universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Decision-Making Between Humans and Artificial Agents: A Gamified Experimental Study of the Exploration–Exploitation Trade-Off

verfasst von | submitted by

Nicole Wimmer BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 840
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Psychologie
Betreut von | Supervisor:	Univ.-Prof. Dr. Frank Scharnowski MSc
Mitbetreut von | Co-Supervisor:	Dr. Filip Melinščak BSc MSc

Acknowledgements

I would like to sincerely thank my co-supervisor Filip for taking the time to always send me detailed and helpful feedback, explain concepts, and sending me so many interesting resources. To the people from the Methods of Psychology Unit I would like to thank Frank for giving me a desk to write my master's thesis at, and Nadja for asking for the desk on my behalf, thanks to Dominik for all the help with the statistical analysis, and to Daniel for always taking the time to answer questions.

I am also very grateful and would like to thank my family and friends for helping me keep my sanity during stressful times. My deepest thanks to my mum, for all the support in the last few days before the submission of my master's thesis, to my dad for keeping my external motivation high with the prospect of a new coffee machine, and to my grandparents for showing interest in what I am doing.

Finally, I am thankful to the participants in the study for all the kind feedback about the experiment, and to everyone who took the time to test it and share their thoughts.

Content

Abstract

Human-machine collaboration increasingly shapes everyday decision-making, yet its influence on a fundamental component of decision-making under uncertainty, the trade-off between exploration and exploitation, remains empirically underexplored. This thesis embeds an artificial agent into a spatially correlated multi-armed bandit task to examine how algorithmic peers affect human search strategies. Two questions are addressed: how the presence of artificial agents affects performance and exploration–exploitation behaviour in a structured search task, and how interindividual differences in exploration-related traits (impulsivity, openness, and curiosity) moderate these relationships. Forty-two participants completed a within-subjects gamified online experiment crossing decision horizon length, landscape structure and search condition (solo, social information bot, competitive bot). Bots did not exert a uniform effect on human search, although their presence tended to shift participants toward exploitation. This shift was associated with higher performance in long-horizon tasks, where search capacity was ample, but with lower performance in short-horizon tasks, where the reduced exploration may have left the landscape insufficiently sampled. Trait-level individual differences did not reliably moderate these human–bot interactions. These exploratory findings indicate that the effect of artificial agents on human decision-making in the employed search task is context-dependent, and provide methodological groundwork for confirmatory follow-up studies and formal computational modelling of human-machine joint search.

Introduction

Decision-Making in the Age of Artificial Collaborators

Across an increasingly diverse range of domains, individuals act not only alongside other people but in collaboration with artificial systems that recommend or co-produce choices once made by humans alone. In governance, machine collaborators are integrated into administrative and policy decisions (Van Rooy, 2024). In aviation, cognitive computing systems form intelligent teams with pilots in safety-critical environments (Dormoy et al., 2021). In finance and the broader corporate sector, machine learning shapes managerial choices and marketing workflows (Haesevoets et al., 2021; Hesel et al., 2022; Li et al., 2023; Semeraro et al., 2023), and in clinical contexts, AI-based decision-support systems are emerging as collaborators in mental-health diagnostics and care (Popat & Ive, 2023).

Given how widely such systems are now deployed, understanding how the presence of artificial agents alters human decision-making is an increasingly relevant question for psychology.

To investigate how artificial agents reshape human decision-making, a paradigm is needed that captures decision-making in a tractable form. Search problems offer such a paradigm. A search problem describes a situation in which an agent must allocate limited resources across a set of options whose values are initially uncertain, deciding both which options to sample and when to stop sampling in favour of acting on what has been learned. Because this structure abstracts away domain-specific content, search problems make it possible to compare decisions that arise in very different contexts but share the same underlying cognitive structure, such as choosing which restaurant to visit, where to invest, or which job applicant to interview. The case for studying decision-making through search problems is strengthened by the observation that the same structure recurs across cognitive domains. In decision research, visual attention, animal foraging, and memory retrieval, agents repeatedly face the problem of distributing effort across options whose value is uncertain (Hills, 2017). Many features of human cognition appear to derive from mechanisms that initially evolved to support spatial foraging, such that cognitive processes for searching memory or attention may use the same machinery that once guided the search for food and resources in physical space (Hills, 2017).

At the heart of every search problem lies the exploration–exploitation trade-off. A person must choose between exploiting an option whose value is already known and exploring alternatives whose value is uncertain but potentially higher (Hills, 2017; Hills et al., 2015). One could return to a restaurant one already likes, or try a new one that might prove even better, at the risk of disappointment. The decision turns on what the person expects to find locally relative to the anticipated opportunity cost of searching elsewhere (Hills, 2017). Because the exploration–exploitation dilemma recurs across cognitive domains and decision contexts, it has become a central framework for understanding decision-making behaviour more broadly (Hills et al., 2015; Wu et al., 2018).

Framing decision-making as a search problem therefore offers an entry point into the question raised above. If everyday choices are at base instances of exploration and exploitation, then the influence of artificial collaborators on human decision-making can be examined by asking how their presence shifts this fundamental balance.

The Structure of Human Search

Multi-armed bandits. Studying how people navigate the exploration-exploitation trade-off in varying situations requires experimental paradigms that recreate the essential features of search under uncertainty in a controlled setting. This consists of a set of options whose values are initially unknown, a limited amount of resources that enable choices to be made, and visible feedback so that each choice can be informed by the last and learning can occur. These features jointly reproduce the core demand of a search problem, which is learning the value of options through experience. The paradigm that dominates research on this question is the multi-armed bandit task (Robbins, 1952; Schulz et al., 2020; Steyvers et al., 2009; Witte et al., 2024; Wu et al., 2018). The task is named after the metaphor of a gambler standing before a row of slot machines, each with its own independent and initially unknown payoff distribution (Schulz et al., 2020; Wu et al., 2018). On each round, the participant selects one option, an "arm", observes the resulting reward, and uses this information to inform subsequent choices, with the overall goal of maximising rewards (Witte et al., 2024). Because no prior information is provided about which arm yields the highest expected payoff, participants must learn the underlying reward structure through experience while simultaneously exploiting whichever arm currently appears most profitable.

The paradigm thus operationalises the exploration–exploitation dilemma in a form that is at once tractable for the participant and tightly controlled for the experimenter. The experimenter sets the reward distributions in advance and records every choice and outcome, while the participant faces a manageable sequence of decisions (Witte et al., 2024; Wu et al., 2018). Rewards can be drawn from a parametric distribution such as a Gaussian, which allows the means and variances of the arms, and therefore the difficulty and informativeness of the environment, to be manipulated systematically (Witte et al., 2024). A reason bandit tasks have proved productive for the study of cognition is that the optimal solution is, in the general case, computationally intractable (Schulz et al., 2020). Participants cannot rely on a closed-form rule for balancing exploration and exploitation. They must instead deploy heuristics. This is precisely what makes their behaviour informative. The strategies people adopt, the conditions under which they shift between exploring and exploiting, and the systematic deviations from optimality they show all provide a window onto the cognitive mechanisms that underlie decision-making in less controlled, real-world settings (Schulz et al., 2020).

A further finding from collective and sequential search is directly relevant to the present design. When search is distributed across agents, performance can be improved by a division of labour in which some agents specialise in exploration while others exploit what has been found. In collective search through spatially correlated reward landscapes, participants who can observe a partner's choices search more efficiently and earn more than those searching alone, because each can capitalise on the regions the other has already sampled (Naito et al., 2022). A parallel idea has been formalised in agent-based models of scientific communities, where a polymorphic population that mixes "mavericks" (agents who deliberately seek out unexplored regions of an epistemic landscape) with "followers" (agents who consolidate around approaches already found to be promising) outperforms a uniform population of either type (Weisberg & Muldoon, 2009). This "explore–exploit division of labour," in which widely exploring agents reveal high-value regions that others can then exploit, makes it likely that an exploratory partner may free a participant to exploit.

Generalisation. Standard multi-armed bandit tasks assume that the rewards on different arms are statistically independent. Experience with one option provides no information about any other. In real-world environments, however, this assumption is not always plausible. Options that share features, that lie close together in some abstract or physical space, or that are otherwise linked tend to yield similar outcomes (Wu et al., 2018, 2024). Since in addition to this exhaustively sampling every possibility is usually impossible, people rely on generalization. Rather than learning each option in isolation, they exploit the correlational structure of the environment to predict the value of options they have not yet tried (Wu et al., 2024). To capture this feature of real-world decision-making, researchers have extended the bandit paradigm to spatially-correlated bandit problems, in which the rewards on different arms are correlated according to some underlying pattern such as spatial proximity, feature similarity, or graph connectivity (Wu et al., 2018, 2024). This correlated structure provides traction for generalization. It allows participants to direct search toward regions of the option space that are likely, on the basis of prior observations, to contain high rewards (Wu et al., 2024).

Landscapes. One way to implement structured environments is to embed the bandit task in a two-dimensional grid, where each cell represents an option that can be chosen. Rewards vary across the grid according to a spatially correlated function. The resulting landscape with initially unknown payoff is shown to participants, who then search it for high-reward

regions, much as an animal might forage across a physical environment (Wu et al., 2018). The spatial framing is not incidental. Hills et al. (2008) suggest that spatial search and abstract cognitive search share a common neural basis. The processes by which one moves through physical environments appear to be similar to the movement through abstract conceptual spaces such as memory or semantic networks. On this view, grid-world bandit tasks engage the same control mechanisms that underlie cognitive search more generally.

Social information. In real-world search, information about the value of options does not come from personal experience alone. Other agents such as friends, colleagues and competitors also sample the environment. Their behaviour conveys information that can guide one's own choices. Recent work has begun to integrate this social dimension into the bandit framework, showing that social information is a valuable resource for decision-making but one that also introduces new strategic tensions. Peers can act simultaneously as sources of information and as competitors for the same limited rewards (Wu et al., 2023). When incorporated into search environments, social information has been shown to function as a form of exploration in its own right, partially substituting for individual uncertainty-directed sampling and helping participants converge on high-value regions of the landscape more efficiently than they would alone (Naito et al., 2022; Witte et al., 2024). Participants are sensitive to the choices of others and integrate them into their own search behaviour (Witt et al., 2023). The competitive face of social information is, by comparison, less well understood. The existing evidence is mixed and sparse. In some competitive contexts, agents have been reported to adopt cautious, conservative strategies rather than bolder exploration (Liu, 2023). An experimental study found no evidence that competition increased participant effort, contrary to prior expectations (Tiokhin & Derex, 2019). How competitive framings shape exploration and exploitation, and under what conditions, remains a largely open question.

Individual differences. A final consideration is that exploration behaviour is not uniform across individuals. People might differ systematically in how readily they sample unknown options, and these differences are linked to broader features of cognition and personality. Impulsivity, for example, has been associated specifically with value-free random exploration, a type of exploration characterised by choosing among options without regard to their estimated value, rather than with “directed” exploration that aims to strategically reduce uncertainty about the underlying reward pattern (Dubois & Hauser, 2022). Wilson

et al. (2014) showed that both value-free random exploration and directed exploration increase as the amount of available decisions lengthens. At the same time, exploration tendencies do not appear to constitute a single domain-general trait. Performance on one type of search task is only weakly predictive of performance on another, suggesting that exploration is shaped substantially by the structure of the task rather than by a stable underlying disposition (Von Helversen et al., 2018; Witte et al., 2024).

From human peers to artificial agents. The agents that explore a search environment need not be human. As artificial systems increasingly co-produce decisions, they too become a source of social information. They are able to sample the environment, make choices, and generally provide information, that a human participants can include into their own search. This raises a question the social-search literature has only begun to address. Does information from an artificial agent shift exploration and exploitation? How people relate to artificial agents is unlikely to be uniform, as exemplified by two competing tendencies suggested by the human-computer interaction literature. On one hand, people often display algorithm appreciation, weighting advice more heavily when they believe it comes from an algorithm than from a person (Logg et al., 2019). On the other hand, a substantial literature documents algorithm aversion. People lose confidence in an algorithmic advisor more quickly than in a human one after observing it making mistakes, and consequently under-use it even when it outperforms human judgement (Dietvorst et al., 2015). Together, these opposing findings establish that awareness of collaborating with an artificial agent changes how people take in social information. Artificial agents can be perceived as a useful tool that augments one's own decision-making, or as a competitor for tasks, resources, or rewards. These framings may have opposite behavioural consequences. An agent perceived as a tool may shift exploration and performance in a quite different direction from the same agent perceived as a competitor. Whether the competitive face of social information operates as it does among humans when the competitor is artificial is even less well understood.

