Novel Exploration Growth Quantifies Anxiety-like Behaviour in the Elevated Plus Maze

Matthew D. Zelko[1,2,3*], Stephen R. Robinson[1,4], Elisa Hill-Yardin[1,5], Helen Nasser[2]

1. School of Health and Biomedical Sciences, RMIT University, Bundoora, VIC, 3083, Australia
2. School of Psychology and Public Health, La Trobe University, Bundoora, VIC, 3083, Australia
3. Epigenes Australia Pty Ltd, Mooroolbark, VIC, 3138, Australia
4. Institute for Breathing and Sleep (IBAS), Austin Health; Heidelberg, Victoria, 3084, Australia
5. Department of Anatomy & Physiology, The University of Melbourne, VIC, 3010, Australia

* Correspondence: Matthew Zelko; Email: mattdzelko@gmail.com, Tel: +61403204267.

Abstract

The elevated plus maze is intended to provoke anxiety-like behaviour, which is primarily measured by the degree to which the rodent explores or avoids the risker, unenclosed arms. Measures such as arm entries and total time spent within the arm are conventionally used, but their analysis often produces inconsistent inferences about the level of anxiety-like behaviour being observed. This inconsistency occurs because the measures do not correlate with one another, raising the question of how validly they capture both the exploratory and avoidance motivations that typify anxiety-like behaviour. Given this inconsistency, we propose a new measure, novel exploration growth, that captures avoidance and exploratory behaviours within the maze. The growth of novel exploration tracks the first visit to discrete areas of the maze over time. The absolute amount of novel exploration, combined with the phasic nature of the exploration, reveals behavioural phenotypes with both avoidance and exploration quantified in a single time series. By addressing both motivations evoked by the maze through a single coherent measure, we provide a superior estimation of

anxiety-like behaviour and resolve inconsistencies that arise when applying conventional measures alone.

Keywords: Anxiety-like Behaviour; Elevated Plus Maze; Exploration, Avoidance

Introduction

The elevated plus maze (EPM) is a popular paradigm for testing rodents' anxiety-like behaviour (ALB). The EPM consists of four elevated arms, two of which are surrounded by high walls, known as closed arms. The remaining two arms, known as the open arms, are not enclosed. These two distinct arm types allow researchers to observe whether the rodent seeks safety in the closed arms or risk detection by exploring the open arms. (Carola et al., 2002; Leo & Pamplona, 2014). Avoidance of the open arms in favour of the safety of the closed arms is considered to indicate a greater degree of ALB. In contrast, exploration of the open arms indicates a lower degree of ALB. Exposure to the maze, therefore, evokes a conflict between the innate motivations of safety and exploration, characterising the ALB as an indicator of emotional state (Hoffman et al., 2022; Perusini & Fanselow, 2015).

Several measures have been used to compare differences in preferencing of an arm type. The standard measures include the total number of entries and time spent in an arm type. As these measures lack temporal sensitivity, researchers also track the initial latency to enter an arm type (Walf & Frye, 2007). However, using latency to enter an arm type assumes that every entrance into an arm is comparable, ignoring

differences in the depth of exploration that follows it. This limitation can result in an early, shallow entry into the open arms, which is interpreted as lower anxiety-like behaviour than a delayed but more deep exploratory entry (Tejada et al., 2009).

Alongside the lack of temporal or spatial sensitivity of these conventional measures, it is unclear how to prioritise outcomes from these measures when making inferences about the level of ALB being observed in rodents. For example, a rodent may only enter the open arms a few times but may spend more time investigating the environment in depth. Ethologically, this rodent displays high amounts of exploratory drive. Conversely, another rodent may make many shallow entries into the open arms but never fully commit to exploring them beyond the entry boundary, a pattern that should be inferred as a higher degree of ALB. Despite this, the greater number of entries into the open arms would typically be interpreted as a lower degree of ALB compared to the former exploratory counterpart. Thus, the most appropriate selection of available behavioural measures is contingent upon potentially subjective decisions determining which behaviour best typifies ALB. This need for subjectivity could lead to difficulty in replicating outcomes or translating intervention effects to human populations.

The present paper overcomes the abovementioned limitations by providing a single, coherent measure of rodent exploratory and avoidant behaviours in the EPM. Using a new measure of ALB, novel exploration growth (NEG), safety and exploration motivational preferences are observed and differentiated into qualitatively distinct behavioural phenotypes. By measuring NEG over time, each behavioural phenotype

is mapped to an appropriate level of ALB to allow justified and replicable inferences to be made about the traversal behaviours of rats in the EPM.

A multistage analysis workflow enables inferences to be made about the level of ALB using NEG. This workflow includes collecting and preprocessing tracking data, which are then analysed using Bayesian change-point regression and Bayesian generalised additive model (GAM) regression. Change-point analysis is used to identify the phases of exploratory behaviour during testing. These phases are then analysed using Bayesian GAM regression to estimate the effects of interest (e.g., between phenotypes or arm types).

The NEGs of the three behavioural phenotypes across the entire EPM and for each arm type were simulated and analysed using the aforementioned regression analyses to demonstrate how to apply the proposed workflow. In addition, we compared the three phenotypes across the conventional measures to compare the difference in ALB level inferred from these measures and NEG. This comparison highlights the incongruence between these conventional measures and demonstrates how NEG, as a unified and coherent measure, remedies these inconsistencies.

Novel Exploration Growth as a Unified Measure of Exploration and Avoidance

NEG is measured by tracking the first physical visit to discrete areas of the EPM. The arena is divided into a 1 cm × 1 cm grid in a similar manner to that used in previous studies (Costa & Tinós, 2016; Salum et al., 2000). The initial visit to each square is recorded, and unique visits are aggregated over time and expressed as a percentage of the total grid units of the arm type of interest or the entire maze. For

example, if a rodent has physically visited all of the available grid squares in the closed arms and half of the available squares in the open arms at a given point in time, they would have a total NEG of 75%, a closed arm NEG of 100% and an open arm NEG of 50%.

The choice to use the first visit to a grid square is based on the utility of the existing latency to enter the arm measure to the entire arm or maze. This utility exists because there is a need to distinguish between naïve exploration and informed preference when evaluating the motivation to enter an arm type. In this example, the first visit is considered naïve, whereas subsequent visits are informed. This distinction links the ALBs observed in the EPM test to foraging behaviours in wild-type rodents. Foraging behaviour encompasses any action taken to locate and exploit a food source. (Davidson et al., 2019; Owen-Smith et al., 2010). Foraging behaviour, particularly in new environments, evokes the conflict between safety and exploration that typifies rodent ALB in the EPM (Lima & Dill, 1990; McArthur et al., 2014). Theoretically, foraging decisions are modelled using Optimal Foraging Theory, which explains movement as the result of an attempt to maximise resource accumulation while minimising resource expenditure (Pyke, 2019; Stephens, 2008)

A key component of Optimal Foraging Theory is the Marginal Value Theorem, which predicts movement between resource patches (Charnov, 1976; King & Marshall, 2022; Pyke & Starr, 2020). Briefly, this theory predicts how long an animal should remain in a patch with resources when faced with deciding to stay or move to novel patches that may provide greater resources. For example, when that resource is food,

113 optimal foraging occurs when minimal net energy is expended while maximal net
114 energy is gained, which can only be achieved when the animal correctly predicts the
115 resource value of the current patch and any possible patches to which it may travel.
116 This decision obviously includes a degree of uncertainty that is proportional to the
117 novelty of the possible destinations. Critically, modern theories of optimal foraging
118 balance the risk of predation against the value of energy gain when assessing
119 movement patterns (Found, 2022; Hintz & Lonzarich, 2018; Křivan, 1996).

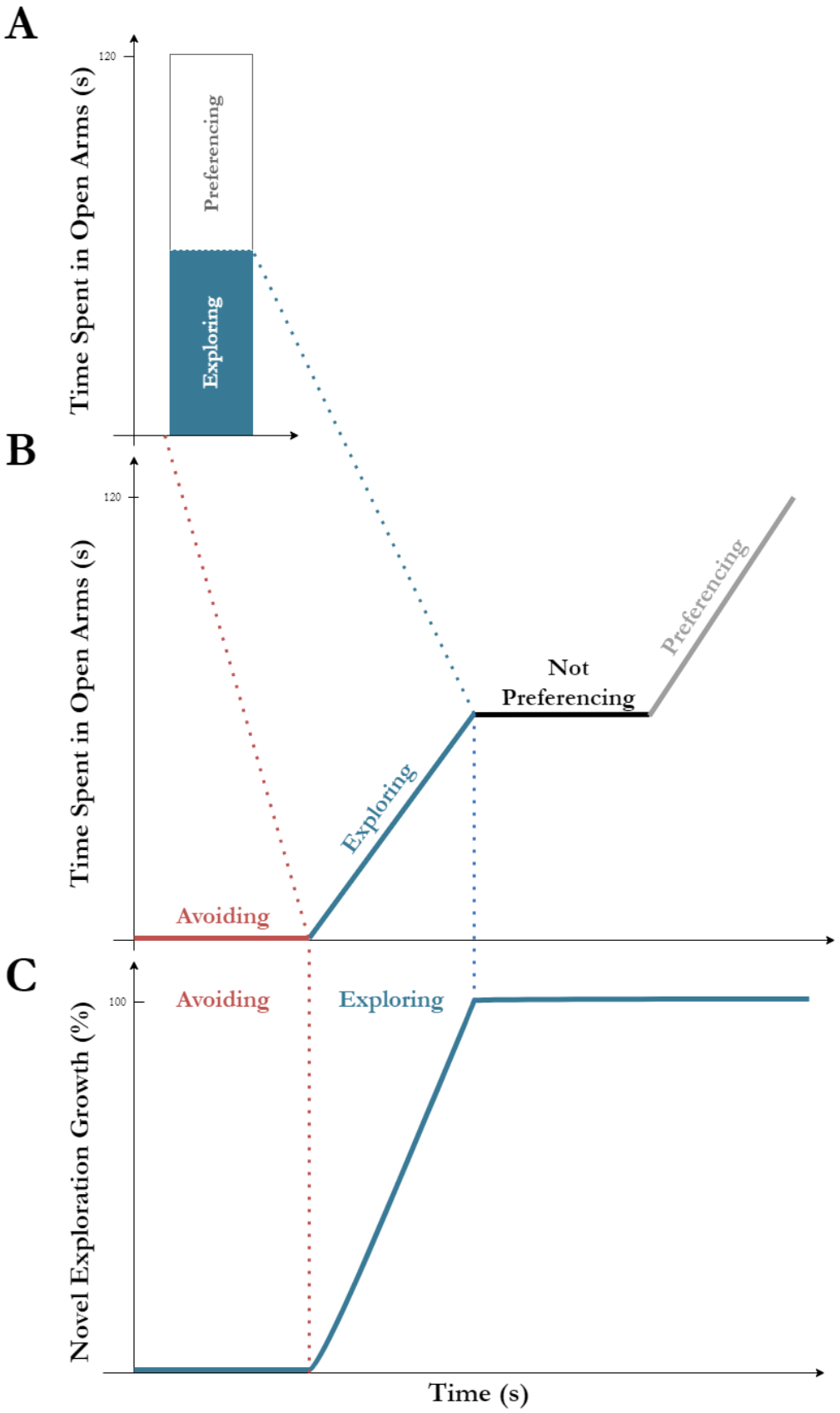
120 Rodents signal a greater predicted predation risk by shifting from cautious
121 exploration to behaviours such as surveillance, avoidance, freezing and escape.
122 Theoretically, ALBs occur in the absence of a predator but are typified by the
123 motivation to minimise the likelihood of an encounter (Hoffman et al., 2022; Mobbs et
124 al., 2018). Given this definition, ALBs are considered a cost to net energy gain,
125 resulting in suboptimal foraging. Thus, if the net energy of a foraging activity is a
126 function of energy expenditure and energy gained, both are modified due to slower
127 accumulation and increased nonforaging behaviors such as surveillance and caution.

