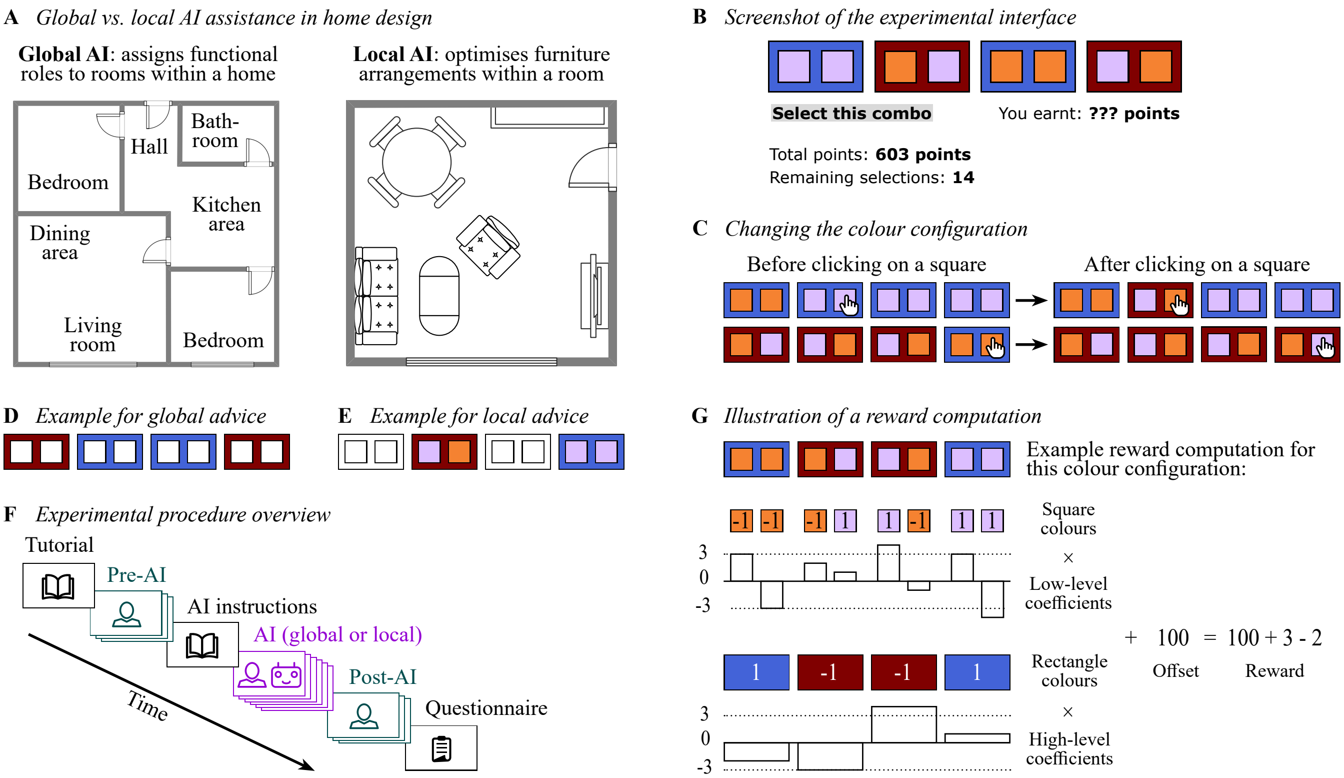

Figure 1: **Task overview.** (**A**) An intuitive example of global and local AI assistance in home design: a global AI might suggest which part of the home would best serve as the kitchen, dining area, or living room, whereas a local AI might help choose the placement of the couch, TV, or coffee table within the living room. (**B**) A screenshot of the hierarchical contextual bandit task, represented as a row of coloured squares (low-level features) and rectangles (high-level features). (**C**) Participants clicked on a square to modify its colour, while the colour of the surrounding rectangle would change automatically based on whether the pair of squares was the same or different colours. (**D**) The global AI recommended the colours of the four oblong rectangles. (**E**) The local AI suggested the specific colours of two pairs of squares and the two surrounding rectangles. (**F**) Experiment procedure. Participants completed 3 pre-AI rounds (without AI), 6 AI-assisted rounds (either global or local; between-subjects), and 3 post-AI rounds (without AI). (**G**) The payoff for each solution was scored as a function of the squares and rectangles that were selected, where each round corresponded to a different payoff function with different low-level and high-level coefficients.

