nature portfolio

https://doi.org/10.1038/s41586-026-10301-4

Supplementary information

Dopaminergic mechanisms of dynamical social specialization

In the format provided by the authors and unedited

Supplementary information

Dopaminergic mechanisms of dynamical social specialization

C. Solié, A. Nicolson, R. Justo, Y. Layadi, B. Morin, C. Batifol, LM Reynolds, T. Le Borgne, SL Fayad, A. Gulmez, Y Rodriguez Quevedo, J. Allegret--Vautrot, G. Centene Guglielmi, F. de Chaumont, S. Didienne, N. Debray, J-P. Hardelin, B. Girard, A. Mourot, J. Naudé, C. Viollet, F. Marti, B. Delord, P. Faure

Supplementary information table of contents:

Supplementary methods

Animals

Male and female wild-type (WT) C57BL/6J mice (8-12 weeks old, Janvier Labs, France) or DAT-iCre mice (C57BL/6J background, BAC-DAT-iCre)[1] were housed in groups in an animal facility and maintained on a 12-hour light-dark cycle (7:00 a.m.-7:00 p.m.), at an average temperature of 22°C (22 ± 2°C) and ~50% humidity. All the animals lived in triads in the animal facility before entering in the task.

Live Mouse-Tracker (LMT) system

This setup, in our paradigm, provided a controlled context for examining how individual mice learn and adapt to the task, either in isolation or within a standard social setting, allowing comparisons between lone and group behaviors. We employ comprehensive 24 hours/7 days data collection within the experimental environment, a method that provides rich longitudinal datasets[2]. The data collected fall into two main categories: discrete events and continuous video streams. Discrete events are automatically captured by sensors integrated into the environment, including infrared beams at the lever and nose poke, and food dispenser activations (MedAssociates). Each detected event generated a time-stamped TTL (Transistor–Transistor Logic) pulse, which was logged with the sensor identifier (and, when possible, the animal identity) in a structured database, enabling detailed analysis of individual and group-level behavioral patterns over time. In addition to discrete events, we record continuous 24/7 video streams, which provide a complete visual record of the animal's behavior within the environment. To manage the substantial storage demands of such video data, we implement real-time processing techniques to extract essential features. This includes tracking each animal's center of mass and body contours, allowing us to quantify locomotion, spatial organization, and interactions[2]. These extracted metrics are stored in a compressed format for subsequent analysis, ensuring that critical information is retained without overwhelming storage capacity. This dual approach to data collection provides a nuanced understanding of behavior, capturing both fine-grained event data and broader spatial and social dynamics.

The LMT and MedAssociates softwares were relaunched every day to prevent huge file weight. The number of lever-presses, the number of nose pokes (beam breaks) and the number of complete sequences were extracted directly from the database using PyCharm (Python) and MySQL. We defined two specific zones around the lever and the food magazine to be able to know which animal was located there during a lever press or a nose poke. We then determined which TTL (Transistor-Transistor Logic) was associated with which animal in the zone. For each day of analysis, it was not possible to associate a small percentage of trials with specific mice since it happened that several mice were present at the same time either at the lever or at the magazine. Once the association was made between the position of the animal at a specific timestamp and the time of the TTL, we obtained the number of lever-presses

and nose pokes for each animal. For complete sequences, we quantified the number of times that the mouse which had pressed the lever was the first arrived at the subsequent nose poke in less than 6 seconds, otherwise the sequence was not counted (typically, if another mouse was already at the magazine waiting for the food).

Tube test

In some mouse triads we conducted a tube test to determine the social hierarchy before and after the microsociety task. Three days before the task, the mice were habituated to traverse a 30 cm long tube with a 2.5cm diameter. The tube was open to two cages, one on each side, containing fresh bedding. The mice were considered habituated when they had passed through the tube at least 10 times.

Two days before the task we conducted the first and second trials of the tube test by placing each mouse of a pair at the entrance of the tube on opposite sides. Since we tested triads of mice, each mouse participated in two confrontations per trial. Once the mice engaged in the tube and met at the center, the confrontation started. The outcome of the confrontation was recorded as a win or a loss when the hind limbs of one of the mice (the loser) touched the cage floor. One day before the task, we performed trial 3, 4 and 5. We performed five tube-test trials before the microsociety task. One week later, at the end of the microsociety task, we performed another five trials for each triad. After each session, the tube and cages were cleaned and bedding was changed. In extended Figure 4a, we analyzed each mouse's rank within its triad before and after the microsociety experiment. On the left, each line represents one mouse, showing its rank across the five initial pre-experiment trials, which sometimes resulted in rank crossings. On the right, ranks from the last trial (Trial 5 pre) before the microsociety experiment were compared with those (Trial 5 post) obtained after the microsociety experiment, allowing us to visualize shifts by plotting each mouse's final pre-experiment rank against its post-experiment rank.

Archetypal analysis

All archetypal analyses and visualizations were performed in R using the archetypes package (version 2.2-0.1). Archetypal analysis[3,4] identifies k idealized behavioral profiles (archetypes) that capture the boundaries of variability in a multivariate dataset. Given an n×m matrix X (with n animals and m behavioral features), the algorithm finds a matrix Z of k archetypes by minimizing the residual $|| X-\alpha Z^T ||^2$, where:

- α is a coefficient matrix such that $\alpha_{i,1..k} \geq 0$ and $\sum\alpha_{i,1..k} = 1$,
- each archetype z_j is itself a convex combination of the data: $Z = X^T\delta$ with $\delta_{j,i} \geq 0$ and $\sum\delta_{i,1..k} = 1$.

Each individual is thus positioned within a convex hull bounded by the archetypes. The α-coefficients quantify the contribution of each archetype to an individual's profile. For k = 3, the results can be visualized in a ternary plot, and individuals are assigned to an archetype based on their maximal α value. The full set of α-coefficients also allows quantification of distances to each archetype, which we use to assess proximity to idealized behavioral types.

To determine the optimal number of archetypes *k*, we computed the residual sum of squares (RSS) as a function of *k*, and visually inspected the archetypal solutions obtained for different *k* values. This procedure was applied both in the lone and social conditions. The resulting elbow plots and corresponding archetypal configurations are presented in Extended Data Fig.1d-e and Extended Data Fig.2c-d to illustrate how the choice of *k* balances residual error and interpretability of the behavioral space.

In Fig. 1, archetypal analysis was conducted in two dimensions using the total number of lever presses (#LP) and the percentage of complete sequences (%CS), yielding two behavioral archetypes: *Achiever* and *Storer*. This low-dimensional archetypal decomposition was used to assess individual variability in the lone context.

In Fig. 2, a three-dimensional archetypal space was constructed using eight behavioral features: intra-triad %LP and %CS for days 5 to 7, along with the number of food pellets gained and lost due to conspecifics. This reference space, computed from the full dataset (n = 195 mice), defined three archetypes, *Workers*, *Scroungers*, and *Storers*, and was used as a common projection space for subsequent analyses. In Fig. 5 and Fig. 6, new experimental cohorts (*e.g.,* mixed-sex triads or DA-manipulated triads) were projected into this common archetypal space (defined in Fig. 2). Their α-coefficients were then used to assign behavioral strategies, compute distances to archetypes (*e.g.,* distances to *Scrounger* or *Storer*), and compare group-level compositions across conditions. This approach captures the continuous geometry of behavioral variability, while allowing interpretation in terms of discrete, idealized roles. Indeed, it enables both categorical assignment (based on nearest archetype) and continuous analyses (using the distance to archetypes).

The same archetypal framework was applied to the modeling data. In the case of simulated triads, archetypes were defined using a dataset in which β-values were clearly separated between males (high-β) and females (low-β), allowing robust identification of behavioral strategies (see Fig. 4g). All subsequent simulated conditions, including those presented in Extended Data Fig.7, were then projected into this reference archetypal space. This allowed us to track how manipulations of β values affected the distribution of simulated individuals within the behavioral space, providing a direct link between this model parameter and emergent role distributions.

To analyze the relationship between dopaminergic activity and behavioral roles, we used the α-coefficients from archetypal decomposition to compute the distance of each animal to the three

archetypes. We then modeled peak VTA responses (to food consumption, own LPs or conspecific LPs) as a function of distance to each archetype using linear regression in R (*lm* function). These fits were made separately for each archetype, using all animals (males and females) in the photometry cohort. To estimate the expected dopaminergic response at each archetype, we computed the intercept of the linear model when the distance to the archetype was zero. Confidence intervals around the estimates were obtained using the *predict()* function with standard error propagation. This procedure was applied identically to firing rates from *in vivo* electrophysiology data (Fig. 3i).

Observed cage-level profile distributions were compared against expected distributions generated by random sampling (Fig. 2i). For each sex separately, the overall proportion of animals assigned to each archetype was computed. Expected cage compositions were then generated by repeatedly sampling (10,000 iterations) three individuals independently, while preserving the empirical distribution of archetypes within each sex (Fig. 2h), yielding the null distribution of cage profiles. Simulated frequencies were normalized to the number of observed cages and compared to the empirical counts.