128 The utility of NEG as a measure of ALB relies on the assumption that a rodent
129 within the EPM is acting as an optimal forager in a novel environment. When a rodent
130 anticipates a high threat risk, it will avoid exploring the area within open arms, as this
131 region poses the highest risk of encountering a predator. When the expected risk of
132 predation is low, rodents explore the available environment to evaluate possible
133 foraging opportunities. Critically, the expected energy gain from the open arms is
134 nonzero when the arms are novel to a naïve rodent, but this value diminishes

commensurate with the actual lack of opportunities to forage once rodents are informed of this through exploration. Consequently, the relative expected net energy gain of open arm areas is lower than that of closed arms following exploration. This deficit exists because, while both arm types lack foraging opportunities, the closed arms provide greater proportional safety than the open arms, so fewer ALBs have to be performed, reducing energy expenditure. This transition from naïve to informed and its impact on the value of the arm types means that an emphasis should be placed on the first visit to a novel location.

The transition from naïve to informed creates a dichotomy within the time spent in the open arms. Conventionally, time spent in the open arms considers all time spent within the open arms to have the same value across the testing period and to be generated by a single motivation. In contrast, we propose that the value of time spent in the open arms is nonlinear across time and is generated by two motivations: novel exploration and informed preferencing (see Fig 1A). This distinction between exploration and preferencing extends to when the rodent is not spending time in the open arms, creating four unique states throughout testing: avoidance, exploration, not referencing, and preferring (Fig 1B). These first two states capture the motivational conflict that governs the assessment of anxiety-like behaviour during exposure to the EPM. Consequently, they form the basis for what is measured using NEG (Fig 1C).

NEG growth curves will differ between individual rodents as they vary in their level of anxiety-like behaviour. Nonetheless, these curves tend to match one of three


A
120
Time Spent in Open Arms (s)
Preferencing
Exploring
B
120
Time Spent in Open Arms (s)
Preferencing
Not
Preferencing
Exploring
Avoiding
C
100
Novel Exploration Growth (%)
Avoiding
Exploring
Time (s)

157

158 **Figure 1** Mapping of time spent in the open arms of the EPM to novel exploration
159 growth. The time spent in the open arms comprises two qualitatively different states,
160 exploration and preferencing (A), where exploration represents the first visit to an
161 area and preferencing represents subsequent visits. These two states form part of a
162 larger suite of behaviours when measuring time spent in the open arms throughout
163 testing. This suite includes avoidance of the open arms, which is distinct from not
164 preferencing them once they have been explored (B). Given their relevance to anxiety,
165 avoidance and exploration of the open arms are mapped onto a growth curve of novel
166 exploration via NEG (C). Avoidance is inferred from a lack of initial growth in the
167 curve, whereas exploration is evident from the phase of growth.

168

169 distinct behavioural phenotypes: Exploratory, Delayed and Avoidant As shown in
170 Figure 2, a rodent exhibiting an Exploratory behavioural phenotype traverses the open
171 arms during the first phase of exploration, either exclusively or mixed with
172 exploration of the closed arms. This pattern contrasts with an Avoidant phenotype,
173 where the rodent exclusively explores the closed arms, indicating a preference for
174 safety. Critically, a rodent exhibiting the Delayed phenotype treats the arms
175 differently over time, displaying an initial avoidance of the open arms that transitions
176 into an exploration of them. Although subsequent exploration of the open arms could
177 occur in multiple phases, the key indicator of this phenotype is whether this
178 exploration reliably occurs after a period of preferencing the closed arms. This
179 preferencing causes a plateau in NEG, showing that exploration has ceased and that
180 safer areas are preferred, indicating a higher overall ALB.

181 The growth curves in Figure 2 highlight that to infer the level of ALB, there is
182 a need to estimate both the phases of exploration and the proportion of the total area
183 of the open arms that were explored. As outlined above, assessing the phases of
184 exploration establishes whether a preference existed between arm types. If the closed
185 arms were preferred earlier than the open arms, this indicated a high level of ALB. In
186 addition, the amount (i.e., area covered) of open arm exploration is important because
187 a low amount indicates a lack of exploration depth. This lack of commitment to
188 physically exploring the open arms of the EPM indicates a high level of ALB,
189 particularly if this amount of exploration is much lower than the exploration of closed
190 arms.

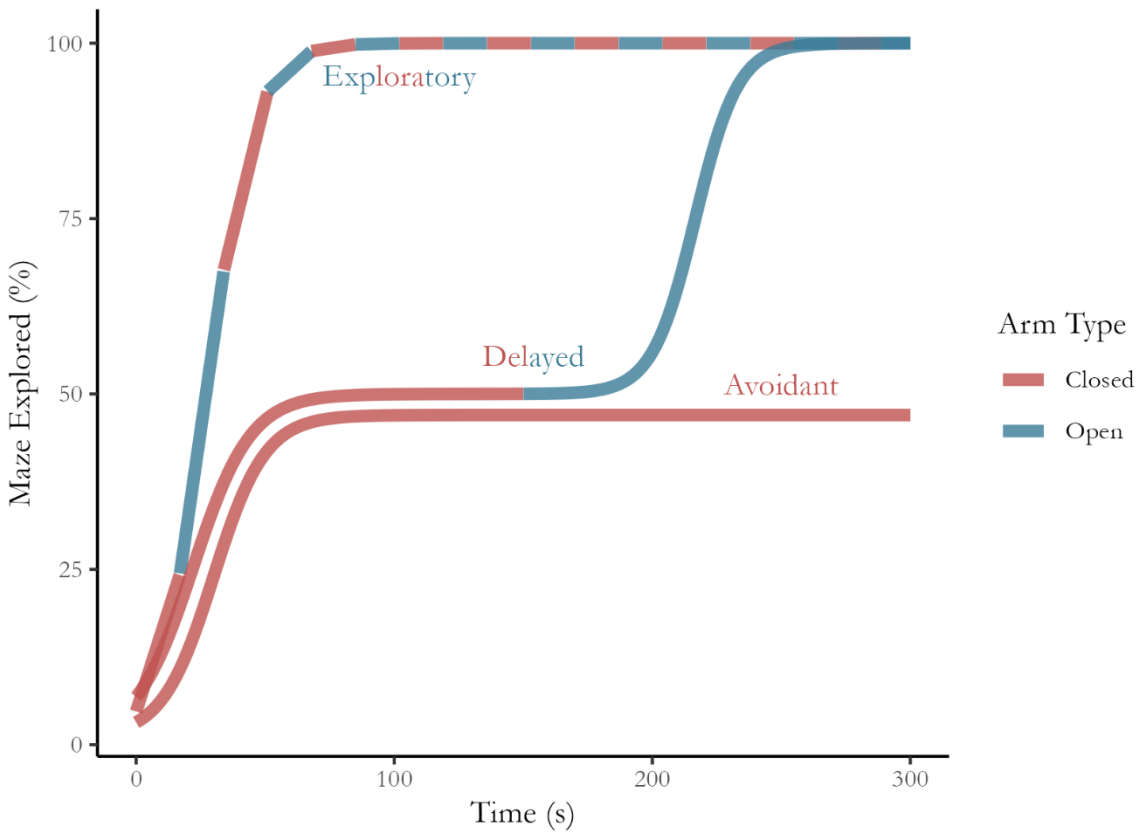
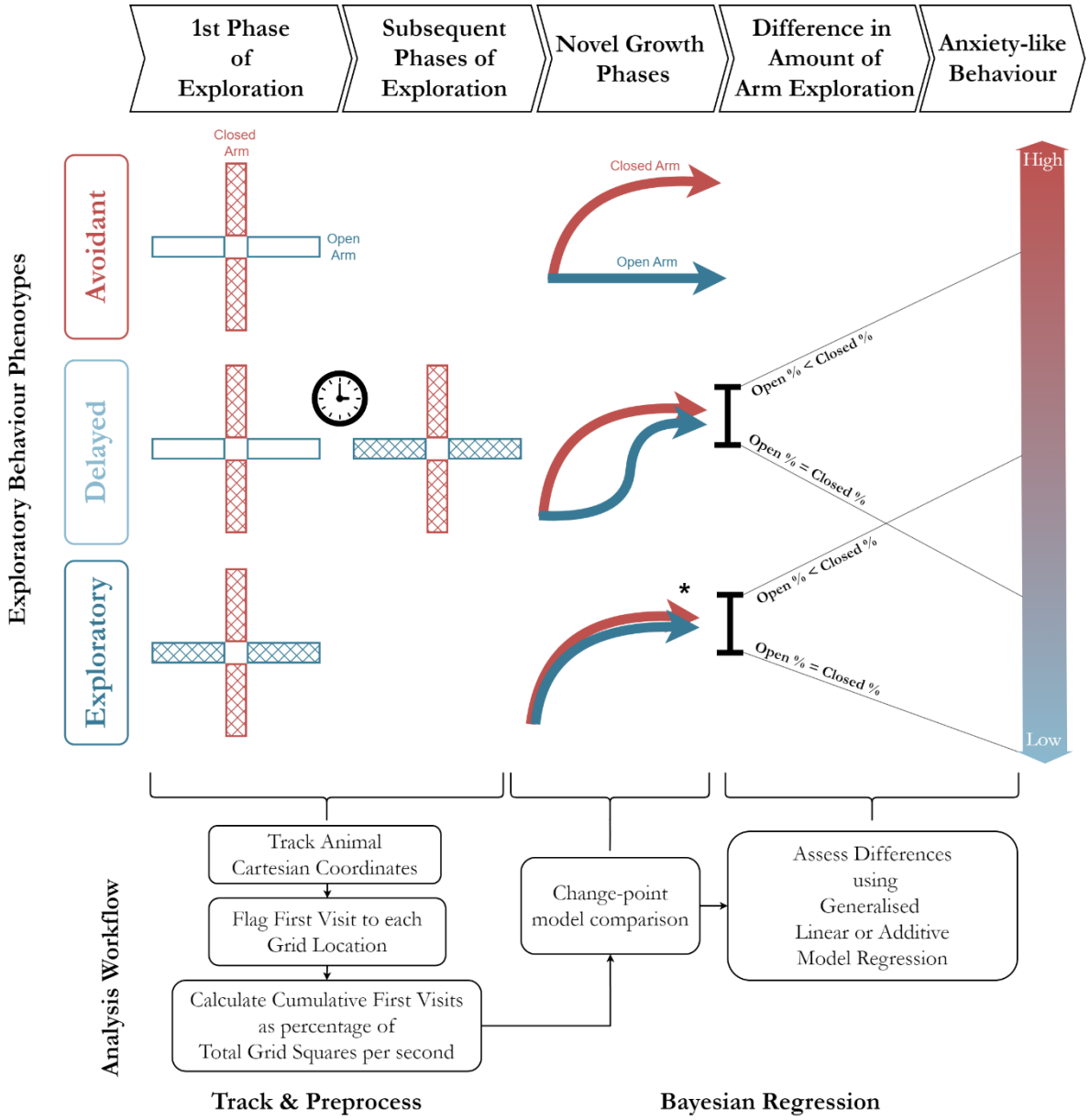