The Present Gap

Despite the growing prevalence of human-machine collaboration and a mature literature on the exploration–exploitation trade-off, the intersection of these two domains remains largely unexamined. Three gaps are especially salient.

First, although structured-search paradigms have characterised in detail how people balance exploration and exploitation when searching alone (Wu et al., 2018) and alongside human peers (Naito et al., 2022; Witte et al., 2024), they have not been used to investigate how an artificial agent impact this balance. Social information has been treated almost exclusively as something supplied by other humans, leaving open whether an algorithmic peer that samples the same environment shifts human search in comparable ways. Given the opposing tendencies of algorithm appreciation and algorithm aversion (Dietvorst et al., 2015; Logg et al., 2019), there is no a priori reason to assume that information from a bot is integrated as human social information would be.

Second, the competitive face of social information is poorly understood even among humans, and the existing evidence is mixed and sparse (Liu, 2023; Tiokhin & Derex, 2019). Whether framing an artificial agent as a competitor rather than a benign source of information changes how people explore and exploit has, to our knowledge, not been tested within a controlled structured-search task. The same bot behaviour may carry very different behavioural consequences depending on whether it is construed as a tool or a rival, yet this distinction has not been isolated experimentally.

Third, exploration is known to vary across individuals and to be linked to traits such as impulsivity (Dubois & Hauser, 2022), but it is unclear whether such dispositions shape how people respond to an artificial collaborator. Because exploration tendencies appear to be substantially task-dependent rather than reflecting a single domain-general trait (Von Helversen et al., 2018; Witte et al., 2024), whether trait-level differences moderate human–bot interaction in a search environment is an open empirical question.

The Present Study

The present study addresses these gaps by embedding an artificial agent into a spatially correlated multi-armed bandit task, manipulating both the presence of bots and their framing as social-information or competitive partners, and testing whether exploration-related traits moderate the resulting effects. We fit a set of linear mixed-effect models to investigate how individual differences and task characteristics jointly shape performance and exploration–exploitation behaviour. The analysis focuses on three central task features hypothesised to shape behaviour: the number of decisions available, the structure of the landscape, and the level of competition. In addition to their direct effects, we examine whether these features interact, recognising that performance may depend not only on

isolated factors but on how environmental constraints combine. To capture individual variability, we include questionnaires on openness, impulsivity, and curiosity as covariates, testing whether these traits moderate the relationship between task conditions and observed behaviour.

Additionally, we aim to replicate key findings from the structured-search literature, particularly those reported by Wu et al., (2018). We expect that participants will obtain higher rewards in smooth than in rough environments, and that short decision horizon length will yield levels of performance comparable to long ones. Beyond replication, we extend the paradigm by manipulating two additional task factors, social information and the level of competition introduced through artificial agents, and by adding two further landscape types. The task is implemented as a gamified online experiment.

Significance and Contribution

The present study is exploratory. Its aim is to fit statistical models that characterise how the presence and framing of an artificial agent affect human performance and search behaviour, generating hypotheses that can be tested in a confirmatory sample. In this sense it offers methodological groundwork for the formal computational modelling of human–machine interaction in social search problems. Its principal contribution is to introduce artificial agents into a spatially-correlated bandit task that has previously been studied with human participants alone. This allows the influence of an algorithmic peer, both its mere presence and its framing as a cooperative or competitive partner, to be examined systematically. In doing so, the study connects two literatures that have largely developed in parallel: the work on exploration and exploitation in structured search, and the work on how people respond to algorithmic collaborators.

Finally, because the exploration–exploitation trade-off recurs across cognitive domains and decision contexts (Hills, 2017; Hills et al., 2015), insight into how artificial agents shift this balance carries beyond any single application. The findings bear on decision-making in the many settings now populated by AI systems, from governance and healthcare to finance, aviation, and everyday consumer choice, while at the same time illuminating the basic search mechanisms that underpin human behaviour more generally.

Methods

Operationalisation of Research Questions

The study is guided by two research questions:

RQ1. How does the presence of artificial agents affect performance and exploration-exploitation behaviour in a structured search task?

RQ2. How do interindividual differences in traits related to exploration moderate the relationships identified in RQ1?

To address these research questions, we defined the following dependent variables (DVs), independent variables (IVs), and covariates.

Dependent variables. Average reward per round (performance) and the exploration-exploitation ratio (behaviour).

Independent variables. Decision horizon length (number of clicks available), landscape type (structure of the payoff environment), and search condition (presence or absence of bots, and competitive vs. social information framing).

Covariates. Questionnaire scores capturing individual traits related to exploration (impulsivity, openness, curiosity).

Participants

We recruited 53 participants from the online platform Prolific. Data collection continued until the planned recruitment target was reached. As an exploratory study, no formal a priori power analysis was conducted, and the sample size was determined by practical constraints on data collection. Of the recruited participants, 11 were excluded based on predefined criteria (see Data Cleaning below). The final sample comprised 42 participants (19 female, 23 male; mean age = 34.17, range 22–63). The sample was ethnically diverse and included participants from 16 nationalities, with the largest groups from South Africa, the United Kingdom, India, and Mexico. Half of the participants held an undergraduate degree, with the remainder holding graduate, secondary, or technical qualifications. Full sociodemographic characteristics can be found in Table 1.

Table 1

Sociodemographic characteristics of participants

	n	%	*Mean*	*Min*	*Max*
Age			34.17	22	63
Gender					
Female	19	45			
Male	23	55			
Ethnicity simplified					
Asian	10	23.8			
Other	9	21.4			
Black	8	19			
White	8	19			
Mixed	7	16.7			
Highest education level					
Undergraduate degree (BA/BSc/other)	21	50			
Graduate degree (MA/MSc/MPhil/other)	10	23.8			
High school diploma/A-levels	6	14.3			
Technical/community college	3	7.1			
Secondary education (e.g. GED/GCSE)	2	4.8			
Nationality					
South Africa	10	23.8			
United Kingdom	8	19			
India	4	9.5			
Mexico	4	9.5			
United States	3	7.1			
Canada	2	4.8			
Chile	2	4.8			
China	1	2.4			
Spain	1	2.4			
Indonesia	1	2.4			
Poland	1	2.4			
Turkey	1	2.4			
Morocco	1	2.4			
Greece	1	2.4			
Venezuela	1	2.4			
Cambodia	1	2.4			

Note. $N = 42$

Each participant received a base reimbursement of £3.42 and a performance-contingent bonus of up to £2.98 (average bonus earned: £2.07). The average completion time was approximately 34.41 minutes. The study and its methodology were approved by the Ethics Committee of the Faculty of Psychology at the University of Vienna. All participants provided informed consent through an online consent form presented at the beginning of the experiment.

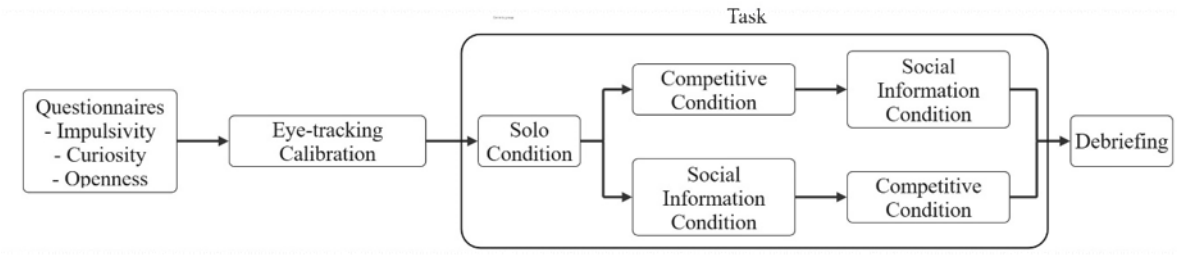
Data Cleaning

The raw dataset comprised 1,176 round-level observations (one record per participant per complete round). Rounds with a score of zero were removed, as these reflected errors caused by, for example, a poor internet connection rather than genuine search behaviour. We used a threshold of 0.9 × IQR, a stricter criterion than the conventional 1.5 × IQR rule. This more conservative approach was adopted given the less controlled nature of the online testing environment. A participant was excluded if they were flagged on any one of these three criteria, as these indicate implausibly fast or slow responding, anomalously low reward, or repetitive tile-clicking inconsistent with the task. This procedure removed 11 of the original 53 participants, yielding a final sample of 42 participants and 908 observations.

Procedure

Figure 1

Overview of the procedure



The study was implemented as an online experiment conducted remotely via participants' own devices. Upon accessing the study link, participants were first presented with an informed consent form. After providing consent, they proceeded through the following phases, as summarised in Figure 1:

Questionnaires. Participants first completed the three self-report questionnaires measuring impulsivity, curiosity, and openness to experience. These measures were

administered prior to the search task to avoid any influence of task performance on self-reported traits.

Eye-tracking. Participants were instructed through the calibration for Labvanced's webcam-based eye-tracking system (Kaduk et al., 2024). Eye-tracking data were collected during each condition. Their analysis is outside the scope of this thesis but will be used in subsequent work.

Solo exploration condition. Participants then received detailed on-screen instructions for the solo exploration condition, including example grids illustrating one landscape per type (kind, wicked, smooth, rough; see Materials for details). They were informed that their goal was to collect as many points as possible by clicking tiles within a grid-based landscape, that a larger number of points yielded a larger bonus payment, and that a click counter indicated the number of remaining clicks per round. In each round of the solo condition, participants were presented with a fully concealed landscape and could freely click concealed tiles to reveal them or re-click revealed tiles. Each time a tile was sampled, the observed value was the tile's true landscape value perturbed by Gaussian noise ($M = 0$, $SD = 10$), and the displayed value for that tile was updated to the mean of all observations so far. The value of the tile was then added to their score. The round ended when the click counter reached zero. After each round, participants received performance feedback consisting of (a) the raw score achieved, (b) the percentage of the maximum attainable score reached, and (c) the corresponding bonus payment in pence. Participants completed eight rounds of the solo condition and, after the final round, were shown a summary of the total bonus accumulated across all solo rounds.

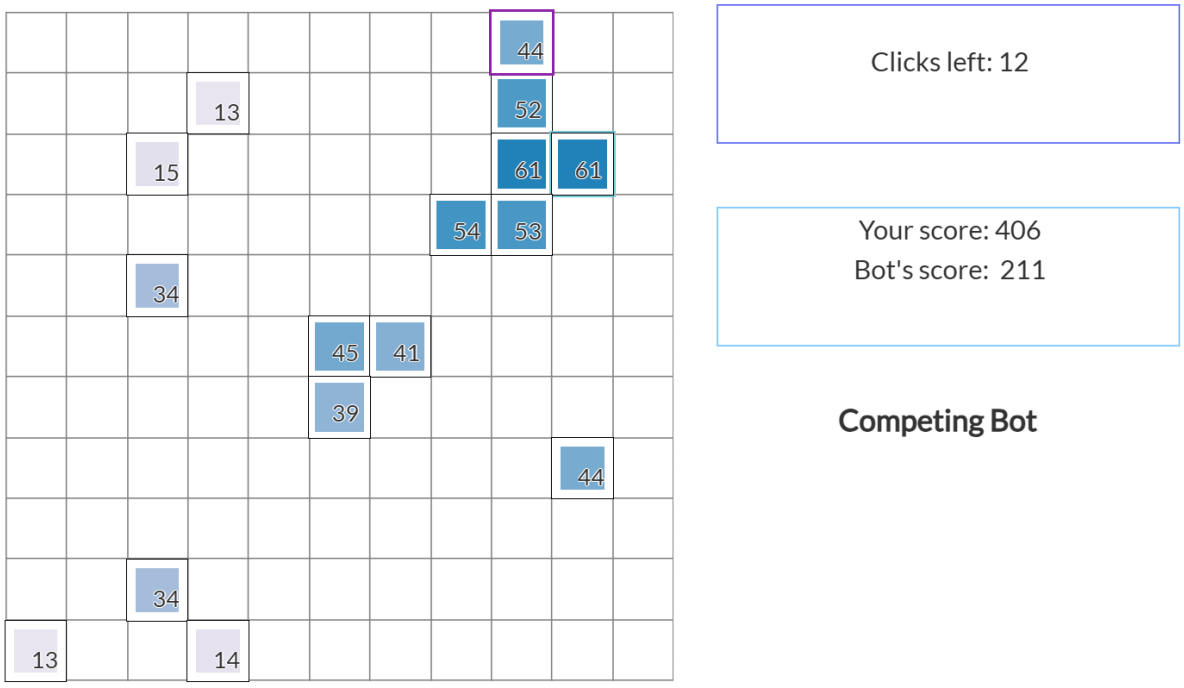
Competitive and social information condition. Following the solo condition, participants were randomly assigned to complete either the competitive or the social information condition first, with the remaining conditions completed afterward. In both conditions, participants were informed that they would take turns with a computer-controlled agent (the bot) and that the bot would receive half of their available clicks per round. A brief reminder of the core instructions was also provided. In the social information condition, participants were told that the bot would help them explore the landscape by revealing tiles and that the bot did not collect points. In the competitive condition, participants were informed that they were competing against the bot and would earn the bonus for a given round only if their score exceeded the bot's score in that round. Both scores were displayed on screen throughout each round (see Figure 2). The mechanics of the conditions were otherwise identical across conditions. On each turn the participant

clicked a covered tile, after which the bot automatically selected a tile. Standardised performance feedback was provided at the end of each round, as well as a summary of total bonus earnings at the end of each condition.