Methods

Participants completed a hierarchical decision-making task where they aimed to maximise rewards across a series of contextual bandit trials (Fig. 1B). On each trial, they selected a colour configuration with a noisy reward determined by two hierarchical levels: (1) *low-level features*: eight individual squares with a binary colour (e.g., orange/lavender); (2) *high-level features*: four rectangles with their binary colours (e.g., maroon/blue) determined by an exclusive-NOR gate (1 if the two squares inside match, or -1 if they differ). Participants could modify the colour configurations by clicking on the low-level squares (Fig. 1C), whereas the high-level rectangles updated automatically.

Participants and design

We recruited 101 participants on Prolific. After excluding 4 participants via multivariate outlier detection (Fig. 3A), we arrived at a sample size of $N = 97$ participants ($M_{\text{age}} = 34.0 \pm 10.5$; 47 female, 48 male). Participants were assigned to either the *global AI assistance* or the *local AI assistance* condition in a between-subjects design (Fig. 1D-E). Participants were compensated with a base payment of £3.60 and a performance-based bonus of up to £3.60. On average, participants spent 23.2 ± 11.3 minutes on the study, and earned £4.76 $\pm$ £0.38. The study was approved by the Ethics in Psychological Research Commission of the University of Tübingen (Wu_2021/0124/213) and informed consent was obtained from all subjects.

Materials and procedure

Participants completed 12 rounds in three phases—pre-AI (3 rounds), AI-assisted (6 rounds), post-AI (3 rounds)—to measure baseline performance and retention of gains after AI removal (Fig. 1F). Each of the 12 rounds comprised 20 trials where participants chose a colour configuration and then received a reward. Every round started with a random configuration of the coloured shapes that the participants could modify to their liking. To minimise memory effects, the interface showed all prior choices and rewards from that round.

Before starting the experiment rounds, participants received instructions and completed a tutorial explaining the task's mechanics. They learned to toggle square colours by clicking them, that rewards depended on both square and rectangle configurations, and that optimal settings for elements were independent. To avoid biases due to colour preference, rounds cycled through a palette of five visually distinct and colour-blind accessible colours. After the tutorial, participants were required to pass a comprehension test with perfect accuracy before proceeding to the experimental rounds.

After the pre-AI rounds, participants received condition-specific instructions about their AI assistant (global or local) and had to pass a second comprehension check about the AI advisor. During AI-assisted rounds, participants received a different AI suggestion on each trial. They were informed that the AI adapted its suggestions based on their within-round choices and performed better when participants explored diverse configurations. Participants could view and freely accept or ignore the AI's advice.

Critically, the AI's guidance differed between conditions. In the local AI condition, the suggestions targeted two of the four rectangles and the containing squares (Fig. 1E), with the chosen subset changing between rounds. In the global AI condition, suggestions always targeted all four rectangles (but not any squares; Fig. 1D). These two settings were strictly controlled to ensure the same quantity and quality of information (see Balancing AI conditions). After completing all rounds, participants answered a questionnaire evaluating the perceived usefulness of the AI and their adherence to its suggestions.

Reward landscapes

A reward landscape is a mapping between colour configurations and numerical rewards, determining how each configuration is scored within an experiment round (Fig. 2). We generated the reward landscapes based on the parametric interaction model from Buzas and Dinitz (2013) and Reeves and Wright (1995), which provides a flexible, hierarchical framework that allows for defining rewards with relatively few parameters while encoding dependencies between global and local elements. Mathematically, these landscapes are similar to NK landscapes (Kauffman & Levin, 1987) that have been widely used in studies of individual and collective decision-making (Barkoczi, Analytis, & Wu, 2016).

In each reward landscape, every high-level element (rectangle) is associated with a high-level coefficient $\mathbf{h} \in \mathbb{R}^H$, and every low-level element (square) is associated with a low-level coefficient $\mathbf{l} \in \mathbb{R}^{HL}$, where H and HL are the number of high-level and low-level coefficients respectively. Intuitively, the coefficients determine the importance of each element (rectangle or square) for the total reward (Fig. 1G).

