**Active Inference for Route Choice Under Uncertainty.**

Jamal Amani Rad, Thomas O. Hancock, and Stephane Hess

Choice Modelling Centre & Institute for Transport Studies, University of Leeds

## Short summary

This paper adapts the Free Energy Principle (FEP) to route choice under uncertainty and examines whether an active inference account can explain travel learning better than conventional reinforcement learning (RL). We develop a route choice model in which travellers update beliefs about hidden traffic conditions and select routes by minimising expected free energy, thereby jointly capturing preferences, uncertainty, and exploration. The model is estimated in a hierarchical Bayesian framework and applied to two complementary experimental datasets: an incentive-compatible driving simulator study and a laboratory route choice study with information manipulations. Across both datasets, the FEP model provides a better account of observed behaviour than RL, especially for participants showing exploratory or belief-adaptive responses. The results suggest that active inference offers a promising neurocomputational foundation for richer models of adaptive travel behaviour.
**Keywords**: Active inference; Free energy principle; Learning; Route choice; Travel behaviour.

## 1 Introduction

Travel behaviour modelling has increasingly moved beyond static utility-based representations toward process-oriented accounts that can capture how preferences evolve during deliberation, especially in uncertain and information-rich environments. In transport, this has motivated adaptations of several influential frameworks from mathematical psychology, including decision field theory (Hancock et al., 2018), the multi-attribute linear ballistic accumulator (Hancock et al., 2021), and quantum choice models (Hancock et al., 2020). These approaches have helped explain why observed choices may depend not only on final outcomes, but also on the underlying dynamics of comparison, attention, and evidence accumulation. This is particularly relevant for route choice, where repeated experience, feedback, and uncertainty can shape behaviour over time.

In parallel, transport research has also begun to engage more directly with neurocognitive evidence. For example, Bogacz et al. (2021) incorporated electroencephalography (EEG) into a hybrid choice framework to study decision processes in transport settings. However, most existing models still remain descriptive abstractions of behaviour rather than mechanistic accounts grounded in theories of how the brain learns from and acts under uncertainty. This leaves an important gap between transport modelling, mathematical psychology, and contemporary computational neuroscience (Wong and Wang, 2006).

A particularly promising framework for narrowing this gap is the Free Energy Principle (FEP) and its decision-theoretic corollary, active inference (Friston et al., 2006; Friston and Kiebel, 2009). Under this framework, agents update beliefs and select actions so as to minimise expected free energy, thereby integrating preference fulfilment with uncertainty reduction within a single objective function. Unlike conventional reinforcement learning (RL), where exploration often requires auxiliary assumptions or ad hoc mechanisms, active inference treats epistemic behaviour as endogenous to decision-making itself. This makes the framework especially attractive for route choice, where travellers routinely face partial feedback, uncertain travel times, and an ongoing trade-off between exploiting known routes and exploring uncertain alternatives. Despite this appeal, FEP-based applications in transport have so far been largely confined to perceptual-motor control, visual sampling, or cognitive-state estimation in driving-related contexts rather than travel behaviour modelling itself (Engström et al., 2018; Engström et al., 2024; Wei et al., 2025).

This paper develops an active inference model of route choice under uncertainty and provides one of the first empirical applications of the FEP to travel behaviour. We adapt the framework to repeated route learning by specifying hidden traffic states, travel-time outcomes, and prior preferences over outcomes, and by allowing travellers to update beliefs through experience while selecting routes by

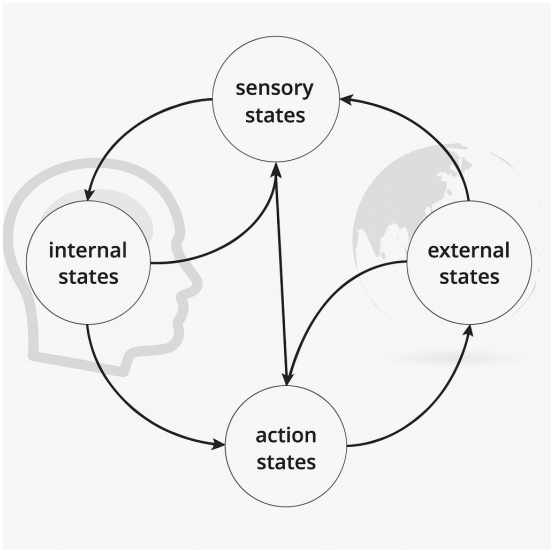



Figure 1: Key notions underlying the free energy principle.