In vivo electrophysiology

Mice were deeply anesthetized with isoflurane delivered continuously (3% induction, 1-2% maintenance; TEMSega). The scalp was opened and a hole was drilled in the skull above the location of the VTA. Extracellular recording electrodes were constructed from 1.5 mm outer diameter / 1.17 mm inner diameter borosilicate glass tubing (Harvard Apparatus) using a vertical electrode puller (Narishige). The tip was broken straight and cleaned under microscopic control to obtain a diameter of about 1 µm. The electrodes were filled with a 0.5% NaCl solution yielding impedances of 6-9 MΩ. Electrical signals were amplified by a high-impedance amplifier (Axon Instruments) and monitored audibly through an audio monitor (A.M. Systems Inc.). The signal was digitized, sampled at 25 kHz, and recorded on a computer using Spike2 software (Cambridge Electronic Design) for later analysis. The electrophysiological activity was sampled in the central region of the VTA (coordinates: 3.1 to 4 mm posterior to bregma, 0.3 to 0.7 mm lateral to midline, and 4 to 4.8 mm below brain surface). Individual electrode tracks were separated from one another by at least 0.1 mm in the horizontal plane. Spontaneously active DA neurons were identified based on previously established electrophysiological criteria[5,6]. Activity was quantified by overlapping 60 second windows shifted every 15 seconds. Percentage of spikes within bursts (%SWB) corresponds to the percentage of spikes discharged within a burst in a given time interval. Firing frequency response was quantified as a percentage of variation from baseline and %SWB response was quantified as a difference of percentage from baseline [7] .

Stereotaxic surgeries

For fiber photometry experiments: injection of AAV1-Syn-FLEX-GCaMP7c (Titer ≥ $9.4x10^{12}$ $vg.mL^{-1}$, Addgene, Watertown, MA, USA) was performed in 6-8 week-old DAT-Cre mice. Mice were anesthetized with a mixture of oxygen (1 L/min) and isoflurane 2% (TEMSega, Pessac, France). After shaving the skin, mice were placed in a stereotaxic frame (David Kopf, Tujunga, USA), the skin was disinfected and locally anesthetized with 70 μL of lurocaine (0.5%). Craniotomies were then performed, and unilateral injections were performed over the VTA at AP: -3.20 mm, ML: ± 0.5 mm from the bregma, and DV: −4.20 ± 0.05 mm from the brain surface. Injections were carried out with a glass micropipette (Drummond Scientific Company, Broomall, PA), at a 100 nL/min rate for a total volume of 300 nL in the VTA. At the end of the surgery, 100 μL of buprenorphine (0.1 mg/kg) was injected subcutaneously. Unilateral optic fiber (200 μm core, NA = 0.39, ThorLabs) implantations were performed 2-3 weeks later above VTA with the same coordinates, and dental acrylic (SuperBond, Sun medical) was used for fixing the ferrule to the skull.

For optogenetic experiments: AAV5-Ef1α-DIO-ChR2(H134R)-EYFP (Titer ≥ 4.7x1012 vg.mL-1, VVF Zürich, Switzerland) or AAV5-Ef1α-DIO-EYFP (Titer ≥ 5.4x1012 vg.mL-1, VVF Zürich, Switzerland) were injected in 6-8 week-old DAT-Cre male mice. Mice were anesthetized and disinfected as previously described. The animals were placed in a stereotactic frame and bilateral injections were performed in the VTA at the same coordinates than above, 300 nL per side, using a glass micropipette. After 2-3 weeks, one optic fiber (200 μm core, NA= 0.39, ThorLabs) was implanted unilaterally above the VTA with a 10° angle (AP: -3.20 mm; ML: ±0.9 mm from the bregma; DV: −3.95 mm from the brain surface), and fixed to the skull using dental acrylic.

For chemogenetic experiments: AAV5-hSyn-DIO-hM4Di-mCherry (Titer ≥ 5.6x1012 vg.mL-1, VVF Zürich, Switzerland) or AAV5-hSyn-DIO-mCherry (Titer ≥ 7.5x1012 vg.mL-1, VVF Zürich, Switzerland) were injected in 6-8 week-old DAT-Cre female mice. Mice were anesthetized and disinfected as previously described. The animals were placed in a stereotactic frame and bilateral injections were performed in the VTA at the same coordinates than above, 300 nL per side, using a glass micropipette.

After all the surgeries, mice were placed alone in a heated cage for one hour recovery, and were subsequently checked every day. The behavioral experiments started at least one week after the surgery. Injections and implantations were all confirmed with *post-hoc* immunochemistry.

Immunohistochemistry

Following euthanasia, brains were rapidly extracted and fixated in a solution of 4% paraformaldehyde for a minimum of three days at 4°C. Subsequently, serial sections of 60 μm in thickness were obtained from the midbrain using a vibratome.

For immunostaining experiments (TH-staining), free-floating brain sections were initially incubated for one hour at 4°C in a phosphate-buffered saline (PBS) blocking solution containing 3% bovine serum albumin

(BSA) by volume/volume ratio (Sigma; A4503) and supplemented with 0.2% Triton X-100. This was followed by overnight incubation at 4°C with a primary antibody solution containing a mouse monoclonal anti-tyrosine hydroxylase antibody (anti-TH, Sigma, T1299) diluted 1:500 in PBS supplemented with 1.5% BSA and 0.2% Triton X-100.

The following day, sections were thoroughly rinsed with PBS, and then incubated for three hours at room temperature (22-25°C) with the secondary antibody Cy3-conjugated goat anti-mouse (Jackson ImmunoResearch, 715-165-150) diluted at 1:500 in a solution of 1.5% BSA in PBS. After undergoing three additional rinses with PBS, the slices were mounted using Prolong Gold Antifade Reagent with DAPI (Invitrogen, P36930). Microscopic analysis was performed using a Zeiss epi-fluorescent microscope, and image capture was performed by a camera. To preserve the raw data, images were captured in grayscale format using the Zeiss software (ZEN microscopy software). Following acquisition, false coloring was applied using ImageJ to enhance visualization of the immunofluorescence signal.

Analysis of fiber photometry recordings

We first subtracted a biexponential detrend to our fluorescence signal in order to eliminate the slow decay of the signal due to sensor photobleaching, and then added on offset equal to the signal mean before detrending. The ΔF/F ratio was calculated as follow:

$$\frac{\text{fluo signal} - \text{baseline fluo signal}}{\text{baseline fluo signal}}$$

The analysis of VTA DA calcium variations and its correlation with behavior relied on the construction of peri-event time histograms, which were obtained by aligning and centering specific behavioral events. The identification of these events was done with TTLs received on the fiber photometry console at lever-press or nose-poke events. The temporal resolution of the recordings was determined by the bin size chosen for analysis. We used a bin size of 100 ms to achieve a more precise representation of the calcium variations over time. Subsequent analyses were conducted on normalized activity data. To obtain normalized calcium fluorescence variations, a specific formula was applied:

$$Zscore\Delta F/F = (m\Delta F/F - \mu\Delta F/F)/\sigma\Delta F/F$$

where Zscore ΔF/F represents the normalized calcium activity for a specific time bin, mΔF/F represents the average calcium signal variations within that specific time bin, μΔF/F represents the mean calcium signal variations across the baseline activity a few seconds before each events (5 seconds), and ΔF/F represents the standard deviation of the calcium signal variations across the baseline activity a few seconds before each event.

Following normalization, a smoothing technique was employed on the data. This involved convolution with a Gaussian kernel using a sliding window encompassing 100 bins. The "gausswin" function within the MATLAB software was utilized to achieve this smoothing process.

To distinguish between focused and unfocused events in *Scrounger* mice (Fig. 3), we performed a visual scoring analysis exclusively during identified conspecific's lever presses. An event was classified as focused if the recorded mouse was: (i) oriented toward the lever press, (ii) oriented toward the food dispenser, or (iii) located within 5 cm of the food dispenser. Conversely, an event was classified as unfocused if the mouse was: (i) not oriented toward the lever press or food dispenser, or (ii) positioned farther from the food dispenser.

Supplementary Table 1

The details of the statistics (tests, degrees of freedom and exact p-values) for each figure panel.

Figure 1	Figure 1c (left)	One-way RM ANOVA	Time main effect LP : $F_{(3,61)} = 35.10$; $p < 0.0001$ Time main effect BB : $F_{(3,61)} = 86.59$; $p < 0.0001$
	Figure 1c (right inset)	One-way RM ANOVA	Time main effect: $F_{(3,61)} = 30.37$; $p < 0.0001$
	Figure 1d	Wilcoxon rank test	LP before learning : v= 18 ; p = 1
		Paired t-test	Food before learning : $t_{(5)} = -3.77$; p = 0.0131
		Paired t-test	LP After learning complete sequence : $t_{(5)} = -2.82$; p = 0.0371
		Paired t-test	Food After learning complete sequence : $t_{(5)} = -3.29$; p = 0.0216
		Wilcoxon rank test	LP After learning incomplete sequence : W = 9 ; p = 0.128
		Wilcoxon rank test	Food After learning incomplete sequence : W = 2 ; p = 0.011
	Figure 1g	Two-way RM ANOVA	Time main effect: $F_{(3,180)} = 37.90$; $p < 2\times10^{-16}$ Group main effect : $F_{(1,60)} = 87.19$; p =2.68e-13 Interaction Time x Group : $F_{(3,180)} = 16.13$; p = 2.49e-09
	Figure 1h	Wilcoxon rank sum test with continuity correction	W = 326.5 ; p = 0.04428
	Figure 1i	Pearson's Chi-squared test with Yates' continuity correction	$\chi2 = 5.3654$; df = 1 ; p = 0.02054
	Figure 1j	Two-way ANOVA	Sex Main effect : $F_{(1,58)} = 210.57$; $p < 0.0001$
			Archetype Main effect : $F_{(1,58)} = 0.035$; p = 0.852
			Sex x archetypes : $F_{(1,58)} = 0.002$; p = 0.967
	Figure 1k	Wilcoxon rank test	W = 3685; p = 0.0009
	Figure 2c (left)	Welch Two Sample t-test	t = -0.13152 ; df = 190.63 ; p = 0.8955
	Figure 2c (right)	Wilcoxon rank sum test with continuity correction	W = 955 ; p < 0.0001
	Figure 2e (up left) : %LP	One way ANOVA followed by Tukey multiple comparisons of means	LP : Rank effect $F_{(2,192)} = 113.2$, $p < 2\times10^{-16}$. Post hoc: A2 < A1 (p < 0.001), A3 < A1 (p < 0.001), A3 > A2 (p < 0.001).
	Figure 2e (upright) : %CS	One way ANOVA followed by Tukey multiple comparisons of means	CS : Rank effect $F_{(2,192)} = 187.3$, $p < 2\times10^{-16}$. Post hoc: A3 < A1 (p < 0.001), A3 < A2 (p < 0.001); A2 vs A1 (ns).
	Figure 2e (bottom left) Gain	One way ANOVA followed by Tukey multiple comparisons of means	Gain : Rank effect $F_{(2,192)} = 145$, $p < 2\times10^{-16}$. Post hoc: A2 > A1 (p < 0.001), A3 < A1 (p < 0.001), A3 < A2 (p < 0.001).