For a given configuration $\mathbf{x} \in \{-1,1\}^{HL}$, the reward $r(\mathbf{x})$ combines weighted contributions from both levels: the value

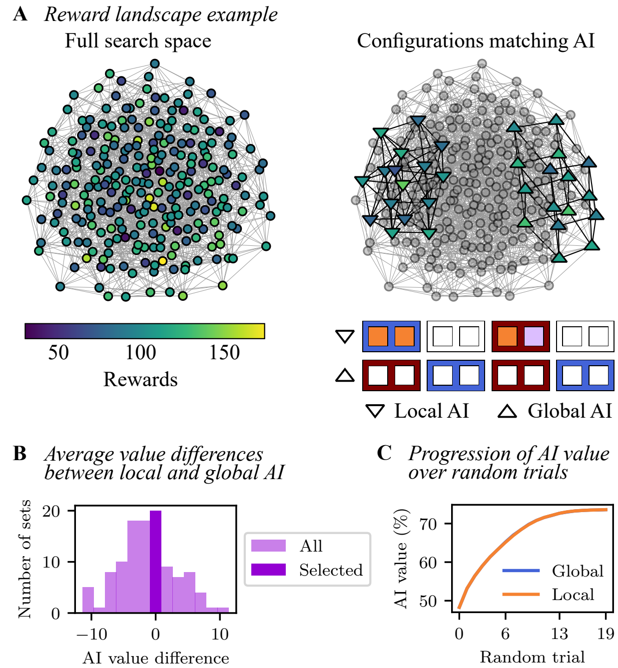

Figure 2: **Reward landscape**. (**A**) Illustration of a reward landscape (left), with an example of local and global AI suggestions (right). Each node corresponds to a colour configuration, where the node colour indicates the associated reward. Connected nodes represent configurations that differ by one square and one rectangle change, or by two squares only. (**B**) Histogram comparing AI values for global and local AI conditions across different reward landscape sets. (**C**) The quality of local and global AI overlapped entirely when they learned from the trials of a random decision-making agent.

of individual squares (local features) is based on their colour (± 1), and the value of each rectangle (global feature) is the product of the corresponding L low-level features $y(\mathbf{x})_i = \prod_{j=(i-1)L+1}^{iL} x_j$ (e.g., 1 if all squares match, -1 otherwise).

Formally, the reward $r(\mathbf{x})$ is defined as:

$$r(\mathbf{x}) = \sum_{i=1}^{H} y(\mathbf{x})_i h_i + \sum_{i=1}^{HL} x_i l_i. \quad (1)$$

To ensure interpretability, the raw reward values were linearly scaled to a range of positive integers. After sampling a minimum reward m uniformly from $[20,70]$, we calculated a scaling factor $s = (m-100)/\min(r(\mathbf{x}))$, where $\min(r(\mathbf{x}))$ is the lowest unscaled reward. When a participant selected a configuration $\mathbf{x}$, the displayed reward was obtained by adding Gaussian noise $sr(\mathbf{x}) + 100 + \mathcal{N}(0,4)$, and rounded to the nearest integer without exceeding the deterministic bounds of the minimum m and the computed maximum. Participants were informed that the maximum rewards could vary in each round, with the average reward fixed at 100 points.

AI assistance

The AI advisors approximated the reward landscape using linear regression models. They incorporated prior knowledge about the structural form of the payoff function (Eq. 1) as well as the value of the average reward. These advisor models learned the feature coefficients $\hat{\mathbf{h}}$ (high-level) and $\hat{\mathbf{l}}$ (low-level) from participant-selected configurations and corresponding rewards. Specifically, the reward r^k for the k^{th} configuration $\mathbf{x}^k$ was approximated as:

$$r^k \approx \hat{r}(\mathbf{x}^k) = \sum_{i=1}^{H} y_i^k \hat{h}_i + \sum_{i=1}^{HL} x_i^k \hat{l}_i + 100, \qquad (2)$$

where $\mathbf{y}^k$ are high-level (rectangle colours) and $\mathbf{x}^k$ are low-level features (square colours). The advisor models updated the high-level and low-level coefficients, $\hat{\mathbf{h}}$ and $\hat{\mathbf{l}}$, to minimise the squared prediction error on the selected configurations.

To generate suggestions, the AI advisors sampled configurations probabilistically using a softmax distribution over estimated rewards $\hat{r}(\mathbf{x})$, favouring high-reward configurations while retaining exploratory variability. The chosen configuration was then partially masked to provide targeted guidance (e.g., global AI masks square values, while local AI masks two of the four rectangle values). From here on, we refer to *AI value* as the mean reward of the colour configurations that match an AI suggestion (i.e., the average of the highlighted subgraph in Fig. 2A right).

Balancing AI conditions

To ensure neither AI condition had an inherent advantage, we controlled for both the quantity and quality of information provided to participants. In terms of information quantity, both global and local AI suggestions transmit 4 bits of information and restrict the search space to the same quantity of matching solutions: from 256 to 16 (Fig. 2A).