minimising expected free energy. The model is then applied to two complementary experimental datasets, one based on an incentive-compatible driving simulator task and one based on a controlled laboratory route choice experiment with information manipulations. Our objective is to assess whether this neuroscience-inspired framework can provide a richer account of adaptive route choice, and whether it improves on a benchmark RL model in capturing exploratory and belief-sensitive behaviour (Rad et al., 2025).

## 2 Methodology

We model repeated route choice as an active inference problem in which travellers form and update probabilistic beliefs about latent traffic conditions and choose routes so as to minimise expected free energy (see Figure 1). Let each policy $\pi$ correspond to a route choice, let $s$ denote a hidden state of the traffic environment, and let $o$ denote the observed outcome, here interpreted as travel time. The agent is assumed to hold a generative model comprising a prior over hidden states, $P(s)$, a likelihood mapping from states to outcomes, $P(o \mid s)$, and a preference distribution over outcomes, $P(o)$, with higher probability assigned to shorter travel times. In this setting, route choice is treated as inference over the future consequences of alternative policies rather than as the maximisation of an externally specified reward function (Friston et al., 2006; Friston, 2009).
For a one-step route choice problem, the expected free energy associated with policy $\pi$ is written as

$$G(\pi) = \mathbb{E}_{Q(o,s|\pi)} \left[ \ln Q(s \mid \pi) - \ln P(o, s) \right], \tag{1}$$

where $Q(s \mid \pi)$ and $Q(o, s \mid \pi)$ denote the traveller's predictive beliefs under route $\pi$. Lower values of $G(\pi)$ indicate routes that are expected to generate both more preferred and more informative outcomes. Following the standard active-inference formulation, this quantity can be expressed as the sum of a risk term and an ambiguity term:

$$G(\pi) = D_{\mathrm{KL}}[Q(o \mid \pi) \,\|\, P(o)] + \mathbb{E}_{Q(s|\pi)}[H(P(o \mid s))]. \tag{2}$$

In route choice terms, the first component penalises policies whose predicted travel-time distribution diverges from the traveller's preferences, while the second penalises policies associated with uncertain or unreliable travel times (Schwartenbeck et al., 2013; Costa et al., 2021; Friston et al., 2017). This gives the model a natural interpretation in terms of the exploitation–exploration trade-off: travellers prefer routes that are expected to be both attractive and uncertainty-reducing.
Learning is represented as sequential Bayesian updating of the parameters governing the agent's generative model. We assume Dirichlet priors over the relevant categorical distributions and update their concentration parameters after each observed outcome. A generic update for the likelihood parameters can be written as

$$a_{o,s}^{(t+1)} = \omega \, a_{o,s}^{(t)} + \eta \, q_t(s) \, \mathbb{I}[o_t = o], \tag{3}$$

where $a_{o,s}^{(t)}$ is the concentration parameter at trial $t$, $q_t(s)$ is the posterior belief over the hidden state after observing outcome $o_t$, $\eta \in [0,1]$ is a learning-rate parameter, and $\omega \in [0,1]$ is a forgetting parameter (Friston et al., 2017; Smith et al., 2022). Intuitively, recent evidence increases confidence in those state–outcome relationships that are supported by experience, while $\omega$ allows older evidence to be discounted when adaptation is needed.