Figure 2	Figure 2e (bottom right) loss	One way ANOVA followed by Tukey multiple comparisons of means	Loss : Rank effect $F_{(2,192)} = 131$, $p < 2\times10^{-16}$. Post hoc: A2 < A1 ($p < 0.001$), A3 < A1 ($p < 0.001$), A3 < A2 ($p < 0.001$).
	Figure 2g (left) distance Storer	Wilcoxon rank sum test with continuity correction	W = 8581 ; p < 0.0001
	Figure 2g (right) distance Scrounger	Wilcoxon rank sum test with continuity correction	W = 1592.5 ; p < 0.0001
	distance to Worker (not shown)	Wilcoxon rank sum test with continuity correction	W = 2485 ; p < 0.0001
	Figure 2h	Pearson's Chi-squared test	$\chi2 = 104.45$; df = 2 ; p < 0.0001
	Figure 2j	One way ANOVA followed by Pairwise comparisons using Wilcoxon rank sum test with continuity correction and P value adjustment method: holm	main effect $F_{(2,165)} = 58.5$; p < 0.0001
			$p_{1-2} = 0.73$; $p_{1-3} < 0.0001$, $p_{2-3} < 0.0001$
Figure 3	Figure 3c	Two-way RM ANOVA Males	Time main effect: $F_{(1,115)} = 46.54$; p < 0.0001 Events main effect : $F_{(2,115)} = 6.46$; p = 0.0030 Interaction Time x Events : $F_{(55,115)} = 6.02$; p = 0.0043
		Two-way RM ANOVA Females	Time main effect: $F_{(1,59)} = 35.60$; p < 0.0001 Events main effect : $F_{(2,59)} = 15.54$; p = 0.0030 Interaction Time x Events : $F_{(27,59)} = 11.80$; p = 0.0002
	Figure 3d	Two-way RM ANOVA Workers	Time main effect: $F_{(1,47)} = 12.39$; p = 0.0020 Events main effect : $F_{(2,47)} = 2.41$; p = 0.11 Interaction Time x Events : $F_{(21,47)} = 2.34$; p = 0.12
		Two-way RM ANOVA Scrounger	Time main effect: $F_{(1,67)} = 35.90$; p < 0.0001 Events main effect : $F_{(2,67)} = 5.49$; p = 0.0091 Interaction Time x Events : $F_{(31,67)} = 4.79$; p = 0.0154
	Figure 3e	Paired t-test	Focused : $t_{(11)} = -3.43$; p = 0.0056
		Paired t-test	Unfocused : $t_{(1)} = -1.33$; p = 0.21
	Figure 3f	Linear regression (LM)	AUC_LP ~ Distance_A1: $\beta = -4.59 \pm 2.38$, $t(14)=-1.93$, p=0.074 (ns trend), $R^2adj=0.15$, n=16
		LM with interaction + post hoc contrasts (emmeans, Tukey)	Archetype × Distance: $F_{(2,42)} = 5.20$, p = 0.0096. Intercepts at Distance = 0: A1 > A2 ($t_{(42)} = 2.84$, p = 0.0188); A1 vs A3 (ns); A3 vs A2 (ns). N = 48.
	Figure 3g	Linear regression (LM)	AUC_OtherLP ~ Distance_A2: $\beta = -6.64 \pm 2.04$, $t_{(15)} = -3.26$, p = 0.0053, $R^2adj = 0.38$, n=17
		LM with interaction + post hoc contrasts (emmeans, Tukey)	Archetype × Distance: $F_{(2,45)} = 7.65$, p = 0.0014. Intercepts at Distance = 0: A2 > A1 ($t_{(45)} = 3.44$, p = 0.0035); A2 > A3 trend (p = 0.056); A3 vs A1 (ns). N = 51.

	Figure 3h (left)	Wilcoxon rank-sum exact test	Firing rate Male vs Female (Before): W = 4487, p = 0.17 (adj. p = 0.35), ns
			Firing rate Male vs Female (After): W = 21622, p = 2.5×10^{-6} (adj. p = 1.0×10^{-5}), ***
			Firing rate Female Before vs After: W = 7242, p = 0.90 (adj. p = 0.90), ns
			Firing rate Male Before vs After: W = 24202, p = 0.0086 (adj. p = 0.026), *
	Figure 3h (right)	Wilcoxon rank-sum exact test	Mean Firing rate Male vs Female (Before): W = 24, p = 0.78 (adj. p = 1.0), ns
			Mean Firing rate Male vs Female (After): W = 45, p = 7.6×10^{-5} (adj. p = 8.5×10^{-4}), ***
			Mean Firing rate Female Before vs After: W = 48, p = 0.94 (adj. p = 1.0), ns
			Mean Firing rate Male Before vs After: W = 154, p = 0.031 (adj. p = 0.11), trend
	Figure 3i (above left)	Linear regression (LM)	mean_Spontf ~ dist_to_A1: β = –1.47 ± 0.79, $t_{(38)}$ = –1.86, p = 0.071 (trend), R^2adj = 0.06, n = 40
	Figure 3i (above right)	Linear regression (LM)	mean_Spontf ~ dist_to_A2: β = –1.14 ± 0.71, $t_{(38)}$= –1.61, p = 0.12 (ns), R^2adj = 0.04, n = 40
	Figure 3i (bottom left)	Linear regression (LM)	mean_Spontf ~ dist_to_A3: β = 1.58 ± 0.55, $t_{(38)}$ = 2.87, p = 0.0067, R^2adj = 0.16, n = 40
	Figure 3i (bottom right)	LM with interaction + post hoc contrasts (emmeans, Tukey)	mean_Spontf ~ Archetype × Distance: $F_{(2,114)}$ = 6.94, p = 0.0014. Intercepts at Distance = 0: A3 < A1 (p = 0.013), A3 < A2 (p = 0.018), A1 vs A2 (ns). n = 120.
	Figure 5a	Pearson's Chi-squared test	Association between sex and archetype distribution: $\chi^2_{(2)}$ = 9.53, p = 0.0085, **
	Figure 5b	Paired t-test	Distance to archetype 2: $t_{(12)}$= –7.11, mean diff = –0.46, 95% CI [–0.60, –0.32], p = 1.2×10^{-5}, ***
Figure 5	Figure 5d	Pearson's Chi-squared test	Archetypal profile distribution across groups (YFP, 1ChR2, ChR2): $\chi^2_{(4)}$ = 15.95, p = 0.0031, **
		One-sided 2-sample proportion test (Holm)	Storer proportion: YFP vs 1ChR2, p = 0.37 (adj. p = 0.37), ns
			Storer proportion: YFP vs ChR2, p = 0.0021 (adj. p = 0.0064), **
			Storer proportion: 1ChR2 vs ChR2, p = 0.011 (adj. p = 0.023), *
	Figure 5f	Pearson's Chi-squared test	Archetypal profile distribution across groups (YFP, 1Dreadd, Dreadd): $\chi^2_{(4)}$ = 17.90, p = 0.0013, **
		One-sided 2-sample proportion test (Holm)	Scrounger proportion: YFP vs 1Dreadd, p = 0.50 (adj. p = 0.50), ns
			Scrounger proportion: YFP vs Dreadd, p = 0.011 (adj. p = 0.022), *