To balance information quality, we controlled for three criteria. First, we generated reward landscapes so that the standard deviations of the distributions of the global and local AI values were equal. Second, the most rewarding colour configuration aligned with both local and global highest-value suggestions. Third, we used a landscape filtering process to ensure that the difference between the value of the local and the global AI suggestion was minimised (Fig. 2B) after m random trials with $m \in \{0, \ldots, 19\}$ (matching the 20 trials of the experiment). Specifically, we generated 100,000 landscapes, and then selected 20 sets of 3 landscapes (60 in total), where the absolute sum of differences between global and local AI suggestion values was smallest across the sets and trials. We chose sets of 3 environments because our experimental phases consisted of multiples of 3 rounds, thus entire sets could be assigned to the different phases. This ensured a balanced comparison between the two AI conditions, which can be observed as near-identical AI values when the local and global AI learn from random trials (Fig. 2C).

Results

We screened participants using multivariate outlier detection, excluding 4 participants who were more than 3 standard deviations from the mean based on performance and selection diversity per round (Fig. 3A). The final cohort comprised 52 participants in the *global* AI condition and 45 in the *local* AI condition. All statistical analyses were conducted using linear mixed-effects models with random intercepts for participants. For these we report the estimated regression coefficient b, the standard error SE, the z-score z, and the p-value p.

Performance

We first examined the effects of AI condition (global vs. local), experiment phase (pre-AI, AI-assisted, post-AI), and their interaction on participant scores. These analyses controlled for the mean pre-AI score as a covariate, since it was a strong predictor of overall performance ($b = 0.415$, $SE = 0.063$, $z = 6.558$, $p < .001$). In general, participants significantly improved during the AI-assisted phase compared to the pre-AI baseline ($b = 7.811$, $SE = 1.229$, $z = 6.359$, $p < .001$), but regressed after AI removal ($b = -5.291$, $SE = 1.229$, $z = -4.307$, $p < .001$; Fig. 3B). Due to sample variability, participants in the local AI condition started with greater scores in the pre-AI phase ($b = 4.432$, $SE = 1.804$, $z = 2.457$, $p = .014$). However, in the AI-assisted phase, global AI participants outperformed their local AI counterparts ($b = 3.043$, $SE = 1.229$, $z = 2.477$, $p = .013$), yielding faster learning curves (Fig. 3C). When AI support was withdrawn, the magnitude of performance decline did not differ significantly between conditions ($b = -3.104$, $SE = 1.804$, $z = -1.721$, $p = .085$).

AI adherence

To understand why participants performed better in the global AI condition, we analysed individual *AI adherence rates*, which we defined as the proportion of trials that were consistent with the AI advice. We looked at how AI adherence was influenced by trial number, condition and their interaction. We found no significant difference in adherence rates between conditions ($b = -0.071$, $SE = 4.457$, $z = -0.016$, $p = .987$), and the interaction between condition and trial number also was not significant ($b = -0.087$, $SE = 0.138$, $z = -0.630$, $p = .528$). This suggests that differences in AI adherence rates (or some temporal differences in when AI advice was followed) cannot account for the superior outcomes in the global AI condition.

Next, we analysed how AI adherence influenced participant reward while controlling for AI suggestion value (mean reward across advice-consistent solutions). Fig. 3D shows the results of this analysis, where reward was modelled as a function of AI value, condition, and prior performance (the maximum reward so far in a round and the participant's mean pre-AI score). When participants adhered to the AI advice, the advice value strongly predicted rewards ($b = 0.482$, $SE = 0.016$, $z = 30.322$, $p < .001$), with global AI users benefiting more from AI adherence ($b = 0.071$, $SE = 0.021$, $z = 3.401$, $p = .001$). This suggests that participants may have found it easier to capitalise on the global AI suggestions. However, when participants ignored the AI recommendation (i.e., their choice was inconsistent with the suggestion), AI