Policy selection is modelled with a softmax mapping from expected free energies to choice probabilities:

$$P(\pi_t = \pi \mid Q_t) = \frac{\exp(-\beta_t\, G_t(\pi))}{\sum_{\pi'} \exp(-\beta_t\, G_t(\pi'))}, \tag{4}$$

where $\beta_t$ is a precision parameter controlling the determinism of choice. Higher values of $\beta_t$ imply stronger selection of the route with the lowest expected free energy, whereas lower values imply more stochastic behaviour. This formulation links the active-inference model to familiar discrete-choice structures while preserving the richer interpretation that route evaluation depends jointly on preferences, beliefs, and uncertainty (Friston et al., 2015).

We compare this FEP-based specification against a benchmark reinforcement learning model described in our earlier work (Rad et al., 2025). Both models are estimated in a hierarchical Bayesian framework, allowing individual-level parameters to vary across participants while being partially pooled through group-level distributions. This approach is especially useful in repeated choice settings because it captures both heterogeneity in learning and uncertainty in parameter estimates more effectively than non-hierarchical maximum-likelihood approaches. Estimation is carried out in Stan using Hamiltonian Monte Carlo with the No-U-Turn Sampler (Gelman et al., 2013). Posterior summaries are then used to characterise behavioural differences across conditions and to assess the extent to which the active-inference model improves on the RL benchmark in accounting for adaptive route choice.

## 3 Data

The empirical analysis uses two complementary experimental datasets that were selected because they capture different but closely related dimensions of adaptive route choice. The first dataset focuses on the role of experienced versus hypothetical feedback in a simple repeated route-learning task, while the second focuses on the role of prior information and travel-time risk in a richer sequential route-choice environment. Considered jointly, the two datasets provide a strong test bed for evaluating whether an active inference account can capture how travellers learn, explore, and stabilise their choices under uncertainty.

***Dataset I.*** The first dataset is drawn from an immersive driving experiment conducted at the Travel Choice Simulation Laboratory (TRACSLab) at the University of Sydney (Fayyaz, 2021). The laboratory houses five advanced driving simulators equipped with a widescreen display, steering wheel, pedals, and standard vehicle controls, allowing participants to experience route outcomes in a more behaviourally consequential setting than a conventional stated choice survey (one of the simulators is shown in Figure 2). Full details of the experiment and its exploratory analysis are provided in Fayyaz (2021); Rad et al. (2025). Participants were instructed to imagine a routine commute from home to work by car and completed a total of 20 route-choice tasks, comprising 10 simulator-based tasks and 10 stated choice tasks. In each task they chose between two routes: a fixed route with a travel time of 5 minutes and an uncertain route with possible travel times ranging between 2 and 7 minutes. The order of the two task modes was counterbalanced across participants, yielding simulator-first and stated-preference-first groups. In the simulator condition, speeds and signals were constrained so that realised travel times matched the experimental design. The analysis sample contains $N = 83$ participants, corresponding to 1,660 observed choices after the exclusion of non-traders. This dataset is particularly useful because it allows us to examine how the salience and consequentiality of feedback affect subsequent learning and route selection.

***Dataset II.*** The second dataset comes from a controlled laboratory route-choice experiment originally conducted by Ben-Elia and Shiftan (2010). The study involved 49 undergraduate participants at the Technion, who were randomly assigned to either an informed group ($N = 24$) or a control group ($N = 25$). Each participant completed 300 sequential route-choice trials on a computer, divided into three blocks of 100 trials. In every trial, participants chose between Route A and Route B under one of three travel-time environments: Fast & Safe, in which Route A was both faster on average and more reliable (25 $\pm$ 5 minutes versus 30 $\pm$ 15 minutes); Fast & Risky, in which Route A was faster on average but more variable (25 $\pm$ 15 minutes versus 30 $\pm$ 5 minutes); and Low-Risk, in which both routes were relatively reliable but Route A remained faster on average (25



Figure 2: One of the five driving simulators at the University of Sydney used in Dataset I (TRACSLab). The simulator includes a full car cockpit with steering and pedal controls and a wrap-around projection screen.