Debriefing. After completing all three conditions, participants were shown a final screen displaying the total bonus earned across the experiment in pounds. They then completed a final questionnaire reporting their search strategy and had the opportunity to describe their overall experience and report any technical issues they might have encountered during the search task. The information obtained in the debriefing was not included in the analysis, as it served to monitor the study quality.

Figure 2

Competitive condition interface as seen by the participants



Materials

Landscapes

The task was modelled as a spatially correlated multi-armed bandit problem (Wu et al., 2018), in which each tile of an 11 × 11 grid represented one of 121 arms. Selecting a tile yielded a noisy reward drawn from an underlying reward function, and participants could freely explore unrevealed tiles or re-select previously observed tiles. Rewards were spatially correlated, such that nearby tiles tended to yield similar outcomes. Four types of

reward landscape were used, varying in the structure and complexity of the underlying reward function. On each observation, normally distributed noise ($\varepsilon \sim N(0, 10)$) was added to the true tile value. Given the total reward range of 0 to 80 points, this noise level ensures that participants cannot easily determine whether they have found the global maximum. An example image of each type of landscape can be found in Figure 3.

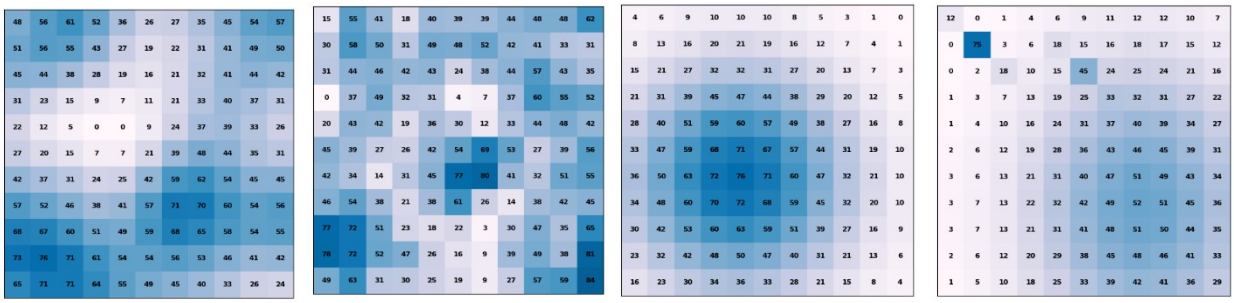
Smooth and rough environments. Following Wu et al. (2018), "smooth" and "rough" landscapes were sampled from a Gaussian-process prior with a radial-basis-function kernel. The length scale parameter controlled the rate at which reward correlations decayed over spatial distance, producing either gradually varying reward surfaces ("smooth") or more rapidly changing, jagged surfaces ("rough"). The two thus differ in their degree of spatial correlation.

*Kind environments. "*Kind" landscapes were newly generated for the present study and were designed to present a straightforward optimisation problem. Each landscape contained a single bivariate Gaussian peak placed at a random location on the grid, with equal spread in both dimensions ($\sigma_1 = \sigma_2 = 0.3$) and no correlation between dimensions ($r = 0.0$). This produced a single, smooth, symmetrical hill of rewards that decreased gradually and uniformly in all directions from the optimum. "Kind" environments were thus unimodal. A participant who found any moderately high-reward tile could reliably improve by searching nearby.

Wicked environments. "Wicked" landscapes were also newly generated and were designed to create a deceptive optimisation problem with multiple competing reward peaks. Each landscape was composed of three elements: a narrow, diagonally elongated global optimum ($\sigma_1 = \sigma_2 = 0.05$, $r = 0.9$); a broader but lower-valued local optimum ($\sigma_1, \sigma_2 \in [0.3, 0.5]$, $r = 0.0$) scaled to 75% of the global peak's height; and one to four additional small, narrow peaks placed at random locations. This structure was intended to be misleading: participants were likely to first encounter the broad local optimum, which covered a larger area of the grid and thus appeared to be the best region, while the true global optimum occupied only a small number of tiles and was difficult to discover without broader exploration.

Figure 3

Example reward landscapes for each of the four environment types



Note. From left to right: smooth, rough, kind, wicked. Darker cells indicate a higher reward.

Computer-Controlled Agent

The computer-controlled agent selected tiles uniformly at random (with replacement) from the full 11 × 11 grid. In both the social information and competitive conditions, half of the participant's available clicks were allocated to the artificial agent, which made one selection each time the participant clicked, so that participant and agent accumulated clicks at the same rate. The artificial agent's observations were subject to the same noise model as the participant's. Because the artificial agent selected randomly with replacement, it could revisit previously sampled tiles, and its selections updated the shared landscape visible to the participant.

Questionnaires

Three self-report questionnaires were administered to assess impulsivity, openness, and curiosity. Impulsivity was measured using the Barratt Impulsiveness Scale (BIS-11; Patton et al., 1995), a widely used 30-item instrument for assessing trait impulsivity. Openness was assessed using the Open-Mindedness domain scale from the Big Five Inventory-2 Short Form (BFI-2-S; Soto & John, 2017). Only the six items corresponding to the Open-Mindedness dimension were administered. Curiosity was measured using the Exploration subscale of the Curiosity and Exploration Inventory (CEI; Kashdan et al., 2004), which captures appetitive strivings for novelty and challenge and consists of four items. For each questionnaire, a trait score was computed by averaging the relevant item responses, following the original scoring instructions. These scores entered the analysis as continuous covariates.

Design

The experiment used a repeated-measures 2 × 4 × 3 factorial design. The three within-subjects independent variables were decision horizon length (2 levels: short = 20 clicks, long = 40 clicks), landscape type (4 levels: kind, wicked, smooth, rough), and search condition (3 levels: solo, social information bot, competitive bot). Both horizon lengths fell well below the total number of tiles on the 11 × 11 grid (121 tiles), ensuring that participants could never exhaustively search any landscape.

The three search conditions varied in their payoff structure and information environment. In the solo condition, participants accumulated points independently by clicking tiles and receiving the corresponding landscape values. In the social information bot condition, half of the available clicks were allocated to a bot that selected tiles at random. The bot's choices revealed tile values on the shared grid, thereby providing additional information. Its score did not count toward the participant's total. The competitive bot condition used the same click allocation. However, the bot accumulated its own score, and participants were informed that they would earn a bonus payment only if their final score exceeded that of the bot.

Each participant completed all 24 unique conditions (2 decision horizons × 4 landscape types × 3 search conditions). The solo condition always preceded the two bot conditions, ensuring that participants first familiarise with the task without social information. The order of the competitive and social information conditions was randomised independently for each participant to mitigate sequence effects. Questionnaires measuring individual differences in impulsivity, openness, and curiosity were administered prior to the main task. These three traits were included as continuous covariates in the analysis to test whether personality moderated the effects of the experimental manipulations.

Outcome Measures

Two primary outcome measures were derived from participants' search behaviour on the 11 × 11 grid landscape.

Average reward. Average reward, the performance DV, is defined as the mean proportion of the global maximum obtained per click within a round, expressed as a percentage. It captures how effectively participants identified high-value cells relative to the global optimum. For each click i, the true landscape value of the selected cell was

divided by the global maximum, and these ratios were averaged across the N clicks in a round:

$$P = \frac{1}{N}\sum_{i=1}^{N}\frac{v_i}{v_{max}}$$

where v_i is the true value of the cell chosen on click i, v_{max} is the global maximum of the landscape, and N is the total number of clicks in the round. The reported percentage score was obtained by multiplying P by 100 and rounding to the nearest integer, yielding a value between 0 and 100. A score of 100 would indicate selection of the globally optimal cell on every click. The per-round bonus was computed as:

$$bonus\ (in\ GBP) = P * \frac{total\ maximum\ bonus}{n_{tasks} * n_{rounds\ per\ task}}$$

Exploration–exploitation ratio. To quantify the degree to which participants explored the landscape versus exploited known high-value regions, we computed a ratio based on spatial distance to the best-known cell. After each click i, the cell with the highest observed value in the participant's search history was identified, and the Euclidean distance between the current click and this best-known cell was calculated:

$$distance = \sqrt{(row_i - row_{max})^2 + (col_i - col_{max})^2}$$

The ratio was then obtained by averaging these distances across all clicks and normalising by the maximum possible distance on the grid:

$$ratio = \frac{\frac{1}{N}\sum_{i=1}^{N}(distance)}{\sqrt{(grid\ size - 1)^2 + (grid\ size - 1)^2}}$$

This ratio ranges from 0 to 1, where values near 0 indicate predominantly exploitative behaviour (repeated selection of cells near the current best) and values near 1 indicate predominantly exploratory behaviour (selection of cells far from the current best). The reference cell updates dynamically as the participant discovers new values, so the measure reflects the explore-exploit trade-off relative to the participant's evolving knowledge of the landscape.

In addition to behavioural measures, eye-tracking data were collected during the task. The analysis of eye-tracking data is beyond the scope of the present thesis and is not reported here.

Statistical Analysis

Two separate sets of linear mixed-effects models (LMMs) were fitted, one for each dependent variable: average reward per round and the exploration–exploitation ratio (with higher values of the latter indicating a stronger tendency to explore rather than exploit). For each dependent variable, two nested models were estimated. The base model tested the experimental manipulations only, specifying the full three-way interaction among search condition (solo, competitive bot, social information bot), landscape type (smooth, wicked, rough, kind), and decision horizon length (long, short), together with all lower-order interactions and main effects. The extended model added the three personality traits hypothesised to moderate the effect of collaborative context, trait curiosity (CEI), openness (BFI-2-S), and impulsivity (BIS-11), by specifying their two-way interactions with search condition while retaining the full experimental design. The full coefficient estimates for each model can be found in the Appendix.

The random-effects structure included a by-participant random intercept and a by-participant random slope for search condition, accounting for both the repeated-measures design and individual variability in how participants responded to the social conditions. Reference levels were set to "solo", "smooth", and "long", so that fixed-effect coefficients represented deviations from the solo condition on a smooth landscape with a long decision horizon length. All models were estimated by restricted maximum likelihood (REML) using the BOBYQA optimizer, and convergence was verified from the optimizer status before further analysis.

Omnibus tests of the fixed effects were evaluated using Type III analyses of variance with Satterthwaite-approximated denominator degrees of freedom. To obtain valid Type III sums of squares, sum-to-zero (deviation) contrasts were applied to all categorical predictors prior to computing the ANOVA tables.

To address the research questions, two theory-driven contrasts over search condition were specified a priori. The solo-versus-bots contrast compared the solo condition against the average of the two bot conditions (weights: solo = +1, social information = −0.5, competitive = −0.5), testing whether the mere presence of a bot altered performance or exploration behaviour. The social-versus-competitive contrast compared the social information bot directly against the competitive bot (weights: social information = +1, competitive = −1, solo = 0), testing whether framing the interaction as cooperative versus competitive made a difference given identical bot behaviour. Both contrasts were

evaluated within each level of decision horizon length and, separately, within each level of landscape type. To unpack the search condition × trait interactions in Model 2, the two contrasts were re-estimated at three representative values of each trait (−1 SD, the mean, and +1 SD) using a simple-slopes procedure.