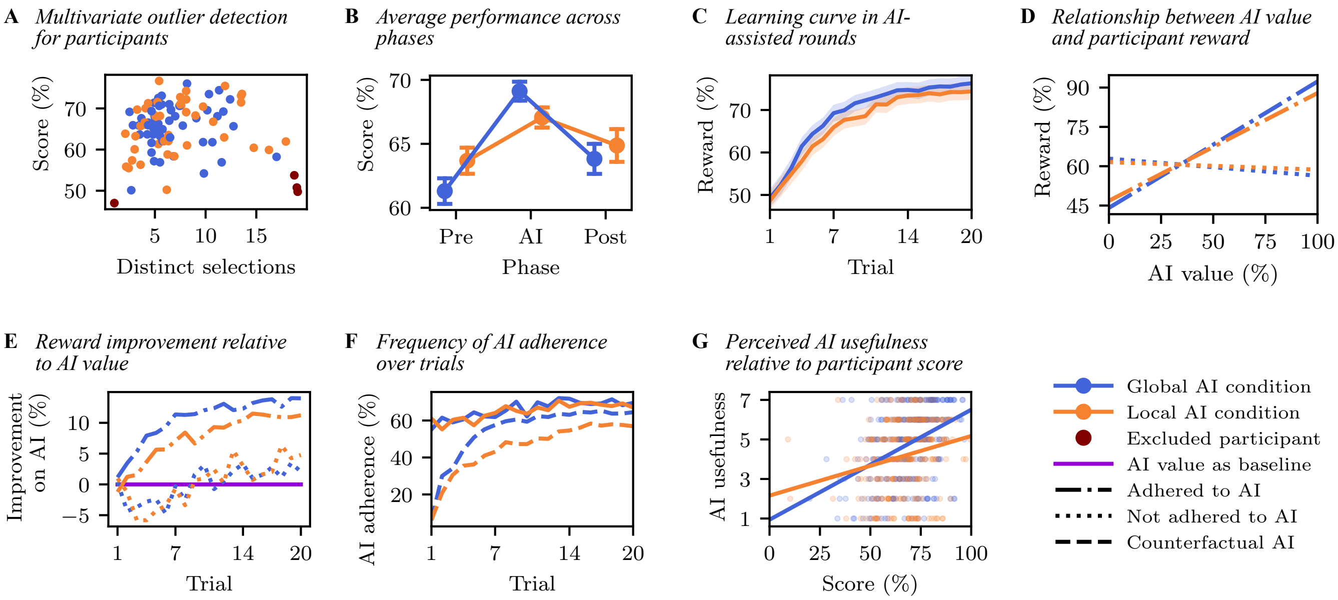

Figure 3: **Results.** (**A**) Identifying the participants excluded from data analysis. (**B**) Mean scores and standard errors in the pre-AI, AI-assisted and post-AI phases. (**C**) Participants in the global AI condition showed greater improvement over trials. (**D**) Regression lines for two cases: when participants adhered to the AI suggestions, and when they did not. (**E**) Difference between participant reward and AI value across trials, showing the average values for two cases: when participants adhered to the AI suggestions, and when they did not. (**F**) The frequency of participants matching the AI suggestions, including the actual AI experienced by the participants, and the simulated counterfactual AI. (**G**) The dots correspond to AI-assisted rounds, indicating the participant's usefulness rating of the AI and their score. The two regression lines correspond to the two conditions.

value negatively predicted participant reward ($b = -0.065$, $SE = 0.024$, $z = -2.691$, $p = .007$), suggesting that disregarding high-value suggestions was costly. In this case, there was no significant interaction between condition and AI value ($b = 0.036$, $SE = 0.033$, $z = 1.078$, $p = .281$), so ignoring AI advice was equally detrimental in both conditions.

Another way of looking at the interaction between the AI and human participants, is to consider how much participants managed to improve upon the AI suggestion (Fig. 3E). We modelled the *relative improvement* on AI based on whether the participant acquired a reward that was higher (positive improvement) or lower (negative improvement) than the value of the AI suggestion. When participants followed the AI recommendations, they obtained significantly higher improvements on the AI values ($b = 11.112$, $SE = 0.454$, $z = 24.468$, $p < .001$), with a stronger effect for the global AI condition ($b = 4.945$, $SE = 0.658$, $z = 7.515$, $p < .001$). This further supports the interpretation that participants found it easier to build on the global AI suggestions. When the AI was not followed, participant performance did not significantly differ from the AI value ($b = -0.217$, $SE = 0.783$, $z = -0.278$, $p = .781$), with no difference across conditions ($b = 1.357$, $SE = 1.146$, $z = 1.185$, $p = .236$).

Counterfactual AI

To test whether the observed differences between global and local AI conditions were due to the quality of the AI suggestions (i.e., an unfair advantage) or because of differences in how participants responded, we simulated *counterfactual AI suggestions* for each participant. This was defined as the opposite AI type: for participants assigned to the global condition, we simulated a local counterfactual AI, and for participants assigned to the local condition, we simulated a global counterfactual AI. For every participant trial, the counterfactual AI suggestions were simulated 100 times. We then computed *counterfactual AI adherence rates*, measuring how often participants unintentionally followed a recommendation that they had not seen (Fig. 3F).