$\pm$ 5 minutes versus 30 $\pm$ 5 minutes). The order of the three scenario blocks was randomised across participants. The informed group received route-specific travel-time range information at the start of each trial, whereas the control group received no such pre-trip information and therefore had to learn purely from experienced outcomes. For both groups, feedback was partial: after each choice, participants observed only the realised travel time of the route they had selected. This dataset is valuable because it directly separates the effects of prior information, experienced feedback, and route reliability on the dynamics of exploration and exploitation.

## 4 Results and discussion

The empirical results provide consistent support for the proposed active inference account across both experimental settings. In what follows, we focus on three questions. First, does the model recover the main behavioural differences induced by feedback salience and information provision? Second, does it offer a psychologically interpretable account of how exploration and exploitation evolve over repeated route choice? Third, does this additional structure translate into improved fit relative to a benchmark reinforcement learning model (Rad et al., 2025)?

*Dataset I.* In the driving simulator and stated-preference experiment, the free-energy model gave a close quantitative account of how participants adapted their choices over the 20 repeated route-choice tasks. The strongest and most theoretically informative contrast was between participants who completed the simulator block first and those who completed the stated-preference block first. As shown in Figure 3, the posterior distributions of the core parameters separate clearly across these two groups. Participants who encountered the simulator first exhibited a substantially higher initial learning rate, lower asymptotic learning, stronger memory retention, and a steeper increase in choice precision than those who began in the stated-preference setting. In substantive terms, this suggests that vivid and consequence-laden feedback produced larger early belief updates, after which behaviour stabilised more rapidly. By contrast, when participants first experienced the same task in a hypothetical format, belief updating was initially more cautious and remained more plastic over subsequent trials.

This pattern is important because it is not simply a statistical difference between groups; it is also exactly the type of effect that an active inference formulation should recover. Within the free-energy framework, more salient and credible observations reduce epistemic uncertainty more quickly, thereby accelerating the transition from exploratory sampling to more deterministic route choice. The parameter estimates therefore support a coherent behavioural narrative: travellers who first experienced consequential outcomes behaved as if early evidence carried greater reliability, whereas those who first encountered the task in a lower-salience context treated the environment as less settled and remained open to further revision. This interpretation is also broadly consistent with the earlier reinforcement-learning analysis of the same dataset, which likewise found stronger learning and a greater exploration–exploitation shift among simulator-first participants, although the present model explains these effects through a unified inferential mechanism rather than through a reward-learning architecture alone (Rad et al., 2025).

A further strength of the Dataset I results is that the free-energy model outperformed the reinforcement-learning benchmark for the clear majority of participants. Figure 5 reports the individual-level