Figure 2 Exploratory, Delayed and Avoidant exploration phenotypes. The three behavioural phenotypes can be separated by their differences over time and by the extent of open arm exploration. The line colour indicates which arm type is being explored during testing. Note: For illustrative purposes, the lines have been separated during the first 20 s across all phenotypes and up to 200 s for the Delayed and Avoidant phenotypes.

200 Using the heuristics outlined above, Figure 3 shows how the exploratory
201 behavioural phenotype is mapped to a gradient of ALBs by assessing the phases of
202 exploration and the difference between the amount of exploration observed in each
203 arm type. An Avoidant phenotype is inferred when no open arm exploration is
204 observed and only the closed arms are explored, indicating the highest level of ALB.
205 A Delayed phenotype displays a similar initial phase of closed arm exploration, but
206 subsequent phases include open arm exploration. This sequence can be interpreted as
207 indicating moderately-low to moderately-high ALB, depending on the extent to which
208 the open arms were explored compared to the closed arms. Conversely, an
209 Exploratory phenotype must explore a mix of arm types or the open arms exclusively
210 within the first exploration phase. This caveat exists because the critical element
211 distinguishing this phenotype is the lack of initial avoidance of the open arms. The
212 absence of avoidance indicates that exploration, not safety, motivates this phenotype's
213 traversal in this phase. Critically, the level of ALB inferred from this sequence still
214 depends on the extent to which the open arms were explored compared to the closed
215 arms. Low ALB is inferred where the amounts are similarly high; however, if the open
216 arms are explored at a much lower comparative rate, then higher ALB is observed.

217 Both conventional measures and NEG emphasise physical commitment to the
218 open arms; however, NEG extends this emphasis by examining the depth and timing
219 of commitment. When a rodent delays or abstains from exploring the open arms, this
220 is considered to indicate that the safety of the closed arms provides

221


1st Phase of Exploration
Subsequent Phases of Exploration
Novel Growth Phases
Difference in Amount of Arm Exploration
Anxiety-like Behaviour
Exploratory Behaviour Phenotypes
Avoidant
Delayed
Exploratory
Closed Arm
Open Arm
Closed Arm
Open Arm
*
Open % < Closed %
Open % = Closed %
Open % < Closed %
Open % = Closed %
High
Low
Analysis Workflow
Track Animal Cartesian Coordinates
Flag First Visit to each Grid Location
Calculate Cumulative First Visits as percentage of Total Grid Squares per second
Change-point model comparison
Assess Differences using Generalised Linear or Additive Model Regression
Track & Preprocess
Bayesian Regression

222

223 **Figure 3** Mapping the exploratory behavioural phenotype to the ALB via novel
224 exploration growth, including the analysis workflow. The position of a rodent in the
225 maze is tracked and preprocessed to enable a temporal analysis of exploration (left to
226 right from the 1st phase to subsequent phases). The cumulative first visits for the total
227 maze and by arm type are evaluated via change-point models to establish their "Novel
228 Growth Phases". The total amount of exploration was then compared between arm
229 types to establish how much ALB was observed.
230

more perceived benefit than the novelty of the unexplored open arms. Critically, a rodent demonstrates an exploratory phenotype whether its first phase of exploration mixes arm types or whether it explores only the open arms. This is because an exploratory rodent would treat the arms equally as novel areas and explore them agnostically. Conversely, a rodent displaying a greater ALB would show a preference for the closed arms, which would manifest as multiple exploratory phases where exploration of the open arms was delayed compared to that of the closed arms.

Casarrubea and colleagues (2013) claim that sequencing behaviours in time during testing in the EPM is critical to its construct validity. However, recent models ignore how behaviours are interrelated over time. (Arantes et al., 2013a; Tejada et al., 2010). This oversight makes them incompatible with the dramatic changes that occur after information has been gathered and marginal value drops. This issue is further magnified by the lack of exploitative opportunities in the EPM, such as the presence of food, rendering the value of the open arms largely limited to their initial novelty. Conversely, the inferential power of NEG in estimating the level of ALB is made possible through a simple decision framework using linear and nonlinear time series models.

Method

The number of exploratory phases and NEG were estimated and compared using a multistage workflow (see "Analysis Workflow" in Figure 3). The workflow

comprises two discrete sections: data collection and preprocessing to calculate NEG and Bayesian regression analysis to assess the change points in the time series and conditional effects of interest, such as the effect of phenotype, arm and phenotype-by-arm type interaction effects.

Data Collection and Preprocessing

First, the position of the rodent's centre of mass is tracked per second and translated into cartesian coordinates, with the centre of the maze being assigned as the origin (0,0). Various standalone programs, such as EthoVision and ANY-maze, can be used to track movement, as can packages such as pathviewr and TrackPy in R and Python. The remaining steps in the workflow were conducted in R using custom-developed scripts.

The coordinates of the rodent are then mapped to a 1x1 cm reference grid using Cartesian coordinates with the origin at the center of the maze. Each square in the reference grid is allocated a unique identifier, which enables each visit to a square by the rodent to be recorded in time (see supplementary figure S-1). This process requires single-point tracking of the animal. Although more accurate than multipoint tracking, this method can neglect the exploration attributable to the head, which is often used near edges such as those in the open arms (Becker et al., 2023; Chetverikov & Verestoy, 1998). Critically, to accurately track the animal, the sampling rate of the tracking video must be high enough to map movement into each square. A rate of 25 frames per second is recommended.

For 300 seconds, the cumulative percentage of the total maze or specific arm type explored is computed by aggregating the flagged grid squares per second. When expressed as a time series, this cumulative measure depicts the temporal phases in which this traversal occurs and the depth to which areas are explored during these phases via the percentage value.

Comparing Phase Models using Change-Point Regression

Bayesian change-point regression analysis is conducted on NEG over time for each phenotype's total maze exploration and by arm type to evaluate exploratory phase dynamics. Change-point models are fitted using the mcp package in R, which produces inferences regarding the location of changes in the trend of a time series (Lindeløv, 2020). Like other bounded exponential growth phenomena, this time series is modelled using an n-sized sigmoid function (Kucharavy & De Guio, 2015). As shown in Figure 5, single and dual growth phase sigmoid functions are used to model change points in the NEG time series (see Supplementary S-2 for model specifications). In practice, the NEG time series for the entire maze and by arm type are modelled individually. The use of individual time series allows for an overall inference to be made regarding exploration and arm type-specific conclusions.