		[illegible]	Scrounger proportion: 1Dreadd vs Dreadd, p = 0.0061 (adj. p = 0.018), *
Extended Figure 1	Extended Figure 1b	Wilcoxon rank test	Moving : W = 6; p = 0.69
		Wilcoxon rank test	Immobile : W = 6 ; p = 0.10
	Extended Figure 1c	Mann-Whitney rank test	Up : U = 52 ; p = 0.0411
		Pearson correlation test	Bottom : R = -0.74 ; p = 0.0236
	Extended Figure 1f	Kolmogorov-Smirnov 2 sample test	K = 0.23; p = 0.36
			K = 1; p < 0.0001
	Extended Figure 1g	Wilcoxon rank test	W = 3685; p = 0.18
Extended Figure 2	Extended Figure 2a (left)	Kolmogorov-Smirnov 2 sample test	K = 0.33 ; p = 0.0140
	Extended Figure 2a (right)		K = 0.83 ; p < 0.0001
	Extended Figure 2b (left)	Kolmogorov-Smirnov 2 sample test	K =0.38 ; p < 0.0001
	Extended Figure 2b (right)		K = 0.28 ; p < 0.0001
	Extended Figure 2e	2-way Repeated-measures ANOVA	Lever presses: main effect of Rank ($F_{(2,190)} = 65.23$, $p < 2\times10^{-16}$); main effect of Day ($F_{(1,192)} = 32.7$, $p = 4.1\times10^{-8}$); Rank × Day interaction ($F_{(2,192)} = 21.7$, $p = 3.1\times10^{-9}$).
	Extended Figure 3b	Fisher exact test with Holm correction	Scrounger : Zone 1: Pre vs Post (ns), Pre < Random (p = 3.7e-12), Post < Random (p = 3.3e-14). Zone 2: Pre < Post (p = 0.0022), Pre < Random (p < 1e-15), Post < Random (p < 1e-15). Zone 3: Pre vs Post (ns), Pre < Random (p = 0.0022), Post < Random (p = 5.4e-07). Zone 4: Pre < Post (p = 8.9e-04), Pre vs Random (ns), Post < Random (p = 2.7e-05).
			Worker: Zone 1: Pre vs Post (ns), Pre < Random (p < 1e-15), Post < Random (p < 1e-15). Zone 2: Pre vs Post (ns), Pre < Random (p = 3.4e-08), Post < Random (p = 9.1e-10). Zone 3: all ns (p > 0.45 after correction). Zone 4: Pre vs Post (ns), Pre < Random (p = 0.0041), Post vs Random (ns after correction).

Extended Figure 3			Storer: Zone 1: all ns after correction (trend Pre vs Random, p = 0.071). Zone 2: all ns after correction. Zone 3: all ns after correction (trend Pre vs Random, p = 0.213). Zone 4: Pre vs Post (ns), Pre vs Random (trend, p = 0.093), Post > Random (p = 0.017).
	Extended Figure 3c (up left)	2-way Repeated-measures ANOVA	Time main effect: $F_{(23,78)} = 35.42$, $p < 0.0001$; Group main effect: $F_{(2,78)} = 19.31$, $p < 0.0001$; Interaction Time x Group: $F_{(46,78)} = 5.75$, $p < 0.0001$.
	Extended Figure 3c (up right)	One-way ANOVA followed by Bonferroni-Holm correction	Group main effect: $F_{(2,78)} = 19.30$, $p < 0.0001$
	Extended Figure 3c (bottom left)	2-way Repeated-measures ANOVA	Time main effect: $F_{(23,78)} = 5.39$, $p < 0.0001$; Group main effect: $F_{(2,78)} = 6.10$, $p = 0.0035$; Interaction Time x Group: $F_{(46,78)} = 3.27$, $p < 0.0001$.
	Extended Figure 3c (bottom right)	One-way ANOVA followed by Bonferroni-Holm correction	Group main effect: $F_{(2,78)} = 6.10$, $p = 0.0035$
	Extended Figure 3d	Linear regression (LM)	LP5 ~ LP2: $\beta = 0.44 \pm 0.05$, $t_{(193)} = 8.27$, $p = 2.1\times10^{-14}$, $R^2adj = 0.26$, n = 195
			LP7 ~ LP6: $\beta = 0.74 \pm 0.04$, $t_{(190)} = 17.25$, $p < 2\times10^{-16}$, $R^2adj = 0.61$, n = 192
	Extended Figure 3e (left)	2-way Repeated-measures ANOVA	Time main effect: $F_{(23,79)} = 19.77$, $p < 0.0001$; Group main effect: $F_{(2,79)} = 8.11$, $p = 0.0006$; Interaction Time x Group: $F_{(46,79)} = 2.78$, $p < 0.0001$.
	Extended Figure 3e (right)	2-way Repeated-measures ANOVA	Time main effect: $F_{(23,79)} = 22.34$, $p < 0.0001$; Group main effect: $F_{(2,79)} = 5.99$, $p = 0.0038$; Interaction Time x Group: $F_{(46,79)} = 3.51$, $p < 0.0001$.
Extended Figure 4	Extended Fig 4b	One-Way anova (pre and post)	main effect pre - Male - Scrounger : $F_{(2,21)} = 0.48$; $p = 0.625$
			main effect post - Male - Scrounger : $F_{(2,21)} = 0.66$; $p = 0.528$
			main effect pre - Female - Storer : $F_{(2,21)} = 0.01$; $p = 0.99$
			main effect pre - Female - Storer : $F_{(2,21)} = 0.05$; $p = 0.95$
	Extended Figure 5a (left)	paired T-test	Male - Food before learning : $t(10) = -2.37$; $p = 0.039$
	Extended Figure 5a (right)	paired T-test	Female - Food before learning : $t(3) = -8.29$; $p = 0.0037$
	Extended Figure 5b (left)	paired T-test	Male - LP before learning : $t(3) = -3.18$; $p = 0.05$
	Extended Figure 5b (right)	paired T-test	Female - LP before learning : $t(2) = -1.58$; $p = 0.25$
	Extended Figure 5d	Linear regression (LM)	AUC_LP ~ Distance_A2: $\beta = 5.40 \pm 1.97$, $t_{(14)} = 2.74$, $p = 0.016$, $R^2adj = 0.30$, n = 16
	Extended Figure 5e	Linear regression (LM)	AUC_OtherLP ~ Distance_A1: $\beta = 7.07 \pm 2.80$, $t_{(15)} = 2.52$, $p = 0.023$, $R^2adj = 0.25$, n = 17
	Extended Figure 5f	Linear regression (LM)	AUC_LP ~ Distance_A3: $\beta = -1.97 \pm 2.69$, $t_{(14)} = -0.73$, $p = 0.47$ (ns), $R^2adj = -0.03$, n = 16
	Extended Figure 5g	Linear regression (LM)	AUC_OtherLP ~ Distance_A3: $\beta = 2.42 \pm 2.75$, $t_{(15)} = 0.88$, $p = 0.39$ (ns), $R^2adj = -0.01$, n = 17

Extended Figure 5	Extended Figure 5i (Left)	Wilcoxon rank-sum exact test	Spontaneous activity (burst) Male vs Female (Before): W = 5075, p = 0.95 (adj. p = 0.95), ns
			Spontaneous activity (burst) Male vs Female (After): W = 19103, p = 3.6×10^{-10} (adj. p = 1.4×10^{-9}), ***
			Spontaneous activity (burst) Female Before vs After: W = 6140, p = 0.052 (adj. p = 0.10), trend
			Spontaneous activity (burst) Male Before vs After: W = 25886, p = 9.4×10^{-5} (adj. p = 2.8×10^{-4}), ***
	Extended Figure 5i (right)	Wilcoxon rank-sum exact test	Mean spontaneous burst activity Male vs Female (Before): W = 33, p = 0.53 (adj. p = 0.53), ns
			Mean spontaneous burst activity Male vs Female (After): W=71, p=0.0020 (adj. p=0.0081), **
			Mean spontaneous burst activity Female Before vs After: W = 34, p = 0.26 (adj. p = 0.52), ns
			Mean spontaneous burst activity Male Before vs After: W = 144, p = 0.083 (adj. p = 0.25), ns
	Extended Figure 5j	LM with interaction + post hoc contrasts (emmeans, Tukey)	mean_Spontb ~ Archetype × Distance: $F_{(2,114)} = 3.71$, p = 0.027. Intercepts at Distance=0: A3 < A1 (trend, p = 0.052); A3 vs A2 (ns); A1 vs A2 (ns). n = 120.
Extended figure 9	Extended Figure 9b (left)	Unpaired t-test	$t_{(34)} = -5.61$; p < 0.0001
	Extended Figure 9b (right)	Unpaired t-test	$t_{(34)} = -1.49$; p = 0.15
	Extended Figure 9c (left)	Wilcoxon rank-sum	Distance to Store: Male vs Female, p = 0.67 (ns)
	Extended Figure 9c (right)	Wilcoxon rank-sum	Distance to Scrounger: Male vs Female, p = 0.00045 (***)
	not shown	Wilcoxon rank-sum	Distance to Worker: Male vs Female, p = 0.0001 (***)
	Extended Figure 9f	Paired t-test	Lever Press week 1 week 2: t(12)=–4.46, mean diff = –93.6, 95% CI [–139.3, –47.9], p=7.7×10^{-4}, ***
	Extended Figure 9j	Kruskal–Wallis + one-sided post hoc Wilcoxon (Holm)	Distance to Storer across 3 groups: $\chi^2_{(2)} = 6.59$, p = 0.037. Post hoc: group3 < group1 (p = 0.025), group3 < group2 (p = 0.039); group1 vs group2 (ns).
	Extended Figure9n	Kruskal–Wallis + one-sided post hoc Wilcoxon (Holm)	Distance to Scrounger across 3 groups: $\chi^2_{(2)} = 7.25$, p = 0.027. Post hoc: group3 < group2 (p = 0.018); group3 vs group1 (ns); group1 vs group2 (ns).
Extended figure 10	Extended Figure 10c	Wilcoxon rank-sum	W = 4204, p-value = 0.002241
	Extended Figure 10d	Wilcoxon rank-sum	W = 5229, p-value = 0.4476
	Extended Figure 10g	Wilcoxon rank-sum	W = 578.5, p-value = 3.918e-06
	Extended Figure 10h	Wilcoxon rank-sum	W = 837.5, p-value = 0.003618