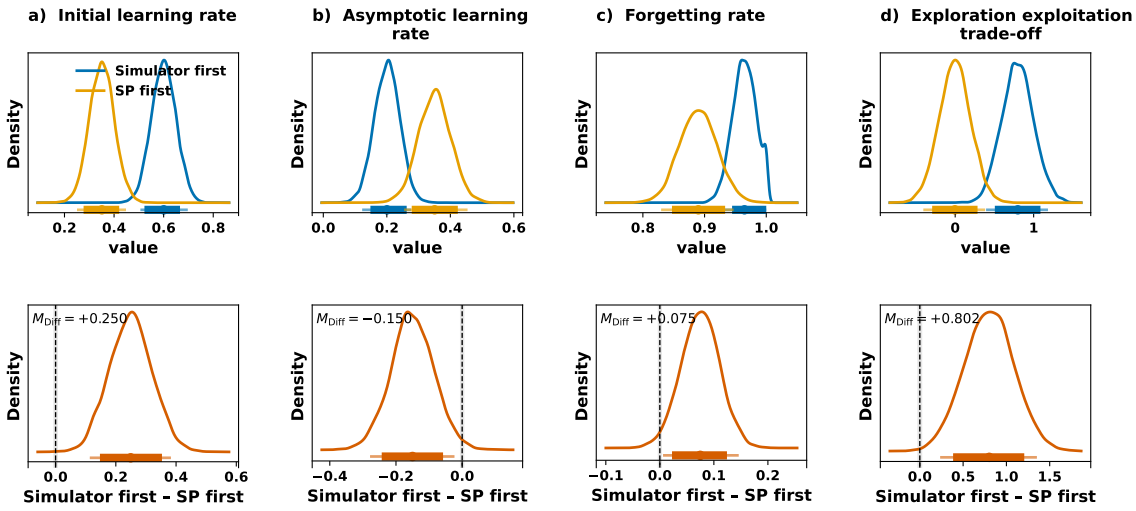



Figure 3: Posterior parameter distributions for Dataset I comparing the Simulator-first and SP-first groups. The panels report the four core active-inference parameters: initial learning rate ($\eta_0$), asymptotic learning rate ($\bar{\eta}$), forgetting rate ($\omega$), and exploration–exploitation progression ($c$). The top panels show group-level posterior densities, while the bottom panels show the posterior distribution of the group mean difference.

model weights. Under DIC, the FEP model was preferred for 70 of the 83 participants, while under WAIC the result was almost identical at 69 of 83. The superiority of the FEP model was not confined to one narrow subgroup, but generalised across simulator-first and stated-preference-first participants as well as across broader behavioural subtypes. Here, the key point is that the active inference architecture was not merely more interpretable; it also provided better statistical support for most individuals' observed route-learning behaviour.

*Dataset II.* The second experiment provides a more demanding test because it combines three route environments with an explicit information manipulation. Here the free-energy model again performed well at both the aggregate and individual levels. Posterior predictive checks, reported in Figure 4, show that the model captured both the time path of fast-route choice and the decline in route switching across the three scenarios. In the Fast & Safe and Low-Risk settings, the informed group was already close to the empirical ceiling from the earliest trial bins, while the control group improved more gradually. In the Fast & Risky setting, where the faster route was also much more variable, the model reproduced the more prolonged exploratory switching of uninformed participants before eventual convergence.

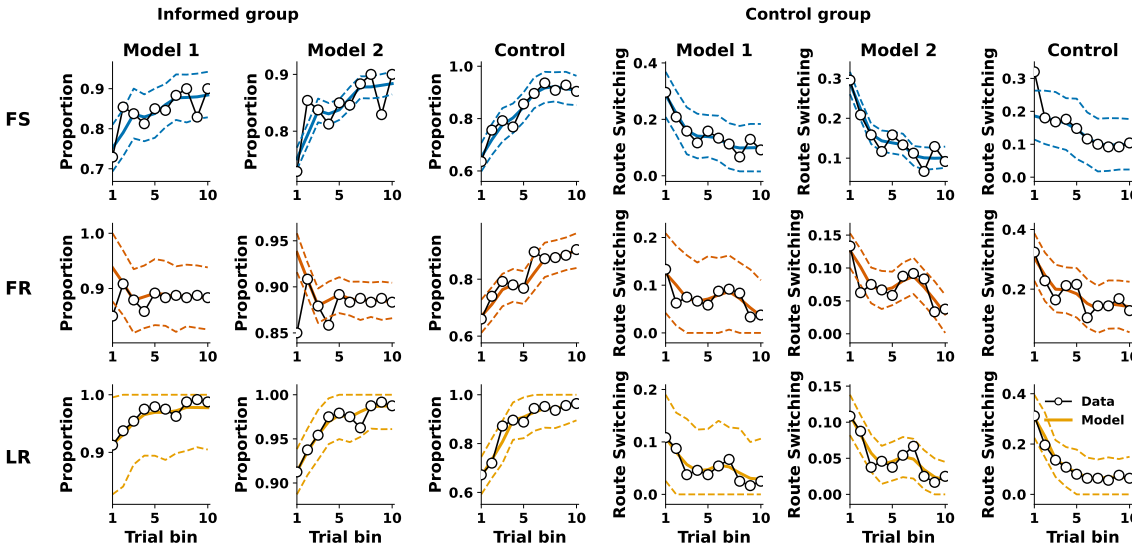



Figure 4: Posterior predictive checks for Dataset II across the Fast & Safe, Fast & Risky, and Low-Risk scenarios.