Estimating Marginal and Conditional Effects using Bayesian Regression

The current methodology follows the recommendations made by Betancourt (2020) when conducting Bayesian analysis for both change-point and generalised linear regression components of this workflow. Prior predictive checks are performed to assess the appropriateness of the chosen prior distributions. These checks involve

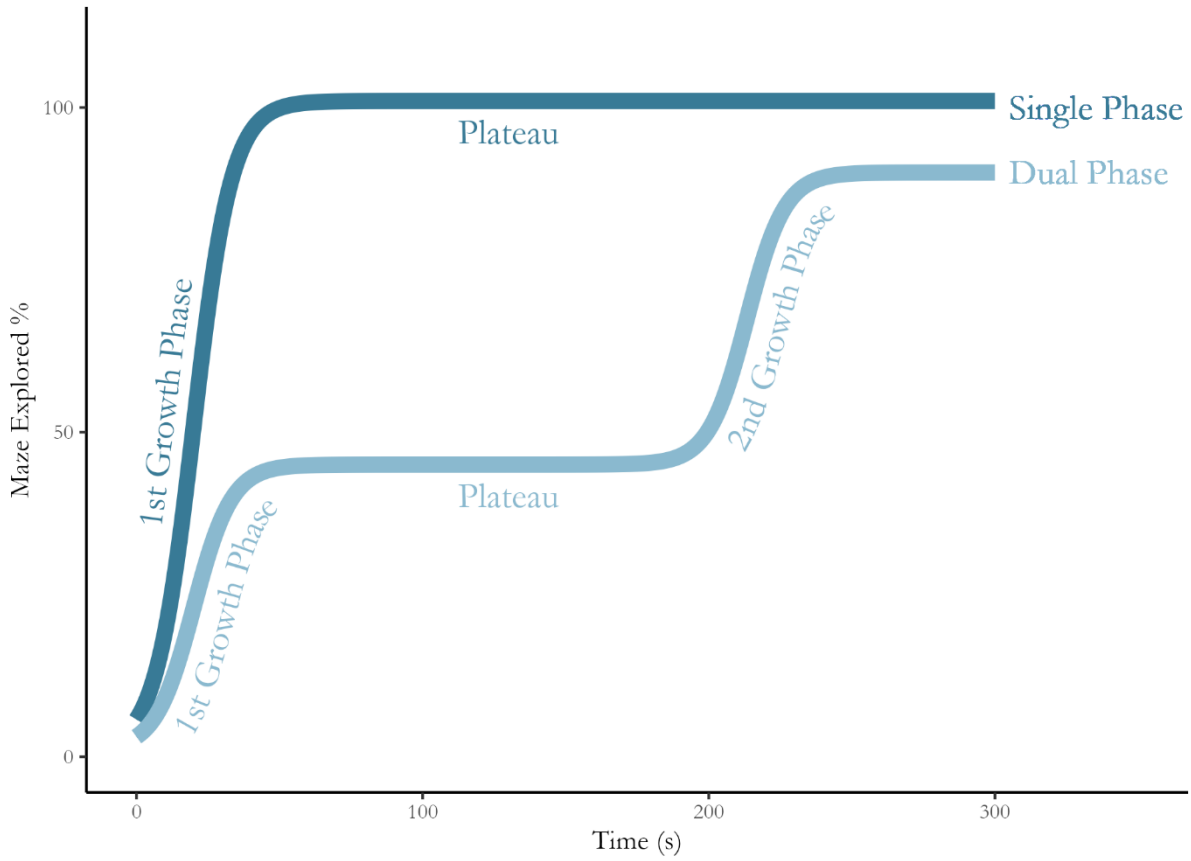


Figure 5 Single- and dual-phase exploration models for the maze. Solid coloured lines indicate novel exploration growth in the EPM over time for single- and dual-phase models. Both models show initial growth that plateaus with the dual-phase model characterised by a second growth phase.

generating data from the prior predictive distribution and evaluating whether these data are consistent with known or plausible values in the context of the study. Posterior predictive checks are performed to evaluate the model's fit and its ability to reproduce observed data patterns. These checks involve visually comparing the observed data with data simulated from the posterior predictive distribution to assess model adequacy.

Model fit is also verified by ensuring that the parameter $\hat{R}$ (Rhat) values are within the tolerance range of [0.9,1.1] and that the effective sample size (ESS) is greater than 1000 (Bürkner, 2017). Finally, leave-one-out cross-validation using Pareto-smoothed importance sampling (LOO_{PS}-CV) is conducted using the loo R package to ensure adequate model fit (Vehtari et al., 2022). LOO_{PS}-CV evaluates model fit by estimating the expected log posterior predictive distribution ($elpd_{loo}$) and the effective number of parameters used to fit an out-of-sample dataset (Vehtari et al., 2017).

Change-point model comparisons are also made via LOO_{PS}-CV. If both single- and dual-phase models meet the predictive checks and convergence criteria, models are ranked according to the highest $elpd_{loo}$ with the lowest number of parameters. The model that returns the higher $elpd_{loo}$ is preferred; when $elpd_{loo}$ is equivalent, the most parsimonious model is selected.

Once phase model selection has been completed, the effects of interest are analysed using Bayesian mixed model regression. For the simulation below, the effects of interest include the marginal phenotype effect for the total maze and the conditional effect of arm type. Generalised linear mixed models can be used to estimate the effects

of interest when growth curves of interest are best captured as single-phase growth curves. If any time series is best captured by a multiphase growth curve, such as the aforementioned dual phase model specified above, then Bayesian generalised additive model regression is applied to estimate effect characteristics. These models are also fit via the brms package using smooth functions from the mgcv r package (Wood, 2011).

Generalised additive models are chosen because, as a semiparametric specification, they can model nonlinear relationships between the response and predictor variables very flexibly while still producing interpretable effect estimates via smoothing functions (Militão et al., 2023; White et al., 2020). The smoothing functions can be fitted componentwise, for example, per phenotype or phenotype/arm type interaction, and are specified using a smoothing spline. For this workflow, a thin plate smoothing spline is chosen because it has excellent multidimensional properties, making it useful for assessing spatiotemporal foraging patterns (Christensen-Dalsgaard et al., 2017; Preisler et al., 2004).

Bayesian mixed model regression analysis is conducted using the brms r package (Bürkner, 2018), with marginal and conditional effects being estimated using the BayestestR and emmeans packages or using custom functions (Lenth, 2022; Makowski, Ben-Shachar, & Lüdecke, 2019). Marginal and conditional effects are evaluated inferentially using the SeXit framework, which reports estimates for centrality, uncertainty, existence, and significance (Kruschke, 2018; Makowski, Ben-Shachar, Chen, et al., 2019; Schwaferts & Augustin, 2020).

Centrality and uncertainty are reported using the median and highest density interval (HDI) of the distribution of the effect. The effect's existence is indicated by assessing the probability of its direction (PD). In contrast, significance is established by constructing a region of practical equivalence (ROPE), which verifies when an effect is negligible because it is practically equivalent to zero. To make the language synonymous with null hypothesis significance testing, the effect is reported as confirmed and equivalent to zero when greater than 97.5% of the posterior remains within the ROPE. Conversely, the equivalence of the effect to zero is rejected when less than 2.5% of the posterior remains within the ROPE. In cases where more than 2.5% but less than 97.5% of the posterior distribution of the effect remains within the ROPE region, the practical equivalence to zero is reported as undecided. For further details on constructing a region of practical equivalence, see Kruschke (2018) and Schwaferts and Augustin (2020).

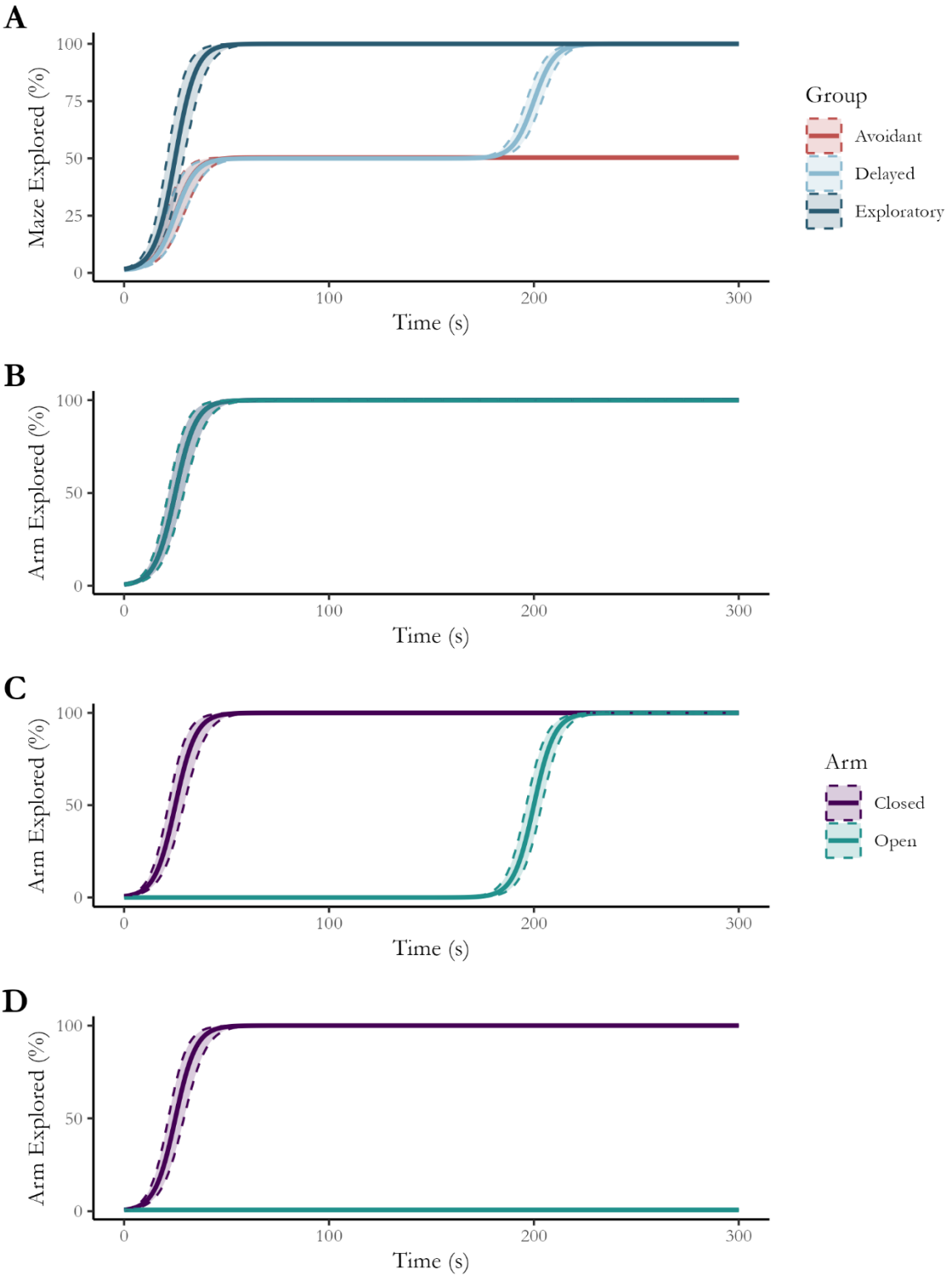
Simulating Phenotypical Exploratory and Conventional Behaviours