Supplementary methods modeling

Building a behavioral model of e-mice for lone and social conditions

The environment of experiments was modeled as 6 states (Fig. 4 and Extended Data 6; rooms 1-4, and lever and dispenser positions). The number and sex of e-mice present in the environment was varied, with 1) one (male or female) e-mouse in the lone experiment, 2) three males or females e-mice in social experiments, or 3) 1 male and 2 females in the mixed-box experiment. In the model, we considered e-mouse satiety, to account for the convergence of pellets' consumption, n_{Eat}, toward a maximum, n_{Sat}. Motivation was considered to decrease with satiety:

$$p_{Motiv} = 1 - p_{Sat} = 1 - exp\left(-\alpha_{Sat}\left[1 - \frac{n_{Eat}}{n_{Sat}}\right]^{+}\right) \quad (1)$$

and scaled both action probabilities and the learning rate (see below). Moreover, e-mice were considered to encounter both pressing and eating fatigue in the model, to get paces consistent with those observed in real mice (without fatigue, paces were unrealistically high in the model). Both fatigue processes were described as action probabilities decreasing upon each individual action and restoring back to their maximal value with first-order dynamics:

$$\begin{cases} \tau_{PressFat}\frac{dp_{PressFat}}{dt} = p_{PressFat}^{max} - p_{PressFat} \\ p_{PressFat} \leftarrow p_{PressFat} - \Delta p_{PressFat}\delta\left(t - t_{Press}\right) \end{cases} \quad (2)$$

and

$$\begin{cases} \tau_{EatFat}\frac{dp_{EatFat}}{dt} = p_{EatFat}^{max} - p_{EatFat} \\ p_{EatFat} \leftarrow p_{EatFat} - \Delta p_{EatFat}\delta\left(t - t_{Eat}\right) \end{cases} \quad (3),$$

with δ being the Dirac delta function. Both fatigue processes and satiety concurred to decrease action probabilities in the model. When at the lever, e-mice pressed the lever with probability

$$p_{Press} = p_{PressFat}p_{Motiv} \quad (4).$$

The lever could not be pressed for 3 time steps after each press (*i.e.,* to mimic the 5 s lever unavailability in experiments). When at the dispenser, e-mice ate with probability:

$$p_{Eat} = p_{EatFat}p_{Motiv} \quad (5),$$

when pellets were present. State transition occurred at each time step, with softmax probabilities based on Q values of all accessible k states from the current state s_i:

$$p\left(s_i, s_j\right) = \frac{e^{\beta Q\left(s_i, s_j\right)}}{\sum_{\forall k} e^{\beta Q(s_i, s_k)}} \quad (6),$$

with β the inverse temperature parameter favoring either exploration (i.e., random choices; lower β values) or exploitation of previous learning (higher β values). Q learning occurred upon each transition from a departure state s_i to an accessible arriving state s_j, as

$$Q\left(s_i, s_j\right) \leftarrow (1-\alpha)\, Q\left(s_i, s_j\right) + \alpha\left(r\,w\,d_{s_j} + r\,w\,d_{s_j}^{future}\right) \quad (7),$$

with α the learning rate, $r\,w\,d_{s_j}$ the reward obtained at the arriving state s_j, and $r\,w\,d_{s_j}^{future}$ the maximal expected reward obtainable in states accessible from the arriving state s_j:

$$r\,w\,d_{s_j}^{future} = \gamma \arg\max_{\forall k}\left(Q\left(s_j, s_k\right)\right) \quad (8),$$

γ being a temporal discount factor. The learning rate scaled with motivation, which was considered to decrease with satiety,

$$\alpha = \alpha_0\left(1 - p_{Sat}\right) \quad (9).$$

As a consequence, learning stopped when e-mice were satiated. The Q matrix was initialized with value 1 at all elements corresponding to possible transitions (and 0 elsewhere). Consistent with observed trajectories in real mice, stops were forbidden at rooms 3 and 4, *i.e.*, on trajectories back and forth between lever and dispenser states.

Observables

The number of lever presses was computed, for each e-mice, from the beginning of the simulation up to its satiety. The percentage of complete sequences was computed as the number of direct sequences from lever to dispenser, following lever presses (within 3 times steps, i.e. 6 seconds, as in experiments), normalized by the number of lever presses:

$$p_{CompleteSequence} = 100\frac{n_{CompleteSequence}}{n_{Presses}} \quad (10).$$

These two observables are depicted in Fig. 3a, e and h and were used to categorize e-mice and e-mouse triads, as done in experiments (Fig. 3c, e, g, h, i). In each experimental condition (alone, social, mixed), an archetypal analysis of e-mice was achieved, as done for experiments. Also, both variables were normalized over e-mice within cages in social conditions:

$$\begin{cases} p_{CompleteSequence}^{Normalized}(i) = \dfrac{{p_{CompleteSequence}}^{(i)}}{\sum_{j=1}^{3} {p_{CompleteSequence}}^{(j)}} \\ n_{Presses}^{Normalized}(i) = \dfrac{{n_{Presses}}^{(i)}}{\sum_{j=1}^{3} {n_{Presses}}^{(j)}} \end{cases} \quad (11).$$

The gain of a given e-mouse was increased by 1 each time it ate a pellet within six seconds following a lever press by another e-mouse. The loss of a given e-mouse was increased by 1 when, in the six seconds following its lever press, a pellet was eaten by another e-mouse.

Reduced model of social interactions.

We built a reduced theoretical model (hereafter termed the *reduced model*) to assess, within a mathematically tractable framework, the causal mechanisms whereby specialized behaviors emerge under social interactions. We derived the reduced model from the full one, based on a continuous time version of Q dynamics [8], *i.e.*, ordinary differential equations (ODEs). In these ODEs, learning and behavioral dynamics operated at a slower time scale – compared to that of individual choices in the full model – so that actions (*i.e.*, state transitions, pressings, eatings) were described probabilistically in the reduced model. In this framework, we performed qualitative analysis of ODEs to determine the number and stability of fixed points of learning and behavioral (state) variables as a function of parameters (with a focus on β, which is essential in setting social interactions in the full model). Moreover, to reduce dimensionality for better tractability, we considered a simpler setup, where the environment contains only two positions for e-mice (the lever, *L,* and the food dispenser, *D*), only two e-mice (which allowed to assess social interactions), and we did not consider fatigue or satiety. Here, *q* dynamics were derived from equation (7), to account for their convergence to the expectation of present and future rewards (as for the dynamic of their discrete time equivalent Q values in the full model). Hence,

$$\begin{cases} \frac{dq_{DDi}}{dt} = \alpha\left(rwd_{Di} + rwd_{Di}^{Future} - q_{DDi}\right) \\ \frac{dq_{LDi}}{dt} = \alpha\left(rwd_{Di} + rwd_{Di}^{Future} - q_{LDi}\right) \\ \frac{dq_{DLi}}{dt} = \alpha\left(rwd_{Li} + rwd_{Li}^{Future} - q_{DLi}\right) \\ \frac{dq_{LLi}}{dt} = \alpha\left(rwd_{Li} + rwd_{Li}^{Future} - q_{LLi}\right) \end{cases} \quad (12),$$

with q_{XYi} being the q value of $X \to Y$ transition, and rwd_{Xi} and rwd_{Xi}^{Future} the present and future expected reward at position X in e-mouse i. Note that, despite all the simplifications made, the dynamical system under consideration is still 8-dimensional. Fortunately, it is straightforward to note that dynamics of q_{DDi} and q_{LDi} are identical (as that of q_{DLi} and q_{LLi}), so that, starting from identical null initial conditions

$$q_{LLi}(t=0) = q_{DLi}(t=0) = q_{LDi}(t=0) = q_{DDi}(t=0) = 0 \quad (13),$$

one has

$$\begin{cases} q_{LLi}(t) = q_{DLi}(t) \\ q_{DDi}(t) = q_{LDi}(t) \end{cases} \quad (14),$$

so that only one *q* value at each position is sufficient to describe dynamics:

$$\begin{cases} q_{Li} \equiv q_{LLi} = q_{DLi} \\ q_{Di} \equiv q_{DDi} = q_{LDi} \end{cases} \quad (15),$$

reducing the theory to a 4-dimensional problem. For each e-mouse, one could express *q* values evolved according to:

$$\begin{cases} \frac{dq_{Li}}{dt} = \alpha\left(rwd_{Li} + rwd_{Li}^{Future} - q_{Li}\right) \\ \frac{dq_{Di}}{dt} = \alpha\left(rwd_{Di} + rwd_{Di}^{Future} - q_{Di}\right) \end{cases} \quad (16).$$