Given the exploratory nature of the study, no corrections for multiple comparisons were applied. Significance was evaluated at $\alpha = .05$ throughout. All analyses were conducted in R, called from Python through the rpy2 interface. Linear mixed-effects models were fitted using lme4, with Satterthwaite-approximated degrees of freedom supplied by lmerTest. Estimated marginal means, planned contrasts, and simple-slopes analyses were computed with emmeans. Omnibus Type III tests were additionally obtained with car::Anova and, for the base models, with afex::mixed using the Kenward–Roger approximation.

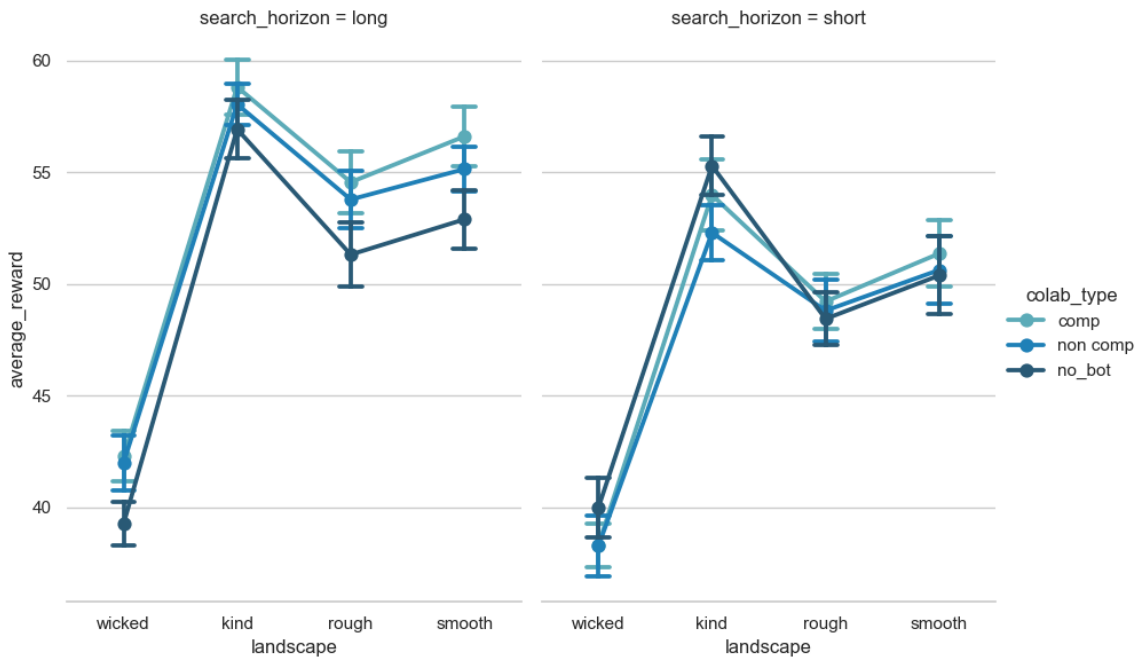
Results

Descriptive Statistics

Across the 908 retained observations, average reward per round was distributed around the low-to-mid 50s (on the 0–100 scale), while the exploration–exploitation (EE) ratio clustered at the low end of its 0–1 range, around 0.12–0.13. Because the ratio is normalised by the maximum grid distance, which is rarely attainable in practice, its absolute magnitude should not be read directly as a level of exploration. It is most informative as a relative measure for comparing conditions. The two outcome measures were themselves moderately and negatively associated. Rounds with higher reward were associated with lower exploration-exploitation ratios ($r = -.42$, $p = <.001$).

Figure 4

Average reward by landscape, decision horizon length, and search condition.

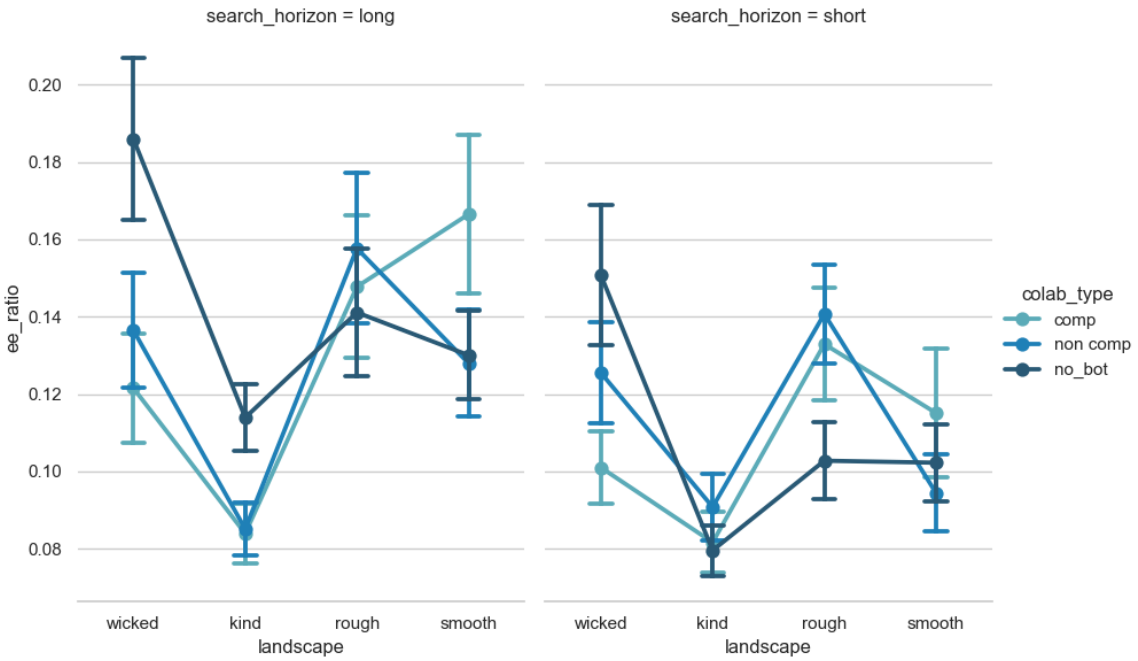


Note. Error bars represent 95% confidence intervals.

As shown in Figure 4, average reward varied markedly across the four landscape types. Performance was highest on kind landscapes, lowest on wicked landscapes, and intermediate on rough and smooth landscapes, with the ordering preserved across both horizon lengths. Reward was higher under the long than the short decision horizon length in every landscape. The three search conditions (solo, social information bot, competitive bot) produced closely overlapping means within most cells, but the lines separated in a horizon-dependent way: under the long horizon the two bot conditions tended to sit at or above the solo condition, whereas under the short horizon this ordering was less consistent.

Figure 5

Exploration-exploitation ratio by landscape, decision horizon length, and search condition.

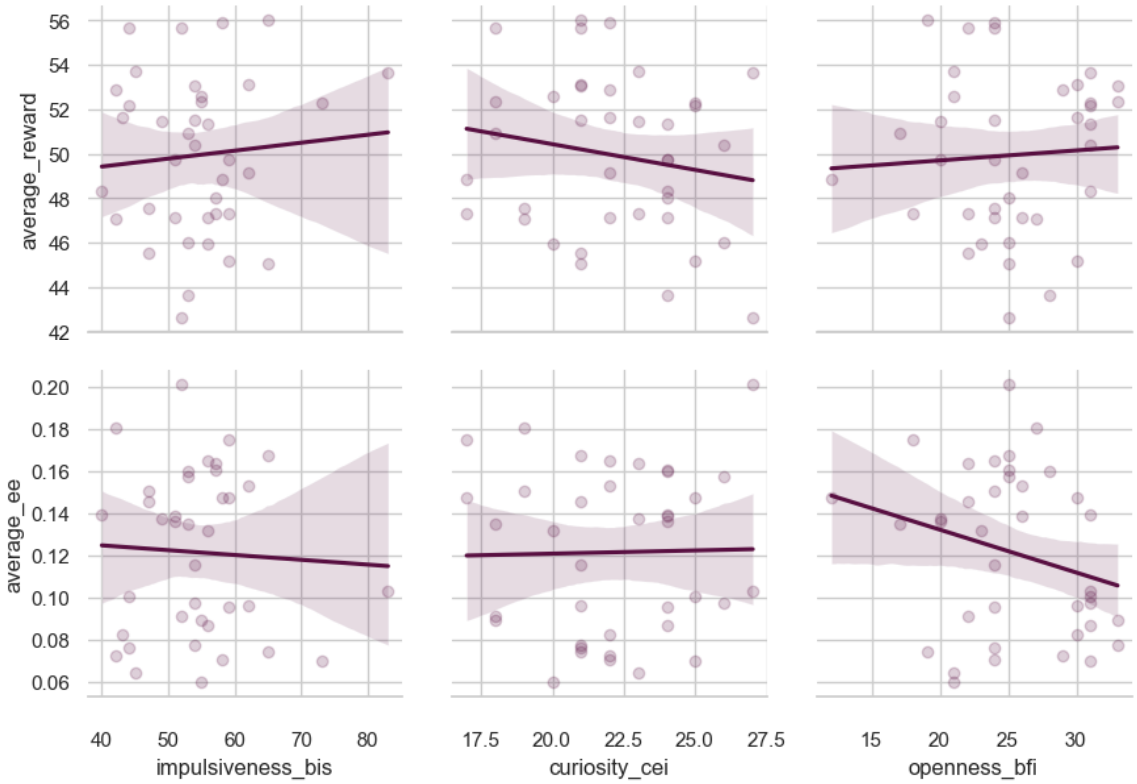


Note. Higher values indicate more exploration. Error bars represent 95% confidence: intervals.

Figure 5 displays the exploration-exploitation ratio across the same breakdown. Exploration was lowest on kind landscapes, where a single smooth peak made extended search unnecessary, and higher on the more structured wicked and rough landscapes. The exploration-exploitation ratio was generally higher under the long than the short horizon, reflecting the additional clicks available for sampling. As with reward, the search conditions overlapped considerably, though some separation was visible in a subset of conditions, where the solo and social information conditions tended toward more exploration than the competitive condition; this separation was not entirely consistent across landscapes.

Figure 6

Relationships between the self-report traits (impulsivity, curiosity, openness) and the dependent variables



Note. Shaded bands represent 95% confidence intervals around the regression lines.

Figure 6 plots each dependent variable against the three self-report traits. The regression slopes were shallow in most panels, anticipating the weak trait–outcome associations. The clearest visible trend was a negative relationship between openness and the Exploration-exploitation ratio (more open participants tended to explore slightly less), while the remaining trait–outcome slopes were close to flat.

Base Model: Experimental Factors

Average Reward in Relation to Experimental Factors

The base model predicted average reward from the three-way interaction of search condition, landscape, and decision horizon length, with by-participant random intercepts and random slopes for search condition. The intercept was 53.43 (SE = 1.16, t = 46.00, p < .001), representing the expected reward in the reference condition (solo, smooth landscape, long horizon).

The Type III ANOVA (Table 2) revealed that search condition did not exert a significant main effect on reward, $F(2, 47.73) = 0.79$, $p = .460$. Landscape had a strong main effect, $F(3, 887.38) = 266.10$, $p < .001$, as did decision horizon length, $F(1, 47.62) = 38.16$, $p < .001$. Of the interactions, the search condition × decision horizon length interaction was significant, $F(2, 887.33) = 11.74$, $p < .001$, indicating that the effect of bot presence on reward depended on the temporal structure of the task. The search condition × landscape interaction was not significant, $F(6, 886.73) = 0.80$, $p = .573$, nor was the landscape × decision horizon length interaction, $F(3, 887.35) = 1.11$, $p = .343$, or the three-way interaction, $F(6, 887.44) = 0.60$, $p = .734$.

Table 2

Average reward (base model): type III ANOVA

Effect	Sum Sq	Mean Sq	NumDF	DenDF	F	p
Search Condition	74.36	37.18	2	47.73	0.79	.460
Landscape	37630.95	12543.65	3	887.38	266.10**	< .001
Decision Horizon	1798.65	1798.65	1	47.62	38.16**	< .001
Search Condition:Landscape	225.09	37.51	6	886.73	0.80	.573
Search Condition:Decision Horizon	1106.93	553.47	2	887.33	11.74**	< .001
Landscape:Decision Horizon	157.23	52.41	3	887.35	1.11	.343
Search Condition:Landscape: Decision Horizon	168.60	28.10	6	887.44	0.60	.734

Note. Type III tests of fixed effects. Sum Sq = sum of squares; Mean Sq = mean square; NumDF = numerator degrees of freedom; DenDF = denominator degrees of freedom; *F* = F value; *p* = p value. * $p < .05$. ** $p < .01$.