A simulation study was conducted to demonstrate how the change-point model comparison and regression analysis elements of the NEG workflow can be applied to samples of interest. Specifically, this study analyses the total maze and arm-specific NEG for the three behavioural phenotypes proposed above to show how to use NEG to infer the level of ALB being observed. The exploratory behaviour for the entire maze was simulated using the double sigmoid function with a sigmoid per arm

367 type (see Supplementary Figure S-3 for parameter values for growth rates and
368 inflection points by phenotype for the total maze by arm type). Parameter values were
369 chosen because they produce reasonable exemplars for the expected time series for
370 each phenotype. The Exploratory phenotype is expected to show rapid, arm-agnostic
371 exploration growth at test onset; thus, arm growth rates are sampled from the same
372 distribution, as are the inflexion time points for each arm. Conversely, the Delayed
373 phenotype is expected to show an earlier inflection time point for the closed arm than
374 for the open arm. Finally, the growth rate and the inflection points will differ between
375 arms for the Avoidant phenotype.

376 Figure 6 visually depicts the simulated NEG by phenotype for the total maze
377 (A) and by arm type for all three phenotypes (B:D). As expected, the Exploratory
378 phenotype showed greater total exploration growth due to being arm-agnostic (B),
379 whereas only closed arms were initially explored by the Delayed (C) and Avoidant
380 phenotypes (D). After 200 s of testing, the Delayed phenotype fully explored the open
381 arms, reaching an equivalent total NEG to the Exploratory phenotype. These NEG
382 time series also included three conventional ALB measurements, latency to enter the
383 open arms, number of open arm entries and total time spent in the open arms,
384 according to the following criteria:

385 • The average velocity is assumed to be identical regardless of the arm
386 type.
387 • The first entry into the open arms is followed by full exploration of that
388 arm type.


A
Maze Explored (%)
100
75
50
25
0
0
100
200
300
Time (s)
Group
Avoidant
Delayed
Exploratory
B
Arm Explored (%)
100
50
0
0
100
200
300
Time (s)
C
Arm Explored (%)
100
50
0
0
100
200
300
Time (s)
Arm
Closed
Open
D
Arm Explored (%)
100
50
0
0
100
200
300
Time (s)

389

390

391 **Figure 6** Total maze exploration and exploration by arm type per phenotype.
392 Exploratory and Delayed phenotypes initially showed similar rates of exploration
393 growth for the entire maze (A); however, while the Exploratory phenotype (B) showed
394 an overlap between NEGs by arm type, Delayed (C) and Avoidant (D) phenotypes
395 displayed differences over time due to their subsequent exploration or complete
396 avoidance of the open arms.
397

398 These assumptions ensure that the conventional ALB behaviours faithfully
399 represented the growth curves and did not create artificial differences between
400 phenotypes and arm types. Basal locomotor activity is typically operationalised by
401 measuring average velocity, where outliers may be removed from analysis due to
402 illness or malaise (Schrader et al., 2018). Therefore, keeping this constant between
403 exploratory phenotypes removes any bias associated with locomotive patterns. In
404 addition, the requirement that an arm type is fully explored following the first entry
405 ensures that no bias is introduced when estimating the number of entries or time spent
406 within an arm type. Although we expect real data to include much more variance in
407 the NEG time series they generate, these phenotypes demonstrate an idealised version
408 of exploratory, delayed and avoidant behaviours. Subsequent publications will show
409 that real-world observations can be reliably characterised using one of these
410 phenotypical behavioural patterns and analysed using the proposed workflow.

411 Given these criteria, the minimum number of open arm entries required by the
412 Exploratory and Delayed phenotypes is two, with time spent in the open arms also
413 being equivalent. Conversely, the Exploratory phenotype enters the open arms early
414 in testing, whereas the Delayed phenotype enters midway through testing. In contrast,
415 the Avoidant phenotype does not enter the open arms at all, so the latency to enter is
416 set to the maximum testing duration, with zero entries into the open arms being
417 observed and no time spent within them.

418 The distributions by phenotype for each open arm measurement value were
419 simulated 1000 times using a truncated Gaussian distribution between 0 and 300 since,

420 in the case of the time spent in the open arms and the latency to enter them, the values
421 must be positive real numbers that do not exceed 300 (see supplementary figure S-4
422 for simulation value summary and generation distribution).
423

424 Results

425 ***Conventional Open Arm Measures***

426 Bayesian linear regression was conducted using phenotype as a predictor to
427 evaluate the effect of phenotype on conventional open arm measures. Default priors
428 were used for all parameters, and all models were well specified. Figure 7 shows each
429 conventional measure's mean and variance by phenotype, including the equivalence
430 decision. As expected, the latency to enter the open arms of rodents with an Avoidant
431 phenotype, total entries and time spent within the open arms were not practically
432 equivalent to the values measured using these parameters for Exploratory or Delayed
433 phenotype rats. Conversely, only the latency to enter the open arms was not
434 practically equivalent between the Delayed and Exploratory phenotypes (see
435 supplementary figure S-5 for effect summary and SeXit reporting table).

436 ***Estimating Novel Exploration Phases via Change Point Regression***

437 To determine whether NEG occurred in a single or dual phase, two change-
438 point regression models were fitted per phenotype to illustrate movement across the
439 entire maze (see Supplementary Figure S-2 for full model specifications, including
440 priors, and S-6:7 for prior predictive checks).