We first compared the real and counterfactual AI values, finding no significant difference between conditions ($b = -0.311$, $SE = 0.788$, $z = -0.394$, $p = .693$), nor any differences in the interaction between suggestion type (actual vs. counterfactual) and condition ($b = 0.223$, $SE = 0.288$, $z = 0.774$, $p = .439$). This confirms equivalent suggestion quality across AI types, ruling out inherent algorithmic advantages.

We then looked at AI adherence rates for both actual and counterfactual advice. We found that global AI users were more likely to adhere to the counterfactual local AI ($b = 0.066$, $SE = 0.016$, $z = 4.001$, $p < .001$). This suggests that, rather than just passive compliance, participants in the global AI condition were more likely to integrate the AI recommendations into their own reasoning process by selecting low-level features that would be consistent with a local AI advisor. There was less integration in the local condition, since participants adhered less to the counterfactual global advice.

Subjective ratings

In the final questionnaire, participants rated their adherence ("I followed the AI's suggestions") and usefulness ("The AI assistant was useful") on a 7-point Likert-scale (1 = Strongly disagree; 7 = Strongly agree), and they also rated usefulness after each AI-assisted round. Higher usefulness ratings were predicted by higher scores ($b = 0.050$, $SE = 0.006$, $z = 8.128$, $p < .001$) and higher AI adherence ($b = 0.009$, $SE = 0.002$, $z = 3.751$, $p < .001$). As seen in Fig. 3G, the relationship between score and the usefulness rating was stronger in the global AI condition ($b = 0.019$, $SE = 0.008$, $z = 2.291$, $p = .022$), suggesting that the global AI users were more likely to attribute their performance improvements to the AI.

We also examined how usefulness rating and observed rate of adherence influenced self-reported adherence. Local AI participants reported significantly lower AI adherence ($b = -0.611$, $SE = 0.152$, $z = -4.030$, $p < .001$). However, observed AI adherence rate and usefulness rating were not statistically significant effects (for both: $b = 0.0$, $SE = 0.0$, $z = 0.0$, $p = 1.0$), suggesting a degree of misalignment between self-perception and reality, while also highlighting the challenge of accurately assessing human-AI interaction based on user-reported data (Papenmeier, Englebienne, & Seifert, 2019). These findings reinforce the idea that local AI participants may have struggled more with AI recommendations, leading them to feel that they were adhering to the AI less, even when objective adherence rates were similar.

Discussion

We examined the impact of global versus local AI assistance on human decision-making in a hierarchical problem-solving task. We found that global AI led to greater performance improvements compared to local AI, with participants better able to improve upon global AI recommendations.

This advantage likely comes from two complementary mechanisms. First, global AI appears to counteract the *myopia of learning* (Levinthal & March, 1993; Nagy, Orban, & Wu, 2025), where decision-makers over-prioritise local elements. Global AI may help scaffold the formation of structured mental models (Zhou, Bamler, Wu, & Tejero-Cantero, 2024) that integrate local and global dependencies. Second, people tend to reduce uncertainty in complex problems by varying one factor at a time (Wüstenberg, Greiff, & Funke, 2012; Bramley, Dayan, Griffiths, & Lagnado, 2017) and by performing dichotomous (Markant, Settles, & Gureckis, 2016) or confirmatory tests (Klayman & Ha, 1987) to obtain easy-to-process feedback. Global AI can constrain the problem space into tractable components while allowing participants to explore locally.

Limitations and future directions. Despite the insights gained, several limitations of our study warrant further investigation. While our task's hierarchical structure was transparent, real-world problems may not have such clear demarcations of global and local features. The explicit high-level features in our task may have made the global AI advice more useful. In real-world tasks, users might not even recognise that an AI suggestion targets high-level features. Additionally, our AI assistants relied on a linear model, reflecting the task's simple structure. In practice, most AI models used in complex, real-world problems are black-box, non-linear models based on deep learning. Future work should test AI advice granularity on tasks with hidden hierarchies, complex structures and opaque AI. In case global AI is only superior in domains where humans naturally think hierarchically (Dedhe, Clatterbuck, Piantadosi, & Cantlon, 2023), identifying those domains will be important for applying these insights.