Planned contrasts decomposing the search condition × decision horizon length interaction (Table 3) revealed a crossover pattern. In long-horizon conditions, participants in bot-present conditions earned significantly more reward than solo participants: the solo-vs-bots contrast was −2.76 ($SE = 0.74$, $t = -3.74$, $p < .001$), and the social-vs-competitive contrast was −3.04 ($SE = 0.84$, $t = -3.60$, $p < .001$), indicating that the social information bot condition earned less than the competitive bot condition. In short-horizon conditions the pattern reversed: the solo-vs-bots contrast was positive at 1.53 ($SE = 0.74$, $t = 2.07$, p

= .041), meaning solo participants now outperformed bot-accompanied participants, while the social-vs-competitive contrast did not reach significance (estimate = 1.33, p = .120).

Table 3

Average reward (base model): contrasts by decision horizon

Contrast	Decision Horizon	Estimate	SE	df	t	p
Solo vs Bots	long	-2.76	0.74	117.41	-3.74**	< .001
Social vs Comp	long	-3.04	0.84	118.18	-3.60**	< .001
Solo vs Bots	short	1.53	0.74	116.76	2.07*	.041
Social vs Comp	short	1.33	0.85	120.41	1.57	.120

Note. SE = standard error; df = degrees of freedom; t = t ratio; p = p value. * p < .05. ** p < .01.

Contrasts by landscape (Table 4) showed that the benefit of bots was most pronounced in rough environments. In rough landscapes, the solo-vs-bots contrast was −1.94 (SE = 0.97, t = −1.99, p = .047) and the social-vs-competitive contrast was −2.28 (SE = 1.12, t = −2.04, p = .042), indicating that competitive bots were associated with higher reward than social information bots. No significant differences between conditions emerged in smooth, wicked, or kind landscapes.

Table 4

Average reward (base model): contrasts by landscape

Contrast	Landscape	Estimate	SE	df	t	p
Solo vs Bots	smooth	-0.89	0.98	316.47	-0.91	.365
Social vs Comp	smooth	-1.25	1.13	326.28	-1.11	.269
Solo vs Bots	wicked	0.48	0.99	323.64	0.48	.631
Social vs Comp	wicked	0.60	1.13	328.96	0.53	.597
Solo vs Bots	rough	-1.94	0.97	307.00	-1.99*	.047
Social vs Comp	rough	-2.28	1.12	317.05	-2.04*	.042
Solo vs Bots	kind	-0.12	0.96	299.48	-0.13	.898
Social vs Comp	kind	-0.48	1.11	308.91	-0.44	.662

Note. *SE* = standard error; *df* = degrees of freedom; *t* = t ratio; *p* = p value. * $p < .05$. ** $p < .01$.

Exploration-Exploitation Ratio in Relation to Experimental Factors

The base model for the exploration–exploitation (EE) ratio mirrored the structure of the reward model. The intercept was 0.128 (*SE* = 0.01, *t* = 12.69, $p < .001$). From Table 5 we can see, that search condition did not have a significant main effect, $F(2, 47.26) = 0.25$, $p = .780$. However, both landscape, $F(3, 886.93) = 32.82$, $p < .001$, and decision horizon length, $F(1, 46.73) = 12.99$, $p = .001$, significantly predicted the exploration-exploitation ratio. Two interactions involving search condition reached significance: search condition × landscape, $F(6, 886.18) = 4.25$, $p < .001$, and search condition × decision horizon length, $F(2, 886.88) = 9.59$, $p < .001$. The landscape × decision horizon length interaction ($F = 2.08$, $p = .101$) and the three-way interaction ($F = 1.29$, $p = .260$) did not reach significance.

Table 5

Exploration–exploitation ratio (base model): Type III ANOVA

Effect	Sum Sq	Mean Sq	NumDF	DenDF	F	p
Search Condition	0.00	0.00	2	47.26	0.25	.780
Landscape	0.28	0.09	3	886.93	32.82**	< .001
Decision Horizon	0.04	0.04	1	46.73	12.98**	.001
Search Condition:Landscape	0.07	0.01	6	886.18	4.25**	< .001
Search Condition:Decision Horizon	0.05	0.03	2	886.88	9.59**	< .001
Landscape:Decision Horizon	0.02	0.01	3	886.91	2.08	.101
Search Condition:Landscape: Decision Horizon	0.02	0.00	6	886.92	1.29	.260

Note. Type III tests of fixed effects. Sum Sq = sum of squares; Mean Sq = mean square; NumDF = numerator degrees of freedom; DenDF = denominator degrees of freedom; *F* = F value; *p* = p value. * $p < .05$. ** $p < .01$.

Contrasts by decision horizon length (Table 6) showed that in long-horizon conditions, solo participants explored significantly more than bot-accompanied participants

(solo-vs-bots: estimate = 0.018, *SE* = 0.006, *t* = 3.14, *p* = .002; social-vs-competitive: estimate = 0.017, *SE* = 0.007, *t* = 2.55, *p* = .012). The positive social-vs-competitive estimate indicates that participants in the social information condition explored more than those in the competitive condition. In short-horizon conditions, neither contrast was significant (solo-vs-bots: *p* = .062; social-vs-competitive: *p* = .158).

Table 6

Exploration–exploitation ratio (base model): contrasts by decision horizon

Contrast	Decision Horizon	Estimate	SE	df	t	p
Solo vs Bots	long	0.02	0.01	120.40	3.14**	.002
Social vs Comp	long	0.02	0.01	125.35	2.55*	.012
Solo vs Bots	short	-0.01	0.01	119.70	-1.88	.062
Social vs Comp	short	-0.01	0.01	127.75	-1.42	.158

Note. *SE* = standard error; *df* = degrees of freedom; *t* = t ratio; *p* = p value. * $p < .05$. ** $p < .01$.

Contrasts by landscape (Table 7) indicated that in wicked environments the social-vs-competitive contrast was significant (estimate = 0.022, *SE* = 0.009, *t* = 2.50, *p* = .013), with social information participants exploring more than competitive participants. In kind environments the solo-vs-bots contrast reached significance (estimate = 0.015, *SE* = 0.008, *t* = 1.99, *p* = .048), with solo participants exploring more than bot-accompanied participants. No significant differences were found in smooth or rough landscapes.

Table 7

Exploration–exploitation ratio (base model): contrasts by landscape

Contrast	Landscape	Estimate	SE	df	t	p
Solo vs Bots	smooth	-0.00	0.01	327.55	-0.33	.745
Social vs Comp	smooth	-0.01	0.01	353.24	-1.56	.119
Solo vs Bots	wicked	0.01	0.01	334.83	1.90	.058
Social vs Comp	wicked	0.02	0.01	356.14	2.50*	.013
Solo vs Bots	rough	-0.01	0.01	317.93	-1.67	.096
Social vs Comp	rough	-0.01	0.01	343.58	-0.96	.340
Solo vs Bots	kind	0.01	0.01	334.57	1.99*	.048

Contrast	Landscape	Estimate	SE	df	t	p
Social vs Comp	kind	0.01	0.01	334.68	1.70	.091

Note. *SE* = standard error; *df* = degrees of freedom; *t* = t ratio; *p* = p value. * $p < .05$. ** $p < .01$.

Extended Model: Adding Personality Moderators

Average Reward in Relation to Experimental Factors and Personality Moderators

The extended model added interactions between search condition and three individual-difference variables: curiosity (CEI), openness (BFI-2-S), and impulsivity (BIS-11). The overall pattern of experimental effects was preserved (Table 8). Landscape remained the strongest predictor, $F(3, 904.59) = 261.50$, $p < .001$, followed by decision horizon length, $F(1, 47.73) = 40.07$, $p < .001$. The search condition main effect remained non-significant, $F(2, 45.38) = 0.18$, $p = .839$.

Among the personality variables, curiosity showed a marginal main effect, $F(1, 44.05) = 4.05$, $p = .050$, suggesting a trend toward lower reward among more curious individuals. Openness ($F = 0.85$, $p = .361$) and impulsivity ($F = 0.14$, $p = .706$) did not reach significance as main effects. Crucially, none of the search condition × personality interactions were significant: curiosity, $F(2, 45.27) = 0.64$, $p = .535$; openness, $F(2, 44.33) = 1.63$, $p = .207$; impulsivity, $F(2, 44.54) = 0.37$, $p = .692$.

Table 8

Average reward (extended model): Type III ANOVA

Effect	Sum Sq	Mean Sq	NumDF	DenDF	F	p
Search Condition	16.92	8.46	2	45.38	0.18	.839
Curiosity	194.63	194.63	1	44.05	4.05	.050
Opennessi	40.99	40.99	1	43.46	0.85	.361
Impulsiveness	6.94	6.94	1	43.68	0.14	.706
Landscape	37727.78	12575.93	3	904.59	261.50**	< .001
Decision Horizon	1926.89	1926.89	1	47.73	40.07**	< .001
Search Condition:Curiosity	61.08	30.54	2	45.27	0.64	.535
Search Condition:Openness	156.93	78.47	2	44.33	1.63	.207
Search Condition:Impulsiveness	35.70	17.85	2	44.54	0.37	.692

Note. Type III tests of fixed effects. Sum Sq = sum of squares; Mean Sq = mean square; NumDF = numerator degrees of freedom; DenDF = denominator degrees of freedom; *F* = F value; *p* = p value. * $p < .05$. ** $p < .01$.

As Table 9 shows, the horizon-specific contrasts reproduced the crossover seen in the base model. In long-horizon tasks, solo participants earned significantly less reward than those in bot-present conditions (solo vs. bots: −2.76, *SE* = 0.75, $t = -3.67$, $p < .001$), and the competitive condition outperformed the social-information condition (social vs. competitive: −3.04, *SE* = 0.86, $t = -3.54$, $p < .001$). Under short horizons the solo advantage re-emerged (solo vs. bots: 1.54, *SE* = 0.75, $t = 2.05$, $p = .043$), while the social-vs-competitive contrast was no longer significant (1.35, *SE* = 0.86, $t = 1.56$, $p = .121$).

Table 9

Average reward (extended model): contrasts by decision horizon

Contrast	Decision Horizon	Estimate	SE	df	t	p
Solo vs Bots	long	-2.76	0.75	105.85	-3.67**	< .001
Social vs Comp	long	-3.04	0.86	106.77	-3.54**	< .001
Solo vs Bots	short	1.54	0.75	105.78	2.05*	.043
Social vs Comp	short	1.35	0.86	108.76	1.56	.121

Note. *SE* = standard error; *df* = degrees of freedom; *t* = t ratio; *p* = p value. * $p < .05$. ** $p < .01$.

Broken down by landscape (Table 10), the search-condition contrasts were largely non-significant, indicating that the reward differences between conditions did not vary systematically across landscape types. The only reliable effect emerged in the rough landscape, where the competitive condition outperformed the social-information condition (social vs. competitive: −2.26, *SE* = 1.13, $t = -2.00$, $p = .046$) and the solo-vs-bots contrast reached the significance threshold (−1.91, *SE* = 0.98, $t = -1.95$, $p = .050$). All contrasts in the smooth, wicked, and kind landscapes were non-significant (all $p > .27$).

Table 10

Average reward (extended model): contrasts by landscape

Contrast	Landscape	Estimate	SE	df	t	p
Solo vs Bots	smooth	-0.90	0.99	282.95	-0.90	.367
Social vs Comp	smooth	-1.25	1.14	292.96	-1.09	.274
Solo vs Bots	wicked	0.48	1.00	289.81	0.48	.631
Social vs Comp	wicked	0.61	1.15	295.40	0.53	.596
Solo vs Bots	rough	-1.91	0.98	274.35	-1.95	.050
Social vs Comp	rough	-2.26	1.13	284.23	-2.00*	.046
Solo vs Bots	kind	-0.12	0.97	267.32	-0.12	.905
Social vs Comp	kind	-0.48	1.12	277.03	-0.43	.668

Note. *SE* = standard error; *df* = degrees of freedom; *t* = t ratio; *p* = p value. * $p < .05$. ** $p < .01$.