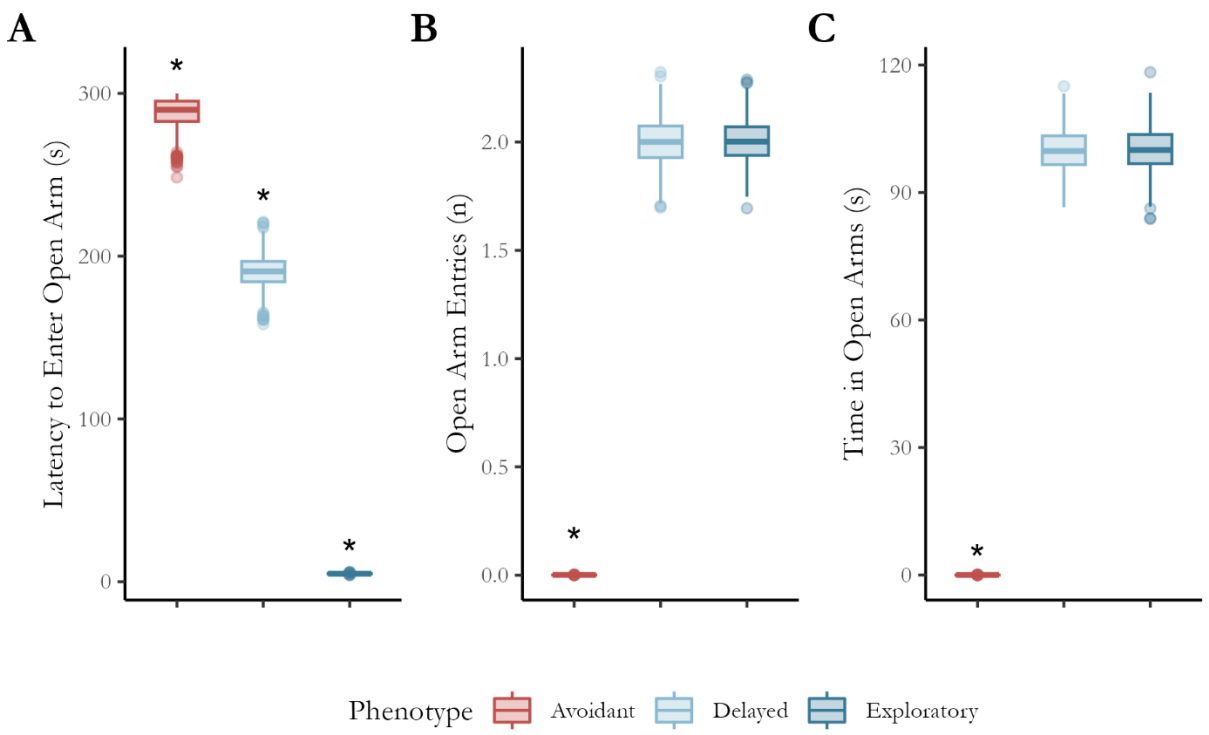
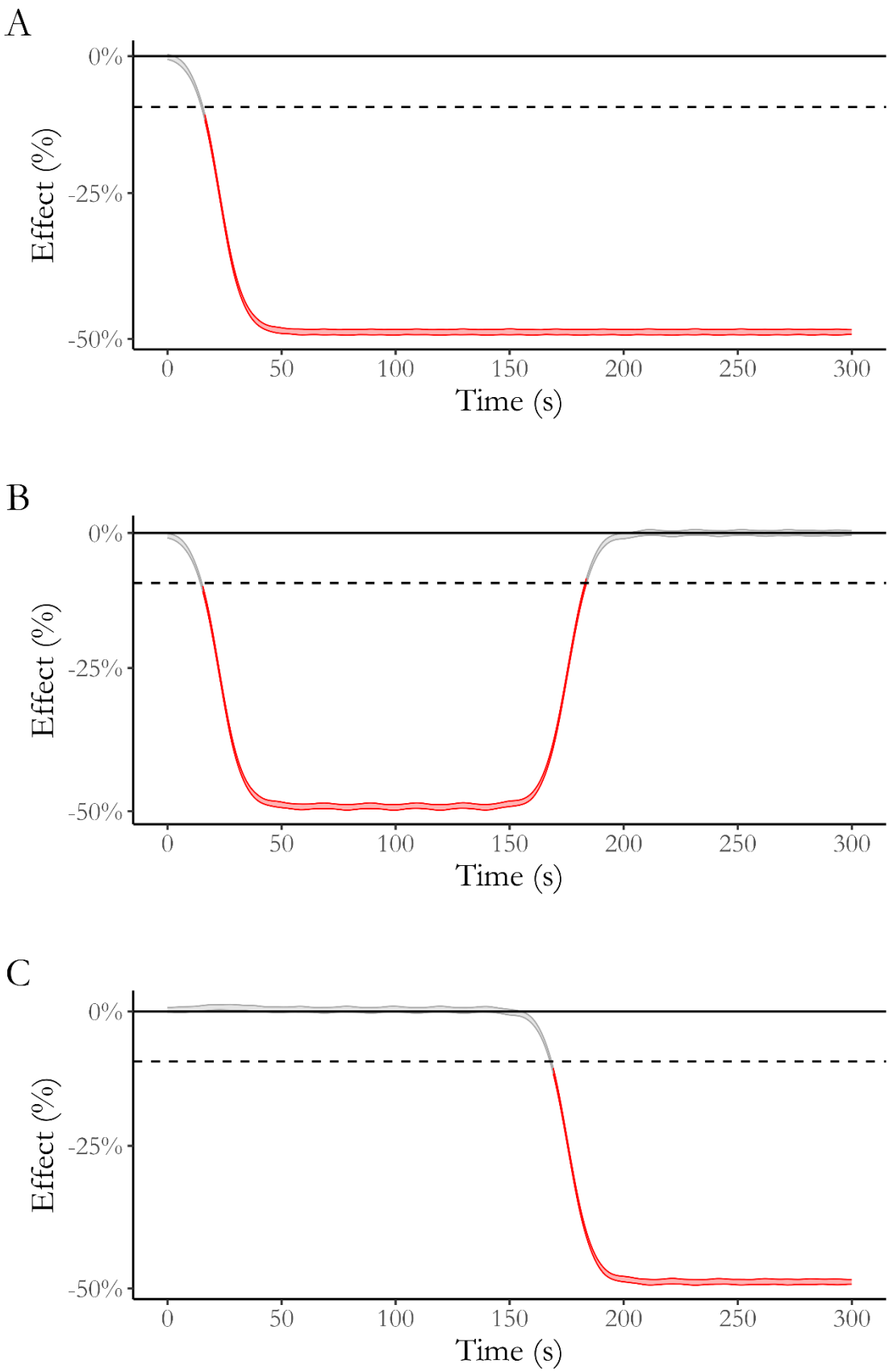


Figure 7 Mean and variance statistics for open arm preferencing metrics. The three phenotypes demonstrated differences in their latency to enter the open arms (A); however, only the Avoidant phenotype displayed differences in the number of entries (B) into and time spent within (C) the open arms. An asterisk (*) indicates that it is not practically equivalent to any other phenotype.

448 The posterior distributions and model fit diagnostic criteria for all phenotypes can be
449 found in the supplementary materials (see S-8 in the supplementary materials). Using
450 the diagnostic criteria outlined earlier, we determined that the single-phase model is
451 the preferred choice for total maze NEG for the Exploratory and Avoidant
452 phenotypes, whereas the Delayed phenotype is best modelled using a dual-phase
453 growth model.

454 Comparing NEG between Phenotypes for the Total EPM Given the
455 heterogeneity of preferred phase models across phenotypes, a generalised additive
456 mixed effects regression model was used to assess the effect of phenotype on NEG for
457 the total maze (see Figure S-9 in the supplementary materials for full model
458 specifications, including priors and S-10 for posterior predictive checks and model
459 diagnostics).

460 Figure 8 shows the posterior of the marginal effect of phenotype over time. The
461 NEG for the Exploratory phenotype is greater than that for the Avoidant (A) and
462 Delayed (B) phenotypes, although this latter difference diminishes after 200 s.
463 Conversely, the exploration growth of the Delayed and Avoidant phenotypes (C) is
464 equivalent for the first 200 s of testing but diverges due to the phase of open arm
465 exploration by the Delayed phenotype (see Figure S-11 in the supplementary materials
466 for a full regression summary table).


A
0%
-25%
-50%
Effect (%)
0
50
100
150
200
250
300
Time (s)
B
0%
-25%
-50%
Effect (%)
0
50
100
150
200
250
300
Time (s)
C
0%
-25%
-50%
Effect (%)
0
50
100
150
200
250
300
Time (s)

467

Figure 8 95% Highest density interval of effect of phenotype over time. The Avoidant and Exploratory phenotypes (A) are not equivalent to zero beyond the first few moments of the testing, whereas the difference between the Exploratory and Delayed phenotypes (B) is not equivalent for the first 200 s. The difference between the avoidant and delayed phenotypes is equivalent to zero for the first 200 s of testing (C). Red indicates that the effect is not practically equivalent to zero. Note: ROPE Region = 0 ± 18%, as indicated by the dashed line.

476 *Comparing Arm Types between and within Phenotypes*

477 To determine whether NEG by arm type for each phenotype occurred in single
478 or dual phases, the change-point regression used for the total maze exploration was
479 applied to these contrast pairs (see s-12 in the supplementary materials for posterior
480 predictive checks and model diagnostics). As expected, all phenotype/arm exploration
481 contrast pairs were best modelled as single-phase growth curves.

482 Given that all contrast pairs were best modelled using a single phase of
483 exploration growth, a generalised linear mixed model was used with a binomial link
484 function to assess conditional differences in hierarchical data such as subject-level
485 time series (see Figure S-13 in the supplementary materials for a full model expression
486 including priors and S-14 for a regression summary table including diagnostic values).

487 Figure 9 shows the median contrast effect and effect over time between each
488 phenotype/arm pair. As expected, the conditional effect of arm type rapidly became
489 negative for the Avoidant phenotype (A). Conversely, the effect of arm type is always
490 equivalent to zero for the Exploratory phenotype (B) and becomes equivalent to zero
491 for the Delayed phenotype after 200 s (C). These effects are mirrored in the marginal
492 effect of phenotype on open arm exploration, where the difference between the
493 Avoidant and Exploratory phenotypes rapidly reached a maximum (A). Conversely,
494 the difference between Delayed and Avoidant (B) became significant only after 185 s
495 of testing. In contrast, the Delayed and Exploratory phenotypes are practically
496 equivalent after 200 s of testing.