The mean expectation of current reward at *L* was null ($rwd_{Li} = 0$), while that at *D* scaled with the probability of being at the dispenser, the probability of eating, the probability of available food at the dispenser, and the reward value of eating, which writes, for e-mouse #1, e.g.,

$$rwd_{D1} = p_{Eat}p_{D1}\underbrace{\left[\underbrace{p_{Press}\left(p_{L1}+p_{L2}\right)}_{p_{common\ production}} - \underbrace{p_{Eat}p_{D2}}_{p_{consumption\ e-mouse\ 2}}\right]^{+}}_{p_{food\ at\ Dispenser}} r \quad (17),$$

with p_{Xi} the residence probability of being at position X for mouse i, p_{Eat} and p_{Press} the probabilities of eating and pressing, and r the unit reward per eating. The positive part function, $[]^+$, is necessary to account for a positive probability of available food at the dispenser. Note also that because $p_{Di} > 0.5$ and $p_{Li} = 1 - p_{Di} < 0.5$ (as there is a reward at D and not at L), the probability of available food at the dispenser is inferior to 1. Note that equation (17) can be written for mouse #2's by swapping indices #1 and #2, as can be done for other equations in the following. Whenever possible, we keep the generic index i for e-mice. In addition, at both positions, the expectation of future rewards expresses as

$$rwd_{Di}^{Future} = \gamma\, rwd_{Di} \quad (18),$$

as getting a reward at *D* is the maximal expected possible reward at the next step from both positions *L* and *D*. Therefore, *q* derivatives can be re-expressed as

$$\begin{cases} \frac{dq_{L1}}{dt} = \alpha\left(\gamma p_{Eat}p_{D1}\left[p_{Press}\left(p_{L1}+p_{L2}\right) - p_{Eat}p_{D2}\right]^{+} r - q_{L1}\right) \\ \frac{dq_{D1}}{dt} = \alpha\left(\left(1+\gamma\right)p_{Eat}p_{D1}\left[p_{Press}\left(p_{L1}+p_{L2}\right) - p_{Eat}p_{D2}\right]^{+} r - q_{D1}\right) \end{cases} \quad (19).$$

In the following, we considered, for the sake of simplicity that $p_{press} = p_{eat} = 1$. Also, because $p_{Li} = 1\text{-}p_{Di}$ for both e-mice *i*, *q* derivatives simply expressed as

$$\begin{cases} \frac{dq_{L1}}{dt} = \alpha\left(\gamma p_{D1}\left[2 - p_{D1} - 2p_{D2}\right]^{+} r - q_{L1}\right) \\ \frac{dq_{D1}}{dt} = \alpha\left(\left(1+\gamma\right)p_{D1}\left[2 - p_{D1} - 2p_{D2}\right]^{+} r - q_{D1}\right) \end{cases} \quad (20).$$

Note that the first equation is actually the second one scaled by a factor $\frac{\gamma}{1+\gamma}$. As both q_{D1} and q_{L1} both start from null initial conditions (equation (13)), their dynamics are homothetic, so that at any point in time

$$q_{Li} = \frac{\gamma}{1+\gamma} q_{Di} \quad (21).$$

In the following, we proceed with the q_{D1} equation, which, together with that for q_{D2}, describes the system with two-dimensional dynamics:

$$\begin{cases} \frac{dq_{D1}}{dt} = \alpha\left(\left(1+\gamma\right)p_{D1}\left[2-p_{D1}-2p_{D2}\right]^{+}r-q_{D1}\right) \\ \frac{dq_{D2}}{dt} = \alpha\left(\left(1+\gamma\right)p_{D2}\left[2-p_{D2}-2p_{D1}\right]^{+}r-q_{D2}\right) \end{cases} \quad (22).$$

In such a setup, with two states, residence probabilities of e-mouse *i*, *i.e.,* p_{Li} and p_{Di} can be directly related to transition probabilities. Indeed, assuming first-order state transitions, p_{Di} evolves as:

$$\frac{dp_{Di}}{dt} = p_{Li}p_{LDi} - p_{Di}p_{DLi} \quad (23),$$

which, after substituting p_{Li} by $1-p_{Di}$ leads to

$$\frac{dp_{Di}}{dt} = p_{LDi} - \left(p_{LDi}+p_{DLi}\right)p_{Di} \quad (24),$$

with a unique fixed point

$$p_{Di} = \frac{p_{LDi}}{p_{LDi}+p_{DLi}} \quad (25),$$

so that replacing p_{LDi} and p_{DLi} by their softmax expression, leads to

$$p_{Di} = p_{LDi} = p_{DDi} = \frac{e^{\beta q_{LDi}}}{e^{\beta q_{LDi}}+e^{\beta q_{DLi}}} \quad (26)$$

and

$$p_{Li} = p_{DLi} = p_{LLi} = \frac{e^{\beta q_{DLi}}}{e^{\beta q_{LDi}}+e^{\beta q_{DLi}}} \quad (27),$$

which, thanks to equation (15), can simply be written

$$p_{Di} = \frac{e^{\beta q_{Di}}}{e^{\beta q_{Li}}+e^{\beta q_{Di}}} \quad (28)$$

and

$$p_{Li} = \frac{e^{\beta q_{Li}}}{e^{\beta q_{Li}}+e^{\beta q_{Di}}} \quad (29).$$

Combining equations (21) and (28) allowed to express the residence probability as a function of the q value at dispenser:

$$p_{Di} = \left(1+e^{-\frac{\beta}{1+\gamma}q_{Di}}\right)^{-1} \quad (30).$$

Then, combining equations (22) and (30) yields

$$\begin{cases} \frac{dq_{D1}}{dt} = \alpha\left(\frac{1+\gamma}{1+e^{-\frac{\beta}{1+\gamma}q_{D1}}}\left[2-\frac{1}{1+e^{-\frac{\beta}{1+\gamma}q_{D1}}}-\frac{2}{1+e^{-\frac{\beta}{1+\gamma}q_{D2}}}\right]^{+}r-q_{D1}\right) \\ \frac{dq_{D2}}{dt} = \alpha\left(\frac{1+\gamma}{1+e^{-\frac{\beta}{1+\gamma}q_{D2}}}\left[2-\frac{1}{1+e^{-\frac{\beta}{1+\gamma}q_{D2}}}-\frac{2}{1+e^{-\frac{\beta}{1+\gamma}q_{D1}}}\right]^{+}r-q_{D2}\right) \end{cases} \quad (31).$$

Note that this system describes the social interaction between two e-mice having similar β values (*i.e.,* two males or females). This two-dimensional ODE system describes the time evolution of learning and behavior variables (as residence, lever press and direct sequence probabilities can be derived from *q* values, see below). Unfortunately, this system (even studied by parts, depending on whether probabilities of food at dispenser are positive or null) has no analytical solution. We therefore assessed a version of the ODEs with linearized approximations of residence probabilities. This linearized version could be solved analytically, which allowed to assess the number and stability of fixed points as a function of β. To do so, we used a linearized approximation of residence probabilities at position *X* in e-mouse *i*, *i.e.*

$$p_{Xi} = \frac{1}{2} + \frac{\beta}{4\left(1+\gamma\right)} q_{Xi} \qquad (32).$$

Such an approximation was typically of the order of one to a few percent. Combined with equation (22), it allowed to provide the tractable system we analyzed to generate Figs. 4 et Extended Data 8 and which, after a few lines of algebraic manipulation, expresses as:

$$\begin{cases} \frac{dq_{D1}}{dt} = \alpha\left(\frac{1+\gamma}{4}\left(1+\frac{\beta q_{D1}}{2(1+\gamma)}\right)\left[1-\frac{\beta(q_{D1}+2q_{D2})}{2(1+\gamma)}\right]^{+} r - q_{D1}\right) \\ \frac{dq_{D2}}{dt} = \alpha\left(\frac{1+\gamma}{4}\left(1+\frac{\beta q_{D2}}{2(1+\gamma)}\right)\left[1-\frac{\beta(q_{D2}+2q_{D1})}{2(1+\gamma)}\right]^{+} r - q_{D2}\right) \end{cases} \qquad (33).$$

In the reduced model, the propensity of e-mouse i to press the lever was expressed as a probability, which scaled both with the probability to be at the lever position and the probability to press, once at that position:

$$p_{LPi} = p_{Press} p_{Li} = p_{Press}\left(1 - p_{Di}\right) \qquad (34).$$

The propensity to achieve direct sequences was also expressed as a probability, scaling both with the be at the lever position, to press the lever and then move toward the dispenser position (which is equal to the residence probability at dispenser (equation (26)):

$$p_{Si} = p_{Press} p_{Li} p_{Di} = p_{Press}\left(1 - p_{Di}\right) p_{Di} \qquad (35).$$

Simulation of the reduced model

The reduced model was simulated using a noisy version of system (19):

$$\begin{cases} \frac{dq_{L1}}{dt} = \alpha\left(\gamma\, p_{Eat} p_{D1}\left[p_{Press}\left(p_{L1}+p_{L2}\right) - p_{Eat} p_{D2}\right]^{+} r - q_{L1}\right) + \sigma_{Noise}\sqrt{dt}\epsilon(t) \\ \frac{dq_{D1}}{dt} = \alpha\left(\left(1+\gamma\right) p_{Eat} p_{D1}\left[p_{Press}\left(p_{L1}+p_{L2}\right) - p_{Eat} p_{D2}\right]^{+} r - q_{D1}\right) + \sigma_{Noise}\sqrt{dt}\epsilon(t) \end{cases} \qquad (36),$$

where σ_{noise} is the noise standard deviation, and ε(t) is a Gaussian stochastic process with zero mean and unit standard deviation, similar equations for e-mouse #2 and residence probabilities computed according to equation (30). Simulations were used to check the validity of the qualitative analyses achieved to produce bifurcation diagrams and illustrate system dynamics in phase portraits (see below).

Qualitative analysis of the reduced model

The system (33) admitted one or several fixed points $\left(q_{D1}^{\star}, q_{D2}^{\star}\right)$, which number and nature depended on parameters. We focused on the dependence of $q_{Di}^{\star}$ steady-states of individual e-mice on β, *i.e.* $q_{Di}^{\star}\left(\beta\right)$, denoted *branches* in the following. Indeed, as in the full model, the value of β separated distinct regimes of social interaction between e-mice. The main results of the qualitative analysis of the system are recapitulated in Table 1.