Exploration-Exploitation Ratio in Relation to Experimental Factors and Personality Moderators

The extended model for the exploration-exploitation ratio added the same three personality moderators. The core experimental findings were replicated: landscape, $F(3, 940.20) = 31.82$, $p < .001$, and decision horizon length, $F(1, 49.59) = 12.38$, $p = .001$, remained strong predictors. Search condition again had no main effect ($F = 1.35$, $p = .268$). None of the three traits showed a significant main effect on the exploration–exploitation ratio (curiosity, $F(1, 49.02) = 0.29$, $p = .590$; openness, $F(1, 47.98) = 0.01$, $p = .908$; impulsivity, $F(1, 48.35) = 0.11$, $p = .741$). Among the personality moderators, openness showed a significant interaction with search condition, $F(2, 49.52) = 3.63$, $p = .034$, suggesting that individuals differing in openness may respond differently to bot presence in terms of exploration. However, the dedicated moderation contrasts for openness did not reach significance individually (solo-vs-bots: −0.046, $SE = 0.028$, $p = .101$; social-vs-competitive: −0.044, $SE = 0.031$, $p = .173$). Neither curiosity ($F = 0.80$, $p = .455$) nor impulsivity ($F = 1.83$, $p = .171$) showed significant interactions with search condition.

Table 11

Exploration–exploitation ratio (extended model): Type III ANOVA

Effect	Sum Sq	Mean Sq	NumDF	DenDF	F	p
Search Condition	0.01	0.00	2	50.84	1.35	.268
Curiosity	0.00	0.00	1	49.02	0.29	.590
Openness	0.00	0.00	1	47.98	0.01	.908
Impulsiveness	0.00	0.00	1	48.35	0.11	.741
Landscape	0.28	0.09	3	940.20	31.82**	< .001
Decision Horizon	0.04	0.04	1	49.59	12.38**	.001
Search Condition:Curiosity	0.01	0.00	2	50.59	0.80	.455
Search Condition:Openness	0.02	0.01	2	49.52	3.60*	.034
Search Condition:Impulsiveness	0.01	0.01	2	49.85	1.83	.171

Note. Type III tests of fixed effects. Sum Sq = sum of squares; Mean Sq = mean square; NumDF = numerator degrees of freedom; DenDF = denominator degrees of freedom; *F* = F value; *p* = p value. * $p < .05$. ** $p < .01$.

The horizon-specific contrasts for the exploration–exploitation ratio (Table 12) mirrored the reward pattern. Under long horizons, participants in bot-present conditions explored more than solo participants (solo vs. bots: 0.02, $SE = 0.01$, $t = 3.20$, $p = .002$), and the competitive condition explored more than the social-information condition (social vs. competitive: 0.02, $SE = 0.01$, $t = 2.56$, $p = .012$). Under short horizons neither contrast was significant (solo vs. bots: −0.01, $t = -1.90$, $p = .060$; social vs. competitive: −0.01, $t = -1.43$, $p = .155$).

Table 12

Exploration–exploitation ratio (extended model): contrasts by decision horizon

Contrast	Decision Horizon	Estimate	SE	df	t	p
Solo vs Bots	long	0.02	0.01	116.82	3.20**	.002
Social vs Comp	long	0.02	0.01	118.11	2.56*	.012
Solo vs Bots	short	-0.01	0.01	116.11	-1.90	.060

Contrast	Decision Horizon	Estimate	SE	df	t	p
Social vs Comp	short	-0.01	0.01	120.45	-1.43	.155

Note. *SE* = standard error; *df* = degrees of freedom; *t* = t ratio; *p* = p value. * $p < .05$. ** $p < .01$.

When broken down by landscape (Table 13), the search-condition differences in exploration were weak and inconsistent. Significant contrasts emerged only in two landscapes: the competitive condition explored more than the social-information condition on the wicked landscape (social vs. competitive: 0.02, $SE = 0.01$, $t = 2.53$, $p = .012$), and solo participants explored more than bot-present participants on the kind landscape (solo vs. bots: 0.01, $SE = 0.01$, $t = 1.99$, $p = .047$). The remaining contrasts, including all comparisons on the smooth and rough landscapes, did not reach significance (all $p > .05$).

Table 13

Exploration–exploitation ratio (extended model): contrasts by landscape

Contrast	Landscape	Estimate	SE	df	t	p
Solo vs Bots	smooth	-0.00	0.01	324.88	-0.32	.748
Social vs Comp	smooth	-0.01	0.01	337.07	-1.57	.117
Solo vs Bots	wicked	0.01	0.01	332.03	1.95	.052
Social vs Comp	wicked	0.02	0.01	339.87	2.53*	.012
Solo vs Bots	rough	-0.01	0.01	315.42	-1.69	.092
Social vs Comp	rough	-0.01	0.01	327.46	-0.96	.339
Solo vs Bots	kind	0.01	0.01	307.44	1.99*	.047
Social vs Comp	kind	0.01	0.01	319.06	1.68	.095

Note. *SE* = standard error; *df* = degrees of freedom; *t* = t ratio; *p* = p value. * $p < .05$. ** $p < .01$.

Discussion

The present study investigated whether the presence of bots in a spatial decision-making environment affects human performance and exploration-exploitation behaviour, and whether individual-difference traits moderate these effects. The results reveal that bots do not exert a simple, uniform influence on performance but instead interact with task structure in ways that can either help or hinder participants.

The Context-Dependent Impact of Bots

A central finding is the crossover interaction between search condition and decision horizon length. In long-horizon tasks, where participants had many clicks to search, those accompanied by bots earned significantly more reward than solo participants. Conversely, in short-horizon tasks, where opportunities for search were limited, bot presence was associated with lower performance relative to the solo condition. This pattern held across both competitive and social information bot conditions, although competitive bots were associated with an additional reward advantage in long-horizon tasks.

The exploration-exploitation ratio data illuminate this mechanism. In long-horizon conditions, solo participants explored significantly more than bot-accompanied participants. That is, bots shifted participants toward exploitation. Critically, this shift coincided with higher reward in the long-horizon condition, suggesting that the exploitation was well-directed. Bots appear to have provided informative social cues, through which options they selected and how those performed, that allowed participants to identify high-reward options without exploring as extensively themselves. This interpretation aligns with the explore-exploit division-of-labour account introduced earlier. An exploratory partner can free a participant to consolidate gains, which is adaptive when there is sufficient time to exploit what the partner reveals.

This bot-induced shift toward exploitation converges with evidence from outside the search-task literature. In a sequential risk decision-making study, Xiong et al. (2023) found that framing a machine as a genuine partner rather than a subordinate led human decision-makers to behave more conservatively, without impairing overall team performance. The present results extend that pattern to a structured-search setting and add an important qualification. A shift toward more conservative, exploitative behaviour is not uniformly beneficial or costly but depends on whether the task affords enough remaining opportunity to benefit from it.

The landscape-dependent contrasts add further texture. The bot-related performance advantage was most pronounced in rough environments, so environments characterised by misleading local optima. In these landscapes, the social information provided by bots appears to have been particularly valuable, perhaps by helping participants avoid getting trapped in local optima. When bots sample broadly, their behaviour provides information about the location of high-reward regions that would be costly for a solo participant to

discover independently. The finding that competitive bots yielded higher reward than social information bots in rough landscapes suggests that competitive pressure may have intensified participants' use of social information. Alternatively, the competitive framing may have added to the gamification aspect of the task and thus heightened task engagement, as seen in educational settings (Ho et al., 2022). It should be noted, however, that the reward contrast in wicked landscapes, which share some of the deceptive structure of rough landscapes, was not significant, so the advantage of bot presence in deceptive environments was not entirely consistent across landscape types. In smooth environments, where most exploration trajectories eventually lead to the global optimum, bot presence had no detectable effect. The task was sufficiently forgiving that social information added little value.

Individual Differences as Moderators

The three personality traits examined, curiosity, openness, and impulsivity, did not meaningfully moderate the effects of bot presence on either reward or exploration. The moderation contrasts were uniformly non-significant and the omnibus interaction tests were likewise null, with one exception. The search condition × openness interaction reached significance in the exploration-exploitation ratio. This omnibus interaction, however, was not localised by the planned contrasts that compared the bot conditions against the solo condition and the competitive bot against the social-information bot; neither of these contrasts was individually significant. The significant omnibus effect therefore does not correspond cleanly to either the bot-versus-solo distinction or the competitive-versus-social distinction, and is better read as a diffuse association between openness and exploration across conditions than as a targeted moderation of the bot effect.

Several explanations warrant consideration. First, the traits measured may not be the most relevant moderators of social influence in exploration–exploitation tasks. In a social context, traits such as susceptibility to social influence, conformity, or social-comparison orientation (Laursen & Faur, 2022) might be more proximal predictors of how individuals respond to bot-generated information. Second, the experimental manipulation may have been strong enough to override individual-level variation. If the presence of social information fundamentally alters the decision-making environment for all participants, dispositional differences in exploration tendency may become secondary.

Third, the sample size may have limited power to detect moderation effects, which typically require larger samples than main effects.

The marginal main effect of curiosity on reward (p = .050) is worth noting, even though it did not interact with search condition. If reliable, it would suggest that more curious individuals earn slightly less reward, perhaps because a dispositional tendency to explore leads them to sample novel options at the expense of exploiting known high-reward ones. It is possible that highly curious individuals intrinsically value information gain over reward maximization, leading them to prioritize exploration even when it is suboptimal for task performance (Murayama, 2022). Given the borderline significance and the absence of a corresponding effect on the exploration-exploitation ratio, this finding should be treated as a hypothesis for future investigation.

Replication of Prior Findings

The present data successfully replicated the robust effect of landscape structure on reward, confirming that smoother environments facilitate higher performance, consistent with a large body of work demonstrating that spatial correlation structure is a fundamental determinant of decision quality. However, the finding from Wu et al. (2018) that short and long decision horizon length yield equivalent performance was not replicated. In the present study, short horizons were associated with significantly lower reward. It might be argued that the significant search condition × decision horizon length interaction reflects a genuine modulation of the horizon effect by social context. It is important to be cautious, however, about attributing the failed replication solely to the introduction of bots. The short-horizon disadvantage is visible even when attention is restricted to the solo condition and to the smooth and rough landscapes that most closely match Wu et al.'s (2018) paradigm, which means the penalty is not produced entirely by the presence of artificial agents. A bot-based explanation can at most account for part of the discrepancy, for example, by amplifying the short-horizon penalty through premature exploitation, rather than for the existence of the penalty itself. Other contributors, including differences in sample composition (discussed below) cannot be ruled out and may be at least as important.

An interesting nuance is that the short-horizon penalty was attenuated in wicked environments. This suggests that in environments with deceptive reward structures, where extended exploration does not reliably identify the global optimum, the advantage of

additional clicks is diminished, because neither extended nor brief search guarantees convergence on the best outcome.

Limitations and Future Directions

The most salient limitation concerns the sample. The final sample comprised only 42 participants, a number that is modest for the detection of interaction and moderation effects in particular. Moderation effects of the kind tested in Model 2 typically require substantially larger samples than main effects to be detected reliably, and the uniformly null trait-level moderation results must therefore be interpreted with caution. The absence of significant moderation by curiosity, openness, and impulsivity may reflect a genuine absence of such effects, but it may equally reflect insufficient statistical power to detect them. The same concern applies to the marginal effects reported throughout, which cannot be distinguished from noise on the basis of the present data alone. The immediate priority for future work is therefore confirmatory replication. Because the present study is exploratory and uncorrected for multiple comparisons, the effects reported here are best treated as hypotheses rather than established findings. A pre-registered study with a substantially larger sample would clarify which patterns are robust, in particular the crossover interaction between search condition and horizon length, which is the central result and the one most in need of independent confirmation.

A related point concerns the composition of the sample rather than its size. The sample was deliberately diverse, with a balanced gender distribution and participants drawn from sixteen nationalities spanning several continents. While such diversity is ordinarily a strength, enhancing the generalisability of findings beyond the narrow undergraduate populations that dominate much of psychology (Henrich et al., 2010), it can also act as a disadvantage in a study of this scale. With only 42 participants distributed across many demographic and cultural backgrounds, any systematic variation associated with culture, language, education system, or familiarity with digital interfaces and gamified tasks is spread thinly across the sample. This heterogeneity introduces additional between-participant variance that the modest sample size is ill-equipped to absorb. The random-effects structure must accommodate considerable individual variability, and genuine effects may be obscured by noise arising from unmodelled demographic differences. Future work could resolve this tension either by recruiting a more homogeneous sample or, preferably,

by recruiting one large enough to model cultural and demographic variation explicitly rather than leaving it as unstructured noise.