The intrinsic informativeness of global or local advice may vary significantly across domains. In domains governed by unifying principles—such as classical physics, where a handful of laws explain most phenomena—global advice might immediately solve the problem, making local advice less valuable. By contrast, in domains full of irregularities, such as natural language learning, local advice may be more informative than global advice. Our results are most relevant for domains balancing global and local patterns and those dominated by global patterns. In addition, practical engineering constraints (e.g., data limitations, domain-specific restrictions) may preclude offering truly global guidance, likely shifting optimal advice granularity in real-world systems.

How global and local AI advisors are defined and presented to users can itself shape effectiveness. For example, in Sah, Yoo, and Sundar (2011), participants trusted a robot more when it was introduced as a specialist in physical exercise, as opposed to a generalist in mimicking human body movements, contrary to our results. Future experiments could explore alternative framings, such as introducing AI as a jack-of-all-trades vs. an expert.

Our use of a between-subjects design to avoid carry-over effects introduced its own limitations. While it allowed us to focus on whether global vs. local AI is better at a population level, we did not measure individual preferences for the two AI types. Prior research has shown that people spontaneously specialise into complementary roles during an immersive foraging task (Wu et al., 2025), suggesting that individuals might also self-select global or local AI based on their preferences. A historical parallel comes from advanced chess, where expert players engaged with global strategic planning themselves while delegating local tactical calculations to machines (Kasparov, 2010). Novices, by contrast, may gain more from global advice as they build their mental model. A within-subjects or interactive choice-based design could reveal individual preferences and expertise effects more directly.

Conclusion. Overall, by "seeing the forest for the trees", global AI assistance outperformed local guidance in enhancing human-AI learning. These findings highlight the need to develop AI systems that align with the hierarchical nature of human decision-making.

Code and data

All materials required to replicate the results, including code and data, are publicly available at `https://github.com/orsiszocs/global-vs-local-ai-advice`.

Acknowledgments

This work was supported by the German Federal Ministry of Education and Research (BMBF): Tübingen AI Center, FKZ: 01IS18039A, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy–EXC2064/1–390727645, and funded by the DFG under Germany's Excellence Strategy – EXC 2117 – 422037984.

References

Barkoczi, D., Analytis, P. P., & Wu, C. M. (2016). Collective search on rugged landscapes: A cross-environmental analysis. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 38).

Bastani, H., Bastani, O., & Sinchaisri, W. P. (2021). Improving human decision-making with machine learning. *arXiv preprint arXiv:2108.08454*.

Bramley, N. R., Dayan, P., Griffiths, T. L., & Lagnado, D. A. (2017). Formalizing Neurath's ship: Approximate algorithms for online causal learning. *Psychological review*, *124*(3), 301.

Buzas, J. S., & Dinitz, J. (2013). An analysis of NK and generalized NK landscapes. *arXiv preprint arXiv:1302.3541*.

Consul, S., Heindrich, L., Stojcheski, J., & Lieder, F. (2022). Improving human decision-making by discovering efficient strategies for hierarchical planning. *Computational Brain & Behavior*, *5*(2), 185–216.

Correa, C. G., Ho, M. K., Callaway, F., Daw, N. D., & Griffiths, T. L. (2023). Humans decompose tasks by trading off utility and computational cost. *PLOS Computational Biology*, *19*(6), e1011087.

Dedhe, A. M., Clatterbuck, H., Piantadosi, S. T., & Cantlon, J. F. (2023). Origins of hierarchical logical reasoning. *Cognitive Science*, *47*, 13250.

Elmalech, A., Sarne, D., Rosenfeld, A., & Erez, E. (2015). When suboptimal rules. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence* (pp. 1313–1319).

Gavetti, G., Levinthal, D., & Rivkin, J. (2005). Strategy making in novel and complex worlds: The power of analogy. *Strategic Management Journal*, *26*, 691–712.

Ho, M. K., Abel, D., Correa, C. G., Littman, M. L., Cohen, J. D., & Griffiths, T. L. (2022). People construct simplified mental representations to plan. *Nature*, *606*, 129–136.

Huys, Q. J., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., ... Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proceedings of the National Academy of Sciences*, *112*(10), 3098–3103.

Ishikawa, T., Fujiwara, H., Imai, O., & Okabe, A. (2008). Wayfinding with a GPS-based mobile navigation system: A comparison with maps and direct experience. *Journal of environmental psychology*, *28*(1), 74–82.

Karuza, E. A., Thompson-Schill, S. L., & Bassett, D. S. (2016). Local patterns to global architectures: Influences of network topology on human learning. *Trends in Cognitive Sciences*, *20*(8), 629-640.