A second key result in Dataset II concerns how prior information is best represented in the model. For the informed group, we compared two alternative implementations. Model 1 treated the supplied information mainly as a contextual bias on decisions, whereas Model 2 treated it as a genuinely informative prior within the generative model. The comparison clearly favoured Model 2. Across scenarios, DIC weights selected Model 2 for 17–18 of the 24 informed participants, and WAIC selected it for 15–17 participants, with average posterior model probabilities of approximately 0.71 under DIC and 0.68 under WAIC. This result is substantively important: it suggests that informed participants were not merely nudged toward one route by a verbal cue, but instead incorporated the announced travel-time ranges into their internal beliefs in a way that is more consistent with Bayesian prior formation.

The informed–control contrasts in the core FEP parameters also followed the expected direction across all three scenarios. Informed participants showed lower learning plasticity, higher memory retention, and a steeper progression toward exploitative choice, with the strongest separation appearing in the Fast & Risky condition. In that setting, the difference in initial learning rate was particularly pronounced, indicating that uninformed participants had to engage in more active empirical sampling before stabilising their beliefs. This pattern is psychologically plausible. When meaningful prior information is available, epistemic value is reduced from the outset, so there is less need to probe the environment through repeated switching. Without such information, travellers continue to explore for longer, especially when the potentially superior route is also the most volatile one. The free-energy model therefore captures a core behavioural regularity that is highly relevant for transport applications: information can substitute for experience, but its value depends strongly on the reliability structure of the environment.

This interpretation is reinforced by the comparison with the RL benchmark. As shown in Figure 5, the FEP model again received majority weight for most participants in every Dataset II scenario, with roughly 20–22 informed participants and 21–22 control participants favouring FEP under both DIC and WAIC. RL still captured a minority of individuals, particularly those with more weakly structured or persistently stochastic choice patterns, but the overall evidence remained clearly tilted toward active inference. The implication is that modelling route choice as expected free-energy minimisation is not only conceptually attractive; it also yields a measurable empirical advantage across distinct feedback and information conditions.

Taken together, the two case studies support a common substantive conclusion. Route choice under uncertainty is better understood as a process in which travellers jointly update beliefs, value informative experience, and gradually shift from epistemic sampling to stable exploitation as uncertainty declines. In Dataset I, this transition is strongly shaped by the salience of early feedback. In Dataset II, it is strongly shaped by whether travellers enter the task with informative priors and by the risk structure of the available routes. Across both datasets, the free-energy framework captures these effects within one coherent mechanism and does so more successfully than a conventional reinforcement-learning benchmark. For travel behaviour research, this is an encouraging finding, because it suggests that active inference can serve as a unified modelling basis for learning, uncertainty, and behavioural adaptation in repeated travel decisions.

## 5 Conclusions

This paper developed an active inference model of route choice under uncertainty and applied it to two complementary experimental datasets. Across both datasets, the results showed that the model captures meaningful variation in learning, exploration, and convergence, while also outperforming a benchmark reinforcement learning specification in most cases.

In Dataset I, the findings showed that the salience of early feedback shapes subsequent learning and route stabilisation. In Dataset II, they showed that prior information reduces exploration and leads to faster convergence, especially in more uncertain environments. Taken together, the results suggest that route choice under uncertainty is better understood as a process of joint belief updating and adaptive decision-making rather than reward learning alone.

Overall, the Free Energy Principle provides a promising framework for modelling travel behaviour in dynamic and information-rich settings.