497

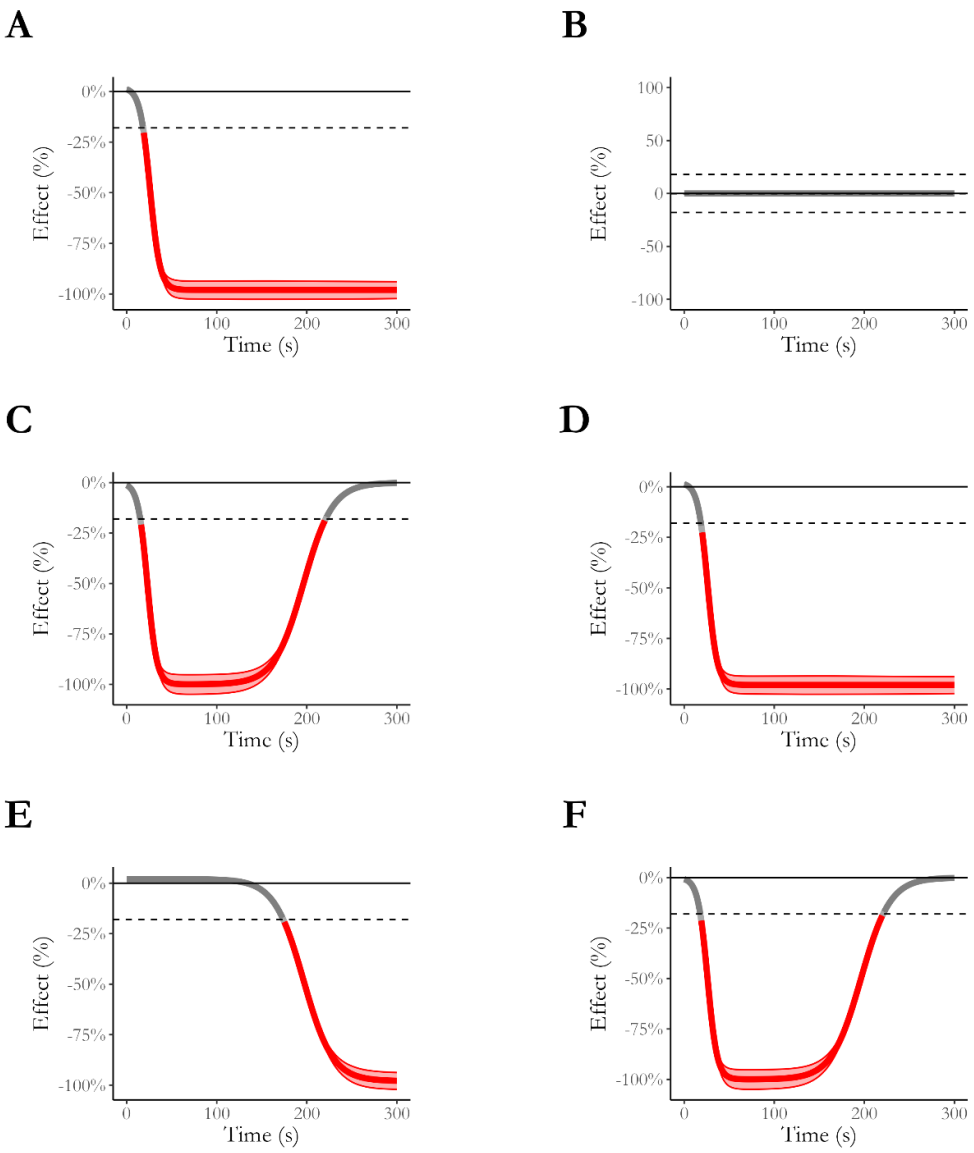


Figure 9 Effect of phenotype on open arm exploration over time. Panels A-C depict the conditional effect of Arm type on the Avoidant, Exploratory and Delayed phenotypes. Panels D-F depict the differences between the Avoidant and Exploratory, Avoidant and Delayed, and Delayed and Exploratory phenotypes on the exploration of the open arms.Red indicates that the effect is not equivalent to zero, where the thick line indicates the median effect over time and the thinner lines indicate the region of the full posterior of the effect. ROPE Region = 0 ± 18%, as indicated by the dashed line.

Discussion

This paper aimed to provide a single, coherent measure of exploratory and avoidant behaviours of rodents in the EPM. To achieve this, we propose evaluating novel exploration growth over time across the entire area of the maze and within each arm type. By simulating three distinct exploratory behavioural phenotypes, we showed that NEG uniquely and coherently maps movement throughout the maze to levels of ALB. These inferences were produced by analysing the phases of traversal of rats using change-point analysis and comparing these phases using generalised additive regression.

According to the NEG analysis, the simulation of the Exploratory phenotype captured the least anxious behavioural characteristics since the change-point analysis indicated that exploration took place in a single arm-agnostic fashion. The simulation of the Delayed phenotype appeared marginally more anxious behaviourally compared to the Exploratory phenotype, as it displayed two, arm-specific phases of exploration that signalled initial avoidance and caution that receded over time. This reduction in exploratory activity reflects a dynamic updating of threat imminence over time via information gathering that lowers avoidance and promotes exploration of the open arms (Davidson et al., 2019; Nonacs, 2001). The simulation of the Avoidant phenotype displayed maximal ALB since it displays a continuous aversion to the open arm regions of the EPM. This continual aversion indicates that the naïve predicted

526 value of the open arms never exceeds that predicted value of the safety provided by
527 the closed arms.

528 Theoretically, a higher ALB is inferred when familiar, safer areas are preferred,
529 while novel, riskier areas are avoided (Montgomery, 1951; Walf & Frye, 2007). In the
530 EPM test, this behavioural pattern is partially observed by measuring the latency to
531 enter the open arms, as it directly captures the avoidance of the riskier areas; however,
532 it cannot substantiate the preferencing of safer ones. Despite this, the latency to enter
533 the open arms was still able to correctly indicate that the Delayed phenotype
534 displayed higher ALB than the Exploratory phenotype. Conversely, the NEG
535 purposefully identifies when closed arms are not only being explored but are also
536 preferenced prior to an initial entrance into the open arms. An example of preference
537 is depicted by the Delayed phenotype's effect on NEG between arm types, where the
538 time series showed that exploration growth plateaued before exploration of the open
539 arms.

540 In contrast to the latency to enter the open arms, the conventional measure of
541 the total number of entries into the open arms could not differentiate between the
542 levels of ALB displayed by the Delayed and Exploratory phenotypes using their
543 respective simulated values. Instead, analysis of this measure erroneously indicated
544 that the two phenotypes were practically equivalent. Notably, this failure would
545 become misleading if the Delayed phenotype made even one additional, shallow entry
546 into the open arms prior to exploring them. In this hypothetical case, all three
547 conventional measures would lead to orthogonally different interpretations: latency

548 to enter the open arms indicating that the Delayed phenotype was more anxious,
549 while the total time spent in them indicating no difference in anxiety-like behaviour
550 and the total number of entries indicating lower anxiety-like behaviour. Critically,
551 there is no justifiable way to select the optimal parameter from the three conventional
552 measures because they evaluate different facets of ALB (Cruz et al., 1994; Walf & Frye,
553 2007).

554 The current study demonstrated that measuring physical exploration across
555 time can isolate early avoidance from eventual exploration. This strength was
556 evidenced by successful identification of the dual phases of the simulated NEG across
557 the entire maze by the Delayed phenotype. Critically, when evaluated using NEG
558 trends, an early, shallow entry into the open arms is correctly distinguished from the
559 true, deep exploration displayed by the Exploratory phenotype. In doing so, NEG, as
560 a single measure, can remedy inconsistencies imposed by using the three conventional
561 measures.

562 A principal limitation of the current methodology is that the speed of
563 movement impacts the calculation of NEG when above a certain rate. When a rat is
564 moving very quickly, typically at speeds greater than 25 cm/s, resolving the areas it
565 has entered becomes difficult. At these speeds, novel exploration estimation requires
566 higher resolution grids and frame rates for accurate tracking. Critically, this lack of
567 resolution of the animal's position is justified since the faster the animal moves, the
568 less information it can gather about any individual location. Given the ethological
569 basis for exploration is as an information gathering behaviour, the speed of traversal

through an area should vary proportionally with the amount of information gathered about the locale. Therefore, while NEG estimation does carry a bias toward slower exploration, this reflects greater certainty that a location has been explored compared to higher movement speeds. This assertion has been validated during testing, where speeds above 25 cm/s captured at 24 frames per second do not impact overall trends in real-world datasets. Despite this, end users should be aware that such speeds may produce small variances in expected NEG values compared to the actual calculated values. This limitation also underscores the need to capture animals at an adequate frame rate to minimise the impact.

The current framework does carry limitations, however, particularly in the implementation of the regression analysis. For example, while generalised additive models are flexible enough to handle noisy real-world data, they require stronger assumptions about how these temporal patterns map to ALB. In particular, the choice of smooth function and number of basis functions impacts the fit and interpretation of the output and its generalisability (Ezhov et al., 2018; Harmening & Neuner, 2016). In contrast, the authors explored the use of fully specified sigmoid functions, which can make interpretation easier and more generalisable, but also leads to label switching when fitting (Ngah et al., 2016; Park et al., 2019). The effectiveness of this approach has been tested against real-world data, which will be presented in a series of follow-up publications, where NEG was able to isolate differences in ALB due to treatment. Additionally, the methodology will be further validated by users who

apply it to data representing previously established anxiolytic interventions, such as benzodiazepines and antidepressants (Hogg, 1996; M. Kurt et al., 2000).

An essential limiting facet of NEG is that it only requires that users explicitly model the initial avoidance and exploration of an area or arm type. Preferencing can be inferred indirectly by comparing differences in growth between arm types, as shown by the plateau between exploration phases demonstrated by the Delayed phenotype. Further extensions could introduce more complex models that directly account for both exploratory and preferencing states, such as using partially observable Markov decision processes to include memory and agency in the model. Additionally, this proposed approach emphasises physical displacement and presence as critical indicators of anxiety during EPM testing. Future studies may incorporate dynamic Bayesian networks that include nonlocomotive behaviors such as grooming or rearing to increase the resolution of the multidimensional presentation of anxiety during exposure to the EPM. Finally, the technical aspects of the workflow may pose a barrier to implementation for some users of the EPM. To combat this, the code used in this workflow is available, and frequentist alternatives to key analytical elements are easily substituted if the user prefers them.

Conclusions

The present study demonstrated that novel exploration of the EPM accurately distinguished between three phenotypical anxiety-like behaviours. This capability

was achieved using Bayesian change-point and generalised additive regression to model rodent behavioural data from EPM assays. The study also demonstrated that conventional ALB measures produce inconsistent and erroneous inferences when analysed using simulated data representing phenotypical exploratory behavioural patterns.

Declarations

Acknowledgements

We thank Dr. Antonina Govic for her assistance in reviewing the theoretical aspects of the proposed methodology. We would also like to thank RMIT University for providing ongoing non-financial support for the project. We would like to thank Epigenes Australia for funding the article processing charge for this manuscript to allow it to be published as open access.

Author Contributions

MZ: conceptualisation, methodology, data curation, formal analysis, writing – original draft, review, and editing. **SR:** conceptualisation, methodology, writing – review and editing. **EH-Y:** conceptualisation, methodology, writing – review and editing. **HN:** conceptualisation, methodology, writing – review and editing

Funding

No funding was received to assist with the preparation of this manuscript. Epigenes Australia provided funding for the article processing charge associated with publishing as an open access article.

Conflicts of interest/Competing interests

The authors declare no competing interests.

Ethics approval

Not Applicable

Availability of data and materials

Not applicable

Code availability

The code used to calculate NEG and replicate the analysis workflow is available at https://github.com/MZelko82/NEG.

References

Becker, S., Hug, R., Hübner, W., & Arens, M. (2023). Center point-based feature representation for tracking. *Artificial Intelligence for Security and Defence Applications*, *12742*, 101–109. https://doi.org/10.1117/12.2680065

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(Plummer 2013). https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modelling with the R package brms. *R Journal*, *10*(1), 395–411. https://doi.org/10.32614/rj-2018-017

Carola, V., D'Olimpio, F., Brunamonti, E., Mangia, F., & Renzi, P. (2002). Evaluation of the elevated plus-maze and open-field tests for the assessment of anxiety-related behaviour in inbred mice. *Behavioural Brain Research*, *134*(1–2), 49–57. https://doi.org/10.1016/S0166-4328(01)00452-1

Charnov, E. L. (1976). *Optimal Foraging, the Marginal Value Theorem* (pp. 129–136).