Kasparov, G. (2010). The chess master and the computer. *The New York Review of Books*, *57*(2), 16–19.

Kauffman, S., & Levin, S. (1987). Towards a general theory of adaptive walks on rugged landscapes. *Journal of theoretical Biology*, *128*(1), 11–45.

Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological review*, *94*(2), 211.

Levinthal, D. A., & March, J. G. (1993). The myopia of learning. *Strategic Management Journal*, *14*(S2), 95–112.

Liu, J., Chen, D., Xiao, X., Zhang, H., Zhou, W., Liang, S., ... Wang, L. (2023). Multi-scale goal distance representations in human hippocampus during virtual spatial navigation. *Current Biology*, *33*(10), 2024–2033.

Markant, D. B., Settles, B., & Gureckis, T. M. (2016). Self-directed learning favors local, rather than global, uncertainty. *Cognitive Science*, *40*(1), 100–120.

Midyett, L. K. (2023). One size fits all versus individualized medicine in type 1 diabetes management. *Diabetes Technology & Therapeutics*, *25*(S3), S42–S47.

Mu, Y., Bai, P., Bontcheva, K., & Song, X. (2024). Addressing topic granularity and hallucination in large language models for topic modelling. *arXiv preprint arXiv:2405.00611*.

Nagy, D. G., Orban, G., & Wu, C. M. (2025). Interplay of episodic and semantic memory arises from adaptive compression. *PsyArxiv*.

Papenmeier, A., Englebienne, G., & Seifert, C. (2019). How model accuracy and explanation fidelity influence user trust in AI. In *Proceedings of the IJCAI 2019 Workshop on Explainable Artificial Intelligence (XAI)*.

Piriyajitakonkij, M., Itthipuripat, S., Ballard, I., & Pappas, I. (2024). What makes a face look like a hat: Decoupling low-level and high-level visual properties with image triplets. *arXiv preprint arXiv:2102.00521*.

Reeves, C., & Wright, C. (1995). An experimental design perspective on genetic algorithms. In L. D. Whitley & M. D. Vose (Eds.), *Foundations of Genetic Algorithms* (Vol. 3). Elsevier.

Roth, E. M., Sushereba, C., Militello, L. G., Diiulio, J., & Ernst, K. (2019). Function allocation considerations in the era of human autonomy teaming. *Journal of Cognitive Engineering and Decision Making*, *13*(4), 199–220.

Rubino, V., Hamidi, M., Dayan, P., & Wu, C. M. (2023). Compositionality under time pressure. In *Proceedings of the 45th Annual Conference of the Cognitive Science Society*. Cognitive Science Society.

Sah, Y. J., Yoo, B., & Sundar, S. S. (2011). Are specialist robots better than generalist robots? In *Proceedings of the*

6th International Conference on Human-Robot Interaction. Association for Computing Machinery.

Sbitan, L., Alzraikat, N., Tanous, H., Saad, A. M., & Odeh, M. (2025). From one size fits all to a tailored approach: integrating precision medicine into medical education. *BMC Medical Education*, *25*(1), 90.

Shi, C., Su, Y., Yang, C., Yang, Y., & Cai, D. (2023). Specialist or generalist? Instruction tuning for specific NLP tasks. In H. Bouamor, J. Pino, & K. Bali (Eds.), *Proceedings of the 2023 conference on empirical methods in natural language processing* (p. 15336-15348). Association for Computational Linguistics.

Song, B., Zhu, Q., & Luo, J. (2024). Human-AI collaboration by design. *Proceedings of the Design Society*, *4*, 2247–2256.

Steyvers, M., Tejeda, H., Kerrigan, G., & Smyth, P. (2022). Bayesian modeling of human–AI complementarity. *Proceedings of the National Academy of Sciences*, *119*(11), e2111547119.

Wu, C. M., Deffner, D., Kahl, B., Meder, B., Ho, M. H., & Kurvers, R. H. (2025). Adaptive mechanisms of social and asocial learning in immersive foraging environments. *Nature Communications*, *16*, 3539.

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving—more than reasoning? *Intelligence*, *40*(1), 1–14.

Zhang, Y., & Li, J. (2025). The effect of turn-by-turn navigation on spatial memory in large-scale environments and ways to improve it. *Advances in Psychological Science*, *33*(1), 77–91.

Zhou, H., Bamler, R., Wu, C. M., & Tejero-Cantero, A. (2024). Predictive, scalable and interpretable knowledge tracing on structured domains. In *The twelfth international conference on learning representations.*