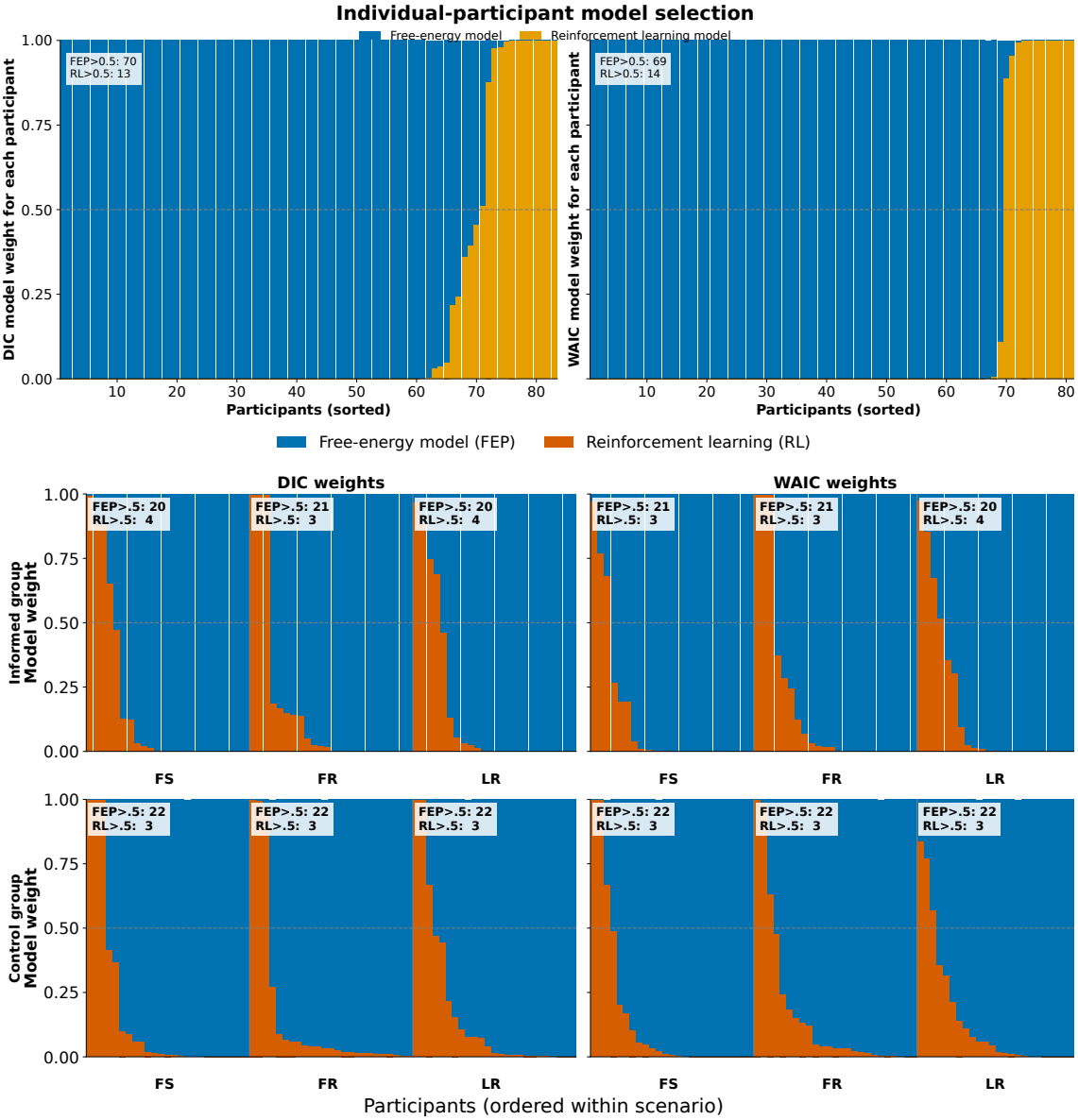



Figure 5: Individual-level model comparison between the free-energy model (FEP) and the reinforcement-learning benchmark (RL). Top: Dataset I. Bottom: Dataset II.

## Acknowledgements

Jamal Amani Rad, Stephane Hess and Thomas Hancock acknowledge the financial support by the European Research Council through the advanced grant 101020940-SYNERGY. Thomas Hancock's time was partially supported by Charisma Choudhury's UKRI Future Leader Fellowship [MR/T020423/1].