Chetverikov, D., & Verestoy, J. (1998). Tracking feature points: A new algorithm. *Proceedings. Fourteenth International Conference on Pattern Recognition (Cat. No.98EX170)*, *2*, 1436–1438 vol.2. https://doi.org/10.1109/ICPR.1998.711973

Christensen-Dalsgaard, S., Mattisson, J., Bekkby, T., Gundersen, H., May, R., Rinde, E., & Lorentsen, S.-H. (2017). Habitat selection of foraging chick-rearing European shags in contrasting marine environments. *Marine Biology*, *164*(10), 196. https://doi.org/10.1007/s00227-017-3227-5

Costa, A., & Tinós, R. (2016). Investigation of rat exploratory behavior by evolving artificial neural networks. *Journal of Neuroscience Methods*, *270*(November 2017), 102–110. https://doi.org/10.1016/j.jneumeth.2016.06.010

Cruz, A. P. M., Frei, F., & Graeff, F. G. (1994). Ethopharmacological analysis of rat behavior on the elevated plus-maze. *Pharmacology, Biochemistry and Behavior*, *49*(1), 171–176. https://doi.org/10.1016/0091-3057(94)90472-3

Davidson, J. D., El, A., Id, H., & El Hady, A. (2019). Foraging as an evidence accumulation process. *PLoS Computational Biology*, *15*(7), 1–25. https://doi.org/10.1371/journal.pcbi.1007060

Ezhov, N., Neitzel, F., & Petrovic, S. (2018). Spline approximation, Part 1: Basic methodology. *Journal of Applied Geodesy*, *12*(2), 139–155. https://doi.org/10.1515/jag-2017-0029

Found, R. (2022). Nutritional stress and population density influence risk/reward decisions by elk. *Wildlife Research*. https://doi.org/10.1071/WR22040

Harmening, C., & Neuner, H. (2016). Choosing the Optimal Number of B-spline Control Points (Part 1: Methodology and Approximation of Curves). *Journal of Applied Geodesy*, *10*(3), 139–157. https://doi.org/10.1515/jag-2016-0003

Hintz, W. D., & Lonzarich, D. G. (2018). sMaximising foraging success: The roles of group size, predation risk, competition, and ontogeny. *Ecosphere*, *9*(10), e02456. https://doi.org/10.1002/ecs2.2456

Hoffman, A. N., Trott, J. M., Makridis, A., & Fanselow, M. S. (2022). Anxiety, fear, panic: An approach to assessing the defensive behavior system across the

predatory imminence continuum. *Learning and Behavior*, *0123456789*. https://doi.org/10.3758/s13420-021-00509-x

Hogg, S. (1996). A review of the validity and variability of the elevated plus-maze as an animal model of anxiety. *Pharmacology Biochemistry and Behavior*, *54*(1), 21–30. https://doi.org/10.1016/0091-3057(95)02126-4

King, A. J., & Marshall, H. H. (2022). Optimal foraging. *Current Biology*, *32*(12), R680–R683. https://doi.org/10.1016/j.cub.2022.04.072

Křivan, V. (1996). Optimal Foraging and Predator–Prey Dynamics. *Theoretical Population Biology*, *49*(3), 265–290. https://doi.org/10.1006/tpbi.1996.0014

Kruschke, J. K. (2018). *Rejecting or accepting parameter values in Bayesian estimation*. https://doi.org/10.1177/2515245918771304

Kucharavy, D., & De Guio, R. (2015). Application of logistic growth curve. *Procedia Engineering*, *131*, 280–290. https://doi.org/10.1016/j.proeng.2015.12.390

Lenth, R. V. (2022). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. https://cran.r-project.org/package=emmeans

Leo, L., & Pamplona, F. (2014). Elevated Plus Maze Test to Assess Anxiety-like Behavior in the Mouse. *Bio-Protocol*, *4*(16), 1–8. https://doi.org/10.21769/BioProtoc.1211

Lima, S. L., & Dill, L. M. (1990). Behavioral decisions made under the risk of predation: A review and prospectus. *Canadian Journal of Zoology*, *68*(4), 619–640. https://doi.org/10.1139/z90-092

Lindeløv, J. K. (2020). mcp: An R Package for Regression With Multiple Change Points. *Journal of Statistical Software*. https://doi.org/10.31219/osf.io/fzqxv

M. Kurt, A. C. Arik, & S. Çelik. (2000). The Effects of Sertraline and Fluoxetine on Anxiety in the Elevated Plus-Maze Test. *Journal of Basic and Clinical Physiology and Pharmacology*, *11*(2), 173–180. https://doi.org/10.1515/JBCPP.2000.11.2.173

Makowski, D., Ben-Shachar, M., & Lüdecke, D. (2019). bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Makowski, D., Ben-Shachar, M. S., Chen, S. H. A. A., & Lüdecke, D. (2019). Indices of Effect Existence and Significance in the Bayesian Framework. *Frontiers in Psychology*, *10*(December), 1–14. https://doi.org/10.3389/fpsyg.2019.02767

McArthur, C., Banks, P. B., Boonstra, R., & Forbey, J. S. (2014). The dilemma of foraging herbivores: Dealing with food and fear. *Oecologia*, *176*(3), 677–689. https://doi.org/10.1007/s00442-014-3076-6

Militão, T., Kürten, N., & Bouwhuis, S. (2023). Sex-specific foraging behaviour in a long-lived seabird. *Marine Biology*, *170*(11), 132. https://doi.org/10.1007/s00227-023-04280-7

Mobbs, D., Trimmer, P. C., Blumstein, D. T., & Dayan, P. (2018). Foraging for foundations in decision neuroscience: Insights from ethology. *Nature Reviews Neuroscience*, *19*(7), 419–427. https://doi.org/10.1038/s41583-018-0010-7

Montgomery, K. C. (1951). The relation between exploratory behavior and spontaneous alternation in the white rat. *Journal of Comparative and Physiological Psychology*, *44*(6), 582–589. https://doi.org/10.1037/h0063576

Nonacs, P. (2001). State dependent behavior and the Marginal Value Theorem. *Behavioral Ecology*, *12*(1), 71–83. https://doi.org/10.1093/oxfordjournals.beheco.a000381

Owen-Smith, N., Fryxell, J. M., & Merrill, E. H. (2010). Foraging theory upscaled: The behavioural ecology of herbivore movement. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1550), 2267–2278. https://doi.org/10.1098/rstb.2010.0095

Perusini, J. N., & Fanselow, M. S. (2015). Neurobehavioral perspectives on the distinction between fear and anxiety. *Learning & Memory*, *22*(9), 417–425. https://doi.org/10.1101/lm.039180.115

Preisler, H. K., Ager, A. A., Johnson, B. K., & Kie, J. G. (2004). Modelling animal movements using stochastic differential equations. *Environmetrics*, *15*(7), 643–657. https://doi.org/10.1002/env.636

Pyke, G. (2019). Optimal foraging theory: An introduction. In J. C. Choe (Ed.), *Encyclopedia of animal behavior* (Vol. 2, pp. 111–117). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-809633-8.01156-0

Pyke, G. H., & Starr, C. K. (2020). Optimal Foraging Theory. *Encyclopedia of Social Insects*, 1–9. https://doi.org/10.1007/978-3-319-90306-4_90-1

Salum, C., Morato, S., & Roque-da-Silva, A. C. (2000). Anxiety-like behavior in rats: A computational model. *Neural Networks*, *13*(1), 21–29. https://doi.org/10.1016/S0893-6080(99)00099-4

Schrader, A. J., Taylor, R. M., Lowery-Gionta, E. G., & Moore, N. L. T. (2018). Repeated elevated plus maze trials as a measure for tracking within-subjects behavioral performance in rats (*Rattus norvegicus*). *PLoS ONE*, *13*(11), 1–15. https://doi.org/10.1371/journal.pone.0207804

Schwaferts, P., & Augustin, T. (2020). *Bayesian Decisions using Regions of Practical Equivalence (ROPE): Foundations* (Tech, Vol. 235, Issue 235). Ludwig Maximilian University of Munich, Department of Statistics: Technical Reports. https://doi.org/10.5282/UBM/EPUB.74222

Stephens, D. W. (2008). Decision ecology: Foraging and the ecology of animal decision making. *Cognitive, Affective and Behavioral Neuroscience*, *8*(4), 475–484. https://doi.org/10.3758/CABN.8.4.475

Tejada, J., Bosco, G. G., Morato, S., & Roque, A. C. (2009). sCharacterisation of rat behavior in the elevated plus-maze using a directed graph. *Journal of Neuroscience Methods*, *184*(2), 251–255. https://doi.org/10.1016/j.jneumeth.2009.08.009

Vehtari, A., Gabry, J., Magnusson, M. S., Y, Y., P, B., Paananen, T., & Gelman, A. (2022). *loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models* [Computer software].

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and Estimating out-of-sample pointwise predictive accuracy using posterior simulations. *J Stat Comput*, *27*(5), 1413–1432.

Walf, A. a, & Frye, C. a. (2007). The use of the elevated plus maze as an assay of anxiety-related behavior in rodents. *Nature Protocols*, *2*, 322–328. https://doi.org/10.1038/nprot.2007.44

White, L. F., Jiang, W., Ma, Y., Armah, K. S., Samet, J. H., & Cheng, D. M. (2020). Tutorial in Biostatistics: The Use of Generalised Additive Models to Evaluate Alcohol Consumption as an Exposure Variable. *Drug and Alcohol Dependence*, *209*, 107944. https://doi.org/10.1016/j.drugalcdep.2020.107944

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric sgeneralised linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(1), 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x