We found that for β values below a critical value β_{Bifurc} (see below, for the analytical calculation of critical values), only one fixed point, q_{Mono}^{FP}, existed, with both e-mice admitting a similar *q* steady-state:

$$q_{Mono}^{FP} = \left(q_{Mono}^{\star}, q_{Mono}^{\star}\right) \qquad (37)$$

with

$$q_{Mono}^{\star} = \frac{2\left(1+\gamma\right)}{3\beta^2 r}\left(2\sqrt{\beta^2 r^2 + 2\beta r + 4} - \left(\beta r + 4\right)\right) \quad (38).$$

Its stability was set by

$$\lambda_{Mono}^{1,2} = -\alpha\left[1 + \frac{\beta r}{8\left(1+\gamma\right)}\left(2\beta q_{Mono}^{\star} \pm \left(2\left(1+\gamma\right) + \beta q_{Mono}^{\star}\right)\right)\right] \qquad (39),$$

with both Eigen values negative, so that q_{Mono}^{FP} was a stable node.

Above β_{Bifurc}^{+}, λ_{Mono}^{2} was positive, so that q_{Mono}^{FP} became a saddle node. Also, two other symmetric fixed points appeared, $q_{Bist}^{FP-A} = \left(q_{Bist}^{Up\star}, q_{Bist}^{Down\star}\right)$ and $q_{Bist}^{FP-B} = \left(q_{Bist}^{Down\star}, q_{Bist}^{Up\star}\right)$, with

$$q_{Bist}^{Up,Down\star} = \frac{2\left(1+\gamma\right)}{\beta^2 r}\left(\beta r - 4 \pm 2\sqrt{\beta^2 r^2 - 6\beta r + 4}\right) \qquad (40).$$

The fixed point q_{Bist}^{FP-A} corresponded to the case in which e-mouse #1 displayed a higher *q* value at dispenser ($q_{Bist}^{Up\star}$), and, as a consequence, a higher probability of residence ($p_{Bist}^{Up\star}$) at dispenser, and lower probabilities of lever press ($p_{Bist}^{LP\star}$) and direct sequences ($p_{Bist}^{CS\star}$), compared to e-mouse #2 (see below). In other words, q_{Bist}^{FP-A} was a fixed point at which e-mouse #1 behaved as a *Scrounger* and e-mouse #2 as a *Worker*, while behaviors were inverted at q_{Bist}^{FP-B}. The stability of both fixed points was set by

$$\lambda_{Bist}^{1,2} = \alpha\left[1 - \frac{\beta r}{2} \pm \frac{\beta r}{8\left(1+\gamma\right)}\sqrt{\left(2\left(1+\gamma\right) + \beta q_{Bist}^{Up\star}\right)\left(2\left(1+\gamma\right) + \beta q_{Bist}^{Down\star}\right)}\right] \qquad (41),$$

with both Eigen values negative, so that both q_{Bist}^{FP-A} and q_{Bist}^{FP-B} were stable nodes.

For β values above a second critical value β_{NoFood} (with $\beta_{NoFood} > \beta_{Bifurc}$, see below), the probability of available food at the dispenser became null for the *Worker* e-mouse (i.e., the argument within the positive part function, $[]^{+}$, becomes negative for e-mouse #1 or #2 in equation (33)). The system thus writes, if the worker is e-mouse #2:

$$\begin{cases} \frac{dq_{D1}}{dt} = \alpha\left(\frac{1+\gamma}{4}\left(1+\frac{\beta q_{D1}}{2(1+\gamma)}\right)\left[1-\frac{\beta(q_{D1}+2q_{D2})}{2(1+\gamma)}\right]^{+} r - q_{D1}\right) \\ \frac{dq_{D2}}{dt} = -\alpha q_{D2} \end{cases} \quad (42),$$

A single fixed point $q_{NoFood}^{FP-A} = \left(q_{NoFood}^{Up\star}, q_{NoFood}^{Down\star}\right)$ existed

$$\begin{cases} q_{NoFood}^{Up\star} = \frac{2(1+\gamma)}{\beta^2 r}\left(\sqrt{\beta^2 r^2 + 16} - 4\right) \\ q_{NoFood}^{Down\star} = 0 \end{cases} \quad (43),$$

which stability was determined by

$$\begin{cases} \lambda_{NoFood}^{Up} = -\frac{\alpha}{4}\sqrt{\beta^2 r^2 + 16} \\ \lambda_{NoFood}^{Down} = -\alpha \end{cases} \quad (44),$$

with both Eigen values negative, so that q_{NoFood}^{FP-A} was a stable node. Obviously, a symmetrical stable node, $q_{NoFood}^{FP-B} = \left(q_{NoFood}^{Down\star}, q_{NoFood}^{Up\star}\right)$, co-existed with q_{NoFood}^{FP-A}, when the worker was e-mouse #1.

Both β_{Bifurc} and β_{NoFood} critical values could be found analytically. At the bifurcation point where q_{Mono}^{FP} becomes a saddle node, the two symmetric fixed points $q_{Bist}^{FP-A} = \left(q_{Bist}^{Up\star}, q_{Bist}^{Down\star}\right)$ and $q_{Bist}^{FP-B} = \left(q_{Bist}^{Down\star}, q_{Bist}^{Up\star}\right)$ switch from complex ($\beta<\beta_{Bifurc}$) to real ($\beta\geq\beta_{Bifurc}$) values. Thus, this point was set by β by a simple criterion, i.e., canceling of the term under the square root function in equation (40):

$$\beta_{Bifurc}^2 r^2 - 6\beta_{Bifurc} r + 4 = 0 \quad (45),$$

which yields

$$\beta_{Bifurc}^{+/-} = r^{-1}\frac{6 \pm \sqrt{20}}{2} \quad (46),$$

with $\beta_{Bifurc}^{-} = r^{-1}\frac{6-\sqrt{20}}{2}$ separating branches of real negative $q_{Bist}^{Up,Down\star}$ values from the branch of complex values between β_{Bifurc}^{-} and β_{Bifurc}^{+}. The branch of real negative $q_{Bist}^{Up,Down\star}$ values below β_{Bifurc}^{-} corresponded to impossible solutions, as q values are positive (with a null initial condition (equation (13)) and positive reward expectations). Finally,

$$\beta_{Bifurc} \equiv \beta_{Bifurc}^{+} = r^{-1}\frac{6+\sqrt{20}}{2} \cong 5.2361 \quad (47),$$

was the critical point (assuming r = 1 ; see parameter values, below) at which the system switched from one to three fixed points (two stable ones separated by an unstable one), at a supercritical pitchfork bifurcation (trifurcation[8]). At this point, where $q_{Bist}^{Up\star} = q_{Bist}^{Down\star}$, i.e., at which $q_{Bist}^{Up,Down\star} = \frac{2(1+\gamma)}{\beta^2 r}(\beta r - 4)$, one could easily check, using equation (39), that $\lambda_{Mono}^2 = 0$., *i.e.*, q_{Mono}^{FP} loses its stability along the second Eigen direction (*i.e.* switching from a stable to a saddle node).

The other critical value, β_{NoFood}, could be found by considering that $q_{Bist}^{Down\star} = 0$ at that point, *i.e.*,

$$\beta_{NoFood} r - 4 \pm 2\sqrt{\beta_{NoFood}^2 r^2 - 6\beta_{NoFood} r + 4} = 0 \qquad (48),$$

which yields

$$\beta_{NoFood} = r^{-1}\frac{16}{3} \cong 5.3333 \qquad (49).$$

Bifurcation diagrams were built (Fig.4 et Extended Data Fig.8) by representing $q_{Mono}^{\star}$ stable and saddle branches (below and above β_{Bifurc}, respectively, $q_{Bist}^{Up\star}$ and $q_{Bist}^{Down\star}$branches for $\beta_{Bifurc} \leqslant \beta < \beta_{NoFood}$, and $q_{NoFood}^{Up\star}$ and $q_{NoFood}^{Down\star}$ branches for $\beta \geqslant \beta_{NoFood}$.

We verified that bifurcation diagrams obtained using linearized approximation of residence probabilities (equation (32)) were indeed qualitatively correct and quantitatively consistent with the original reduced model, using numerical simulations of the reduced model. To do so, the stable and bistable branches were computed by taking steady-state q_{Bi} values of these simulations, and computing p_{Bi}, p_{LPi}, and p_{Si} from equations (30), (34) and (35), respectively. The q_{Mono}^{FP} saddle node branch was acquired by doing the same, in the absence of noise (σ_{noise} = 0), so that system's state converged to the saddle node from its null initial condition (*i.e.*, along the first bisector in the (q_{B1}, q_{B2}) plane, due to system' symmetry) and stayed at it, without further diverging toward q_{Bist}^{FP-A} or q_{Bist}^{FP-B} (which occurred in the presence of noise), as Eigen directions are stable along the first bisector and unstable along the second bisector (see below).