## References

Ben-Elia, E., Shiftan, Y., 2010. Which road do i take? a learning-based model of route-choice behavior with real-time information. Transportation Research Part A: Policy and Practice 44, 249–264.

Bogacz, M., Hess, S., Calastri, C., Choudhury, C.F., Mushtaq, F., Awais, M., Nazemi, M., Eggermond, M.A.B.V., Erath, A., 2021. Modelling risk perception using a dynamic hybrid choice model and brain-imaging data: Application to virtual reality cycling. Transportation Research Part C: Emerging Technologies 133, 103435.

Costa, L.D., Friston, K., Heins, C., Pavliotis, G.A., 2021. Bayesian mechanics for stationary processes. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 477, 20210518.

Engström, J., Wei, R., McDonald, A.D., Garcia, A., O'Kelly, M., Johnson, L., 2024. Resolving uncertainty on the fly: Modeling adaptive driving behavior as active inference. Frontiers in Neurorobotics 18, 1341750.

Engström, J., Bärgman, J., Nilsson, D., Seppelt, B., Markkula, G., Piccinini, G.B., Victor, T., 2018. Great expectations: A predictive processing account of automobile driving. Theoretical Issues in Ergonomics Science 19, 156–194.

Fayyaz, M., 2021. Route choice behaviour of drivers under risky and uncertain travel times—stated choices and simulated experiences. Phd thesis. The University of Sydney, Business School, Institute of Transport and Logistics Studies. Sydney, Australia.

Friston, K., 2009. The free-energy principle: A rough guide to the brain? Trends in Cognitive Sciences 13, 293–301.

Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. Philosophical Transactions of the Royal Society B: Biological Sciences 364, 1211–1221.

Friston, K., Kilner, J., Harrison, L., 2006. A free energy principle for the brain. Journal of Physiology - Paris 100, 70–87.

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., Pezzulo, G., 2015. Active inference and epistemic value. Cognitive Neuroscience 6, 187–214.

Friston, K.J., FitzGerald, T., Rigoli, F., Schwartenbeck, P., Pezzulo, G., 2017. Active inference: A process theory. Neural Computation 29, 1–49.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. Bayesian Data Analysis. 3rd ed., Chapman and Hall/CRC, Boca Raton, FL.

Hancock, T., Broekaert, J., Hess, S., Choudhury, C., 2020. Quantum choice models: A flexible new approach for understanding moral decision-making. Journal of Choice Modelling 37, 100235.

Hancock, T., Hess, S., Choudhury, C., 2018. Decision field theory: Improvements to current methodology and comparisons with standard choice modelling techniques. Transportation Research Part B: Methodological 107, 18–40.

Hancock, T., Hess, S., Marley, A., Choudhury, C., 2021. An accumulation of preference: Two alternative dynamic models for understanding transport choices. Transportation Research Part B: Methodological 149, 250–282.

Rad, J.A., Hess, S., Hancock, T.O., Bliemer, M., Beck, M., Fayyaz, M., Ben-Elia, E., 2025. Where does this road go? insights from reinforcement learning models on route choice behaviour. Preprint on PsyArXiv.

Schwartenbeck, P., FitzGerald, T., Dolan, R., Friston, K., 2013. Exploration, novelty, surprise, and free energy minimization. Frontiers in Psychology 4, 710.

Smith, R., Friston, K.J., Whyte, C.J., 2022. A stepbystep tutorial on active inference and its application to empirical data. Journal of Mathematical Psychology 107, 102632.

Wei, R., McDonald, A.D., Garcia, A., Markkula, G., Engstrom, J., O'Kelly, M., 2025. Learning an active inference model of driver perception and control: Application to vehicle car-following. arXiv:2303.15201. preprint on arXiv.

Wong, K.F., Wang, X.J., 2006. A recurrent network mechanism of time integration in perceptual decisions. Journal of Neuroscience 26, 1314–1328.