This consideration bears directly on the failure to replicate the horizon-equivalence finding reported by Wu et al. (2018). The two samples may have differed in composition. The demographic and cultural makeup of Wu et al.'s participants is not reported in a way that permits direct comparison, and their nationality and ethnicity are unknown to us. It is therefore possible that differences in sample composition, rather than, or in addition to, the experimental manipulation, contributed to the divergent results. Without knowing the characteristics of the comparison sample, the present study cannot adjudicate between these possibilities, and the replication failure should be read with this ambiguity in mind.

A further limitation concerns the moderators themselves. The traits examined here, curiosity, openness, and impulsivity, were chosen for their links to exploratory behaviour, but they may not be the most relevant predictors of how people respond to a social, and specifically artificial, partner. As suggested in the discussion, traits more proximal to social influence, such as susceptibility to persuasion, conformity, social-comparison orientation, (Laursen & Faur, 2022) or general attitudes toward and trust in automation (Schaefer et al., 2016), may moderate human–bot interaction more directly. Future work could incorporate such measures, alongside explicit assessments of attitudes towards algorithms (Bock & Pütten, 2025; Jussupow et al., 2024) to test whether dispositional stance toward algorithms predicts the degree to which bot-generated information is taken up.

Beyond the sample and the choice of moderators, the artificial agent itself was deliberately simple. The bot selected tiles uniformly at random rather than employing a value-sensitive or adaptive search policy. This design choice ensured that any effect of bot presence could be attributed to the social and informational framing rather than to the sophistication of the bot's strategy, but it also limits ecological validity. The artificial agents people encounter in applied settings, from clinical decision-support systems to financial recommendation tools, are typically far more capable than a random sampler, and a competent bot might shift human search in ways a random one does not. The present findings therefore speak to the effect of an algorithmic peer's presence and framing, not to the effect of algorithmic competence. A natural next step is to manipulate the bot's search policy systematically, comparing random agents against value-sensitive or generalising agents that approximate competent human search. This would separate the effect of bot competence from the effect of bot presence and would also create the conditions under which algorithm aversion is most likely to emerge. A paradigm in which the bot visibly

succeeds or fails could thus test whether participants update their reliance on it as the human-computer interaction literature predicts (Daronnat et al., 2021; Yang et al., 2023).

A final limitation of the present design is that the mechanism behind the competitive-framing effect could not be isolated. The design contrasted a cooperative social-information framing with a competitive one while holding bot behaviour constant, and found that competitive framing was associated with higher reward in rough landscapes. Whether this reflects intensified use of social information, heightened task engagement driven by the competitive element, or some combination, could not be disentangled here. Future studies could measure engagement, perceived stakes, and subjective experience directly, or vary the strength of the competitive incentive, to identify the mechanism through which competitive framing shifts behaviour. It would also be valuable to compare an artificial competitor against a human competitor within the same task, to establish whether the competitive face of social information operates differently when the competitor is known to be a machine.

Conclusion

This thesis embedded an artificial agent into a spatially correlated multi-armed bandit task to ask how the presence and framing of an algorithmic peer relate to human performance and exploration–exploitation behaviour, and whether exploration-related traits moderate these relationships. The presence of bots did not exert a uniform effect on search. Instead, it was associated with a shift toward exploitation whose consequences depended on task structure: bot presence accompanied higher reward when the decision horizon was long and search capacity ample, and lower reward when the horizon was short. Trait-level individual differences in curiosity, openness, and impulsivity did not reliably moderate these patterns. Because the study is exploratory and uncorrected for multiple comparisons, these results are best read as hypotheses rather than established effects. Their main value is to connect two literatures that have largely developed apart, structured search and responses to algorithmic collaborators, and to provide the methodological groundwork for a pre-registered confirmatory study and for the formal computational modelling of human–machine joint search.

Data and Code Availability

The data and analysis code supporting this thesis are being prepared for a registered confirmatory replication and subsequent publication. The data, code, and experimental materials will be made openly available once the confirmatory findings have been published.

AI Use

Generative AI tools were used during the preparation of this thesis to support language editing, to help structure and refine prose, and to assist with debugging analysis code. All substantive research decisions, the study design, the statistical analyses and their interpretation, and the final wording remain the responsibility of the author, who reviewed and verified all AI-assisted output. No AI tool was used to generate data or results.

References

Bock, N., & Pütten, A. R. der. (2025). Development and Validation of the Attitudes Toward Algorithms Scale: A Universal Scale to Measure Individuals Attitudes Toward Algorithmic Decision-Making. *Human-Machine Communication*, *10*(1). https://doi.org/10.30658/hmc.10.8

Daronnat, S., Azzopardi, L., Halvey, M., & Dubiel, M. (2021). Inferring Trust From Users' Behaviours; Agents' Predictability Positively Affects Trust, Task Performance and Cognitive Load in Human-Agent Real-Time Collaboration. *Frontiers in Robotics and AI*, *8*. https://doi.org/10.3389/frobt.2021.642201

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0000033

Dormoy, C., André, J.-M., & Pagani, A. (2021). A human factors' approach for multimodal collaboration with Cognitive Computing to create a Human Intelligent Machine

Team: A Review. *IOP Conference Series: Materials Science and Engineering*, *1024*(1), 012105. https://doi.org/10.1088/1757-899X/1024/1/012105

Dubois, M., & Hauser, T. U. (2022). Value-free random exploration is linked to impulsivity. *Nature Communications*, *13*(1), 4542. https://doi.org/10.1038/s41467-022-31918-9

Haesevoets, T., De Cremer, D., Dierckx, K., & Van Hiel, A. (2021). Human-machine collaboration in managerial decision making. *Computers in Human Behavior*, *119*, 106730. https://doi.org/10.1016/j.chb.2021.106730

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X0999152X

Hesel, N., Buder, F., & Unfried, M. (2022). The Next Frontier in Intelligent Augmentation: Human-Machine Collaboration in Strategic Marketing Decision-Making. *NIM Marketing Intelligence Review*, *14*(2), 49–53. https://doi.org/10.2478/nimmir-2022-0017

Hills, T. T. (2017). Trade-Off between Exploration and Exploitation. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science* (pp. 1–2). Springer International Publishing. https://doi.org/10.1007/978-3-319-16999-6_2660-1

Hills, T. T., Todd, P. M., & Goldstone, R. L. (2008). Search in External and Internal Spaces: Evidence for Generalized Cognitive Search Processes. *Psychological Science*, *19*(8), 802–808. https://doi.org/10.1111/j.1467-9280.2008.02160.x

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., & Couzin, I. D. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences*, *19*(1), 46–54. https://doi.org/10.1016/j.tics.2014.10.004

Ho, J. C.-S., Yu-Sheng, H., & Kwan, L. Y.-Y. (2022). The impact of peer competition and collaboration on gamified learning performance in educational settings: A Meta-analytical study. *Education and Information Technologies*, *27*(3), 3833–3866. https://doi.org/10.1007/s10639-021-10770-2

Jussupow, E., Benbasat, I., & Heinzl, A. (2024). An Integrative Perspective on Algorithm Aversion and Appreciation in Decision-Making. *MIS Quarterly*, *48*(4), 1575–1590. https://doi.org/10.25300/MISQ/2024/18512

Kaduk, T., Goeke, C., Finger, H., & König, P. (2024). Webcam eye tracking close to laboratory standards: Comparing a new webcam-based system and the EyeLink 1000. *Behavior Research Methods*, *56*(5), 5002–5022. https://doi.org/10.3758/s13428-023-02237-8

Kashdan, T. B., Rose, P., & Fincham, F. D. (2004). Curiosity and Exploration: Facilitating Positive Subjective Experiences and Personal Growth Opportunities. *Journal of Personality Assessment*, *82*(3), 291–305. https://doi.org/10.1207/s15327752jpa8203_05

Laursen, B., & Faur, S. (2022). What Does it Mean to be Susceptible to Influence? A Brief Primer on Peer Conformity and Developmental Changes that Affect it. *International Journal of Behavioral Development*, *46*(3), 222–237. https://doi.org/10.1177/01650254221084103

Li, J.-M., Wu, T.-J., Wu, Y. J., & Goh, M. (2023). Systematic literature review of human–machine collaboration in organizations using bibliometric analysis. *Management Decision*, *61*(10), 2920–2944. https://doi.org/10.1108/MD-09-2022-1183

Liu, G. K.-M. (2023). *Perspectives on the Social Impacts of Reinforcement Learning with Human Feedback* (arXiv:2303.02891). arXiv. https://doi.org/10.48550/arXiv.2303.02891

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Murayama, K. (2022). A reward-learning framework of knowledge acquisition: An integrated account of curiosity, interest, and intrinsic–extrinsic rewards. *Psychological Review*, *129*(1), 175–198. (2022-22672-001). https://doi.org/10.1037/rev0000349

Naito, A., Katahira, K., & Kameda, T. (2022). Insights about the common generative rule underlying an information foraging task can be facilitated via collective search. *Scientific Reports*, *12*(1), 8047. https://doi.org/10.1038/s41598-022-12126-3

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the barratt impulsiveness scale. *Journal of Clinical Psychology*. https://doi.org/10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1

Popat, R., & Ive, J. (2023). Embracing the uncertainty in human–machine collaboration to support clinical decision-making for mental health conditions. *Frontiers in Digital Health*, *5*, 1188338. https://doi.org/10.3389/fdgth.2023.1188338

Robbins, H. (1952). Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*. https://doi.org/10.1090/S0002-9904-1952-09620-8

Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2016). A Meta-Analysis of Factors Influencing the Development of Trust in Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. https://doi.org/10.1177/0018720816634228

Schulz, E., Franklin, N. T., & Gershman, S. J. (2020). Finding structure in multi-armed bandits. *Cognitive Psychology*, *119*, 101261. https://doi.org/10.1016/j.cogpsych.2019.101261

Semeraro, F., Griffiths, A., & Cangelosi, A. (2023). Human–robot collaboration and machine learning: A systematic review of recent research. *Robotics and Computer-Integrated Manufacturing*, *79*, 102432. https://doi.org/10.1016/j.rcim.2022.102432

Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, *68*, 69–81. https://doi.org/10.1016/j.jrp.2017.02.004

Steyvers, M., Lee, M. D., & Wagenmakers, E.-J. (2009). A Bayesian analysis of human decision-making on bandit problems. *Journal of Mathematical Psychology, Special Issue: Dynamic Decision Making*, *53*(3), 168–179. https://doi.org/10.1016/j.jmp.2008.11.002

Tiokhin, L., & Derex, M. (2019). Competition for novelty reduces information sampling in a research game—A registered report. *Royal Society Open Science*, *6*(5), 180934. https://doi.org/10.1098/rsos.180934

Van Rooy, D. (2024). Human–machine collaboration for enhanced decision-making in governance. *Data & Policy*, *6*, e60. https://doi.org/10.1017/dap.2024.72

Von Helversen, B., Mata, R., Samanez-Larkin, G. R., & Wilke, A. (2018). Foraging, exploration, or search? On the (lack of) convergent validity between three behavioral paradigms. *Evolutionary Behavioral Sciences*, *12*(3), 152–162. https://doi.org/10.1037/ebs0000121

Weisberg, M., & Muldoon, R. (2009). Epistemic Landscapes and the Division of Cognitive Labor. *Philosophy of Science*, *76*(2), 225–252. https://doi.org/10.1086/644786

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Et., A. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/a0038199

Witt, A., Toyokawa, W., Gaissmaier, W., Lala, K. N., & Wu, C. M. (2023). *Social learning with a grain of salt*. PsyArXiv. https://doi.org/10.31234/osf.io/c3fuq

Witte, K., Thalmann, M., & Schulz, E. (2024). *How should we measure exploration?* PsyArXiv. https://doi.org/10.31234/osf.io/tzuey

Wu, C. M., Deffner, D., Kahl, B., Meder, B., Ho, M. H., & Kurvers, R. H. J. M. (2023). *Adaptive mechanisms of social and asocial learning in immersive collective foraging*. Animal Behavior and Cognition. https://doi.org/10.1101/2023.06.28.546887

Wu, C. M., Meder, B., & Schulz, E. (2024). *Unifying Principles of Generalization: Past, Present, and Future*.