Eigenvectors of the reduced model

Eigenvectors in the general case expressed as

$$V_{1,2}^{Mono,Bist} = \begin{pmatrix} \pm\dfrac{\sqrt{\left(2(1+\gamma)+\beta q_{B1}^{\star}\right)\left(2(1+\gamma)+\beta q_{B2}^{\star}\right)}}{2(1+\gamma)+\beta q_{B2}^{\star}} \\ 1 \end{pmatrix} \qquad (50),$$

which, in the case of the $q_{Mono}^{FP} = (q_{Mono}^{\star}, q_{Mono}^{\star})$ branch, trivially simplifies to

$$V_{1,2}^{Mono} = \begin{pmatrix} \pm 1 \\ 1 \end{pmatrix} \qquad (51).$$

In the case with no food for one e-mouse, eigenvectors expressed as

$$V_1^{NoFood} = \begin{pmatrix} -\dfrac{2(1+\gamma)+\beta q_{B1}^{\star}}{\beta\left(q_{B1}^{\star}+q_{B2}^{\star}\right)} \\ 1 \end{pmatrix} \quad V_2^{NoFood} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad (52)$$

which, in the case of the q_{NoFood}^{FP-A}, simplified to

$$V_1^{NoFood} = \begin{pmatrix} -\dfrac{\beta r + 4 + \sqrt{\beta^2 r^2 + 16}}{\beta r} \\ 1 \end{pmatrix} \quad V_2^{NoFood} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad (53).$$

Phase portraits of the reduced model

In phase portraits (Fig.4 et Extended Data Fig.8), the *q* vector field was numerically computed from equation (19), with q_{Di} values set within a fixed range and q_{Li} computed from equation (21).

The p vector field was computed with p_{Di} values set within a fixed range, q_{Di} values computed as

$$q_{Di} = \frac{1+\gamma}{\beta} log\left(\frac{p_{Di}}{1-p_{Di}}\right) \quad (54),$$

i.e., the reciprocal equation of equation (30), q_{Li} computed from q_{Di} with equation (21) and p derivatives computed thanks to the chain rule, i.e.,

$$\frac{dp_{Di}}{dt} = \frac{dp_{Di}}{dq_{Di}}\frac{dq_{Di}}{dt} \quad (55),$$

where

$$\frac{dp_{Di}}{dq_{Di}} = \frac{\beta e^{\beta\frac{\gamma q_{D1}}{1+\gamma}} e^{\beta q_{D1}}}{(1+\gamma)\left(e^{\frac{\gamma\beta}{1+\gamma}q_{D1}} + e^{\beta q_{D1}}\right)^2} \quad (56)$$

was obtained by differentiating equation (30) and $\frac{dq_{Di}}{dt}$ taken from system (19).

The p_{LP} vector field was computed with p_{LPi} values set within a fixed range, p_{Di} values computed as

$$p_{Di} = 1 - \frac{p_{LPi}}{p_{Press}} \quad (57),$$

which was derived from equation (34), q_{Di} values with equation (21), $\frac{dq_{Di}}{dt}$ from system (19), $\frac{dp_{Di}}{dq_{Di}}$ from equation (56), and by applying the chain rule, *i.e.*,

$$\frac{dp_{LPi}}{dt} = \frac{dp_{LPi}}{dp_{Di}}\frac{dp_{Di}}{dq_{Di}}\frac{dq_{Di}}{dt} \quad (58),$$

with

$$\frac{dp_{LPi}}{dp_{Di}} = -p_{Press} \quad (59),$$

which was obtained by differentiating equation (34).

The p_S vector field was computed as the p_{LP} vector field, using equations (34), (21), (19) and (56) and applying the chain rule

$$\frac{dp_{Si}}{dt} = \frac{dp_{Si}}{dp_{Di}}\frac{dp_{Di}}{dq_{Di}}\frac{dq_{Di}}{dt} \quad (60),$$

with

$$\frac{dp_{Si}}{dp_{Di}} = -p_{Press}(1-2p_{Di}) \quad (61),$$

which was obtained by differentiating equation (35).

In phase portraits, vector fields were represented both by a direction field plot and through a color coded map. Nullclines were estimated as zero level contours of the vector fields. Fixed points were taken with bistable Up and Down steady-state values obtained from the bifurcation diagram for $\beta = \beta_{Female}$ (social female condition), $\beta = \beta_{Male}$ (social male condition). In the mixed box condition, we considered one female and one male, and the fixed point was taken as the intersection of both nullclines. Eigenvectors were not represented for clarity. An example of simulation of reduced model's dynamics was also represented in each condition to show dynamics from initial conditions toward the – or one of the – fixed points.

Numerical procedures

Numerical simulations and qualitative analysis were achieved thanks to Matlab custom scripts that are available online. The qualitative analysis of the reduced model was done using the Symbolic toolbox of Matlab. Numerical simulations of the full and reduced model were performed used Euler forward integration. In the full model, standard parameters, unless mentioned, were α_0 = 0.05, γ = 0.99, α_{Sat} = 20, n_{Sat} = 150, $\tau_{PressFat} =$ 50s, $p^{max}_{PressFat} = 1$, $\Delta p_{PressFat}$ = 0.5, τ_{EatFat} = 50s, $p^{max}_{EatFat} = 1$, Δp_{EatFat} = 0.5, and the time step was Δt = 2s. Simulations lasted t_{max} = 20000s. β distributions are discussed in the main text. Parameters values in the reduced model were as following: α = 0.01, β = [1,10], β_{Female} = 1, β_{Male} = 10, γ = 0.99, r = 1, p_{press} = 1, p_{eat} = 1, σ_{noise} = $1e^{-4}$, dt = 1s. Simulations lasted $t_{max} = 1500s$ for example simulations shown in phase portraits and t_{max} = 100000s in simulations with noise, i.e., a much longer time to estimate steady-states in the numerical estimation of phase portraits. The distribution of e-mouse behavioral profiles and box type distributions were obtained on 1000 repetitions of 16 boxes of either the lone condition, or single sex or mixed sex conditions, with 14 animals per box (Fig. 4, g, and Extended Data Fig. 7).

Table 2

The table recapitulates the name of the fixed-points (first column), steady-states for both e-mice and their eigenvalues (second column), as well as their value or type, depending of the ranges of β. NA: not applicable.

<table>
<tr><td colspan="2">model</td><td colspan="2">general case
(food available for both e-mice)</td><td>no food available
for e-mouse #2 (e.g.)</td></tr>
<tr><td colspan="2">β</td><td>$\beta < \beta_{Bifurc}$</td><td>$\beta_{Bifurc} \leqslant \beta < \beta_{NoFood}$</td><td>$\beta \geqslant \beta_{NoFood}$</td></tr>
<tr><td rowspan="5">q^{PF}_{Mono}</td><td>$q^{\star}_{B1} = q^{\star}_{Mono}$</td><td rowspan="2">$> 0$</td><td colspan="2" rowspan="2">> 0</td></tr>
<tr><td>$q^{\star}_{B2} = q^{\star}_{Mono}$</td></tr>
<tr><td>λ^{1}_{Mono}</td><td rowspan="2">< 0</td><td colspan="2">< 0</td></tr>
<tr><td>λ^{2}_{Mono}</td><td colspan="2">> 0</td></tr>
<tr><td>type</td><td>stable node</td><td colspan="2">saddle node</td></tr>
<tr><td rowspan="5">q^{PF-A}_{Bist}</td><td>$q^{\star}_{B1} = q^{Up\star}_{Bist}$</td><td rowspan="5">no real positive solution</td><td rowspan="2">> 0</td><td rowspan="5">NA</td></tr>
<tr><td>$q^{\star}_{B2} = q^{Down\star}_{Bist}$</td></tr>
<tr><td>λ^{1}_{Bist}</td><td rowspan="2">< 0</td></tr>
<tr><td>λ^{2}_{Bist}</td></tr>
<tr><td>type</td><td>stable node</td></tr>
<tr><td rowspan="5">$q^{PF-A}_{BistNoF}$</td><td>$q^{\star}_{B1} = q^{Up\star}_{BistNoF}$</td><td colspan="2" rowspan="5">NA</td><td>> 0</td></tr>
<tr><td>$q^{\star}_{B2} = q^{Down\star}_{BistNoF}$</td><td>$0$</td></tr>
<tr><td>$\lambda^{1}_{BistNoF}$</td><td rowspan="2">< 0</td></tr>
<tr><td>$\lambda^{2}_{BistNoF}$</td></tr>
<tr><td>type</td><td>stable node</td></tr>
</table>

References

1. Turiault, M. *et al.* Analysis of dopamine transporter gene expression pattern -- generation of DAT-iCre transgenic mice. *Febs J* **274**, 3568 3577 (2007).

2. Chaumont, F. de *et al.* Real-time analysis of the behaviour of groups of mice via a depth-sensing camera and machine learning. *Nature biomedical engineering* **7**, 462–942 (2019).

3. Cutler, A. & Breiman, L. Archetypal Analysis. *Technometrics* **36**, 338–347 (1994).

4. Hart, Y. *et al.* Inferring biological tasks using Pareto analysis of high-dimensional data. *Nature Methods* **12**, 233-5–3 p following 235 (2015).

5. Ungless, M. A. & Grace, A. A. Are you or aren't you? Challenges associated with physiologically identifying dopamine neurons. *Trends Neurosci* **35**, 422 430 (2012).

6. Exley, R. *et al.* Distinct contributions of nicotinic acetylcholine receptor subunit alpha4 and subunit alpha6 to the reinforcing effects of nicotine. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7577–7582 (2011).

7. Mameli-Engvall, M. *et al.* Hierarchical control of dopamine neuron-firing patterns by nicotinic receptors. *Neuron* **50**, 911–921 (2006).

8. Strogatz, S. H. *Nonlinear Dynamics and Chaos : With Applications to Physics, Biology, Chemistry, and Engineering.* (1994).