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature Human Behaviour*, *2*(12), 915–924. https://doi.org/10.1038/s41562-018-0467-4

Xiong, W., Wang, C., & Ma, L. (2023). Partner or subordinate? Sequential risky decision-making behaviors under human-machine collaboration contexts. *Computers in Human Behavior*, *139*, 107556. https://doi.org/10.1016/j.chb.2022.107556

Yang, X. J., Schemanske, C., & Searle, C. (2023). Toward Quantifying Trust Dynamics: How People Adjust Their Trust After Moment-to-Moment Interaction With Automation. *Human Factors*, *65*(5), 862–878. https://doi.org/10.1177/00187208211034716

Appendix

Base Model: Experimental Factors

Average Reward in Relation to Experimental Factors

Table 14

Full coefficient estimates for the base model of average reward

Predictor	*b*	*SE*	*t*	*p*
Fixed effects				
Intercept	53.42**	1.16	45.99	< .001
Collab: Competitive	2.09	1.53	1.37	.172
Collab: Social Information	1.43	1.54	0.93	.352
Landscape: Wicked	−14.16**	1.51	−9.37	< .001
Landscape: Rough	−3.85*	1.49	−2.58	.010
Landscape: Kind	1.88	1.47	1.28	.201
Horizon: Short	−3.10*	1.49	−2.07	.038
Competitive × Wicked	0.23	2.11	0.11	.913
Social Information × Wicked	1.00	2.11	0.47	.636
Competitive × Rough	2.09	2.09	1.00	.317
Social Information × Rough	2.06	2.09	0.99	.324
Competitive × Kind	1.47	2.06	0.72	.475
Social Information × Kind	1.16	2.07	0.56	.576
Competitive × Short	−1.68	2.10	−0.80	.425
Social Information × Short	−1.81	2.09	−0.86	.388
Wicked × Short	4.72*	2.13	2.22	.027
Rough × Short	2.51	2.10	1.20	.231
Kind × Short	3.33	2.08	1.60	.110
Competitive × Wicked × Short	−4.16	2.98	−1.40	.163
Social Information × Wicked × Short	−3.77	2.98	−1.27	.206
Competitive × Rough × Short	−2.12	2.96	−0.72	.474

Social Information × Rough × Short	−1.99	2.94	−0.68	.497
Competitive × Kind × Short	−4.48	2.95	−1.52	.129
Social Information × Kind × Short	−3.85	2.93	−1.31	.190
Random effects (participant_id)				
Random effect	*SD*	*Variance*		
Intercept	3.370	11.360		
Slope: Competitive	2.863	8.195		
Slope: Social Information	3.024	9.144		
Residual	6.974	48.639		

Note. Linear mixed-effects model fit with REML (lme4 / lmerTest); degrees of freedom and p values via Satterthwaite approximation. * p < .05. ** p < .001.

Exploration-Exploitation Ratio in Relation to Experimental Factors

Table 15

Full coefficient estimates for the base model of exploration-exploitation

Predictor	*b*	*SE*	*t*	*p*
Fixed effects				
Intercept	0.128**	0.010	12.685	< .001
Collab: Competitive	0.016	0.012	1.308	.191
Collab: Social Information	−0.013	0.012	−1.108	.268
Landscape: Wicked	0.017	0.012	1.391	.165
Landscape: Rough	0.022	0.012	1.856	.064
Landscape: Kind	−0.008	0.012	−0.693	.488
Horizon: Short	−0.033**	0.012	−2.813	.005
Competitive × Wicked	−0.045**	0.017	−2.683	.007
Social Information × Wicked	−0.006	0.017	−0.369	.712
Competitive × Rough	−0.031	0.017	−1.858	.063
Social Information × Rough	0.003	0.017	0.152	.880

Competitive × Kind	−0.054**	0.016	−3.321	< .001
Non-competitive × Kind	−0.022	0.016	−1.340	.181
Competitive × Short	−0.004	0.017	−0.232	.817
Social Information × Short	0.009	0.017	0.550	.582
Wicked × Short	0.007	0.017	0.439	.661
Rough × Short	−0.019	0.017	−1.121	.263
Kind × Short	−0.008	0.017	−0.493	.622
Competitive × Wicked × Short	0.017	0.024	0.740	.460
Social Information × Wicked × Short	0.015	0.024	0.627	.531
Competitive × Rough × Short	0.050*	0.023	2.153	.032
Social Information × Rough × Short	0.047*	0.023	2.020	.044
Competitive × Kind × Short	0.051*	0.023	2.202	.028
Social Information × Kind × Short	0.031	0.023	1.317	.188

Random effects (participant_id)

Random effect	*SD*	*Variance*
Intercept	0.03899	0.0015200
Slope: Competitive	0.02028	0.0004111
Slope: Social Information	0.02066	0.0004268
Residual	0.05531	0.0030595

Note. Linear mixed-effects model fit with REML (lme4 / lmerTest); degrees of freedom and p values via Satterthwaite approximation. Coefficients are on the raw exploration–exploitation ratio scale (higher = more exploration). * p < .05. ** p < .001.

Extended Model: Adding Personality Moderators

Average Reward in Relation to Experimental Factors and Personality Moderators

Table 16

Full coefficient estimates for the personality model of the average reward

Predictor	*b*	*SE*	*t*	*p*
Fixed effects				
Intercept	58.51**	7.30	8.02	< .001
Collab: Competitive	6.14	8.23	0.75	.460
Collab: Social Information	2.23	8.26	0.27	.788
Landscape: Wicked	−14.11**	1.51	−9.33	< .001
Landscape: Rough	−3.80*	1.49	−2.55	.011
Landscape: Kind	1.91	1.47	1.30	.193
Horizon: Short	−3.06*	1.49	−2.05	.040
Curiosity (CEI)	−0.44	0.25	−1.80	.079
Openness (BFI)	0.19	0.15	1.30	.202
Impulsiveness (BIS)	−0.00	0.08	−0.00	.997
Competitive × Wicked	0.22	2.11	0.10	.917
Social Information × Wicked	0.95	2.11	0.45	.653
Competitive × Rough	2.05	2.09	0.98	.327
Social Information × Rough	1.97	2.09	0.94	.345
Competitive × Kind	1.48	2.06	0.72	.474
Social Information × Kind	1.08	2.07	0.52	.603
Competitive × Short	−1.71	2.10	−0.81	.417
Social Information × Short	−1.88	2.09	−0.90	.369
Wicked × Short	4.67*	2.13	2.19	.029
Rough × Short	2.48	2.10	1.18	.238
Kind × Short	3.29	2.08	1.58	.115
Competitive × Curiosity	−0.09	0.28	−0.33	.744
Social Information × Curiosity	0.22	0.28	0.78	.439

Competitive × Openness	0.05	0.17	0.31	.755
Social Information × Openness	−0.24	0.17	−1.43	.159
Competitive × Impulsiveness	−0.06	0.09	−0.67	.508
Social Information × Impulsiveness	0.01	0.09	0.09	.928
Competitive × Wicked × Short	−4.15	2.98	−1.39	.164
Social Information × Wicked × Short	−3.70	2.98	−1.24	.214
Competitive × Rough × Short	−2.07	2.96	−0.70	.484
Social Information × Rough × Short	−1.91	2.94	−0.65	.517
Competitive × Kind × Short	−4.49	2.95	−1.52	.128
Social Information × Kind × Short	−3.74	2.93	−1.27	.203

Random effects (participant_id)

Random effect	*SD*	*Variance*
Intercept	3.344	11.185
Slope: Competitive	3.053	9.322
Slope: Social Information	3.100	9.612
Residual	6.974	48.636

Note. Linear mixed-effects model fit with REML (lme4 / lmerTest); degrees of freedom and p values via Satterthwaite approximation. * $p < .05$. ** $p < .001$.

Exploration-Exploitation Ratio in Relation to Experimental Factors and Personality Moderators

Table 17

Full coefficient estimates for the personality model of the exploration-exploitation ratio

Predictor	*b*	*SE*	*t*	*p*
Fixed effects				
Intercept	0.088	0.078	1.131	.264
Collab: Competitive	0.083	0.061	1.365	.179
Collab: Social Information	0.078	0.060	1.301	.200
Landscape: Wicked	0.017	0.012	1.397	.163
Landscape: Rough	0.022	0.012	1.839	.066
Landscape: Kind	−0.008	0.012	−0.706	.480
Horizon: Short	−0.033**	0.012	−2.827	.005
Curiosity (CEI)	0.002	0.003	0.708	.483
Openness (BFI)	−0.001	0.002	−0.330	.743
Impulsiveness (BIS)	0.000	0.001	0.243	.809
Competitive × Wicked	−0.045**	0.017	−2.700	.007
Social Information × Wicked	−0.006	0.017	−0.381	.703
Competitive × Rough	−0.031	0.017	−1.855	.064
Social Information × Rough	0.003	0.017	0.167	.867
Competitive × Kind	−0.054**	0.016	−3.321	< .001
Social Information × Kind	−0.022	0.016	−1.344	.179
Competitive × Short	−0.004	0.017	−0.231	.817
Social Information × Short	0.009	0.017	0.542	.588
Wicked × Short	0.008	0.017	0.450	.652
Rough × Short	−0.018	0.017	−1.101	.271
Kind × Short	−0.008	0.017	−0.481	.631
Competitive × Curiosity	0.001	0.002	0.471	.640
Social Information × Curiosity	0.001	0.002	0.322	.749

Competitive × Openness	−0.002	0.001	−1.522	.136
Social Information × Openness	−0.002	0.001	−1.730	.091
Competitive × Impulsiveness	−0.001	0.001	−1.141	.260
Social Information × Impulsiveness	−0.001	0.001	−1.477	.147
Competitive × Wicked × Short	0.018	0.024	0.746	.456
Social Information × Wicked × Short	0.015	0.024	0.629	.530
Competitive × Rough × Short	0.050*	0.023	2.147	.032
Social Information × Rough × Short	0.047*	0.023	2.014	.044
Competitive × Kind × Short	0.052*	0.023	2.213	.027
Social Information × Kind × Short	0.031	0.023	1.331	.183

Random effects (participant_id)

Random effect	*SD*	*Variance*
Intercept	0.04032	0.0016260
Slope: Competitive	0.02003	0.0004013
Slope: Social Information	0.01929	0.0003719
Residual	0.05532	0.0030599

Note. Linear mixed-effects model fit with REML (lme4 / lmerTest); degrees of freedom and p values via Satterthwaite approximation. Coefficients are on the raw exploration–exploitation ratio scale (higher = more exploration). * $p < .05$. ** $p < .001$.

List of Figures and Tables

List of Figures

List of Tables

German Abstract

Mensch-Maschine-Kollaboration prägt zunehmend alltägliche Entscheidungsprozesse, doch ihr Einfluss auf eine grundlegende Komponente des Entscheidens unter Unsicherheit, den Kompromiss zwischen Exploration und Exploitation, bleibt empirisch unzureichend erforscht. Die vorliegende Arbeit bettet einen künstlichen Agenten in eine räumlich korrelierte „Multi-Armed-Bandit“ Aufgabe ein, um zu untersuchen, wie algorithmische Agenten menschliche Suchstrategien beeinflussen. Zwei Fragen werden behandelt: wie sich die Anwesenheit künstlicher Agenten auf Leistung und Explorations-Exploitations-Verhalten in einer strukturierten Suchaufgabe auswirkt und wie interindividuelle Unterschiede in explorationsbezogenen Merkmalen (Impulsivität, Offenheit und Neugier) diese Zusammenhänge moderieren. Zweiundvierzig Versuchspersonen absolvierten ein gamifiziertes Online-Experiment im Within-Subjects-Design, das Entscheidungshorizontlänge, Landschaftsstruktur und Suchbedingung (allein, Bot mit sozialer Information, kompetitiver Bot) kreuzte. Bots übten keinen einheitlichen Effekt auf die menschliche Suche aus, obwohl ihre Anwesenheit die Versuchspersonen tendenziell in Richtung Exploitation verschob. Diese Verschiebung war mit höherer Leistung bei Aufgaben mit langem Horizont verbunden, bei denen die Suchkapazität reichlich vorhanden war, jedoch mit geringerer Leistung bei Aufgaben mit kurzem Horizont, bei denen die reduzierte Exploration die Landschaft möglicherweise unzureichend abgetastet hat. Merkmalsbezogene individuelle Unterschiede moderierten diese Mensch-Bot-Interaktionen nicht zuverlässig. Diese explorativen Befunde deuten darauf hin, dass der Effekt künstlicher Agenten auf menschliche Entscheidungsfindung in der eingesetzten Suchaufgabe kontextabhängig ist, und liefern eine methodische Grundlage für konfirmatorische Folgestudien und die formale computationale Modellierung gemeinsamer Mensch-Maschine-Suche.