

Supplementary material

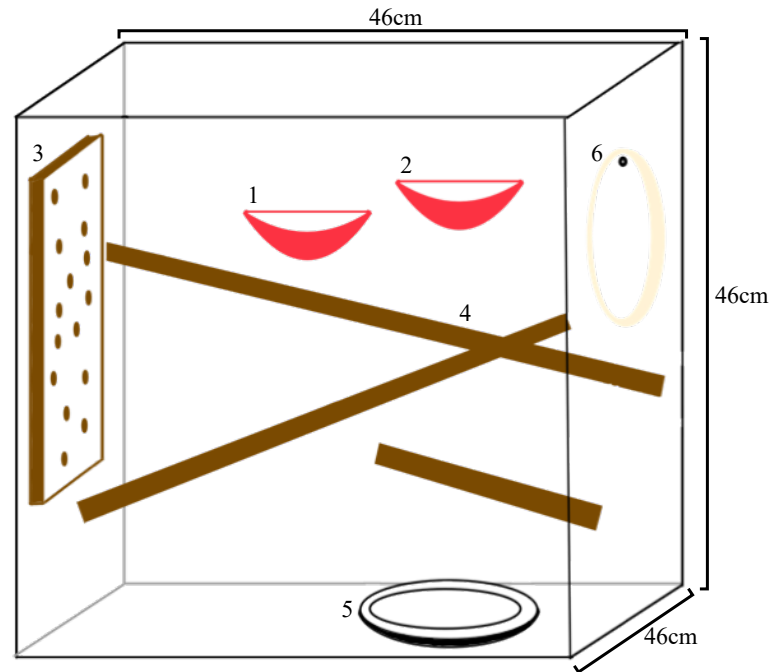


Fig. S1 Diagram of the individual housing cages (46 x 46 x 46 cm) with a food dish (1), water dish (2), caching block (3), three perches (4), a shallow bowl of sand (5), and a cuttlebone (6).

Table S1. Collected data from the individuals.

Bird color bands	Spatial learning data	Cognitive flexibility data	Long-term memory data	Plumage data	Songs data
Red/	Included	Included	Included	Included	Not included
Orange/	Included	Included	Included	Included	Included
Yellow/	Included	Included	Included	Included	Included
Green/	Included	Included	Included	Included	Included
Blue/	Included	Included	Included	Included	Included
Light blue/	Included	Included	Included	Included	Included

Brown/	Included	Included	Included	Included	Included
Grey/	Included	Included	Included	Included	Included
Orange/ Orange	Included	Included	Included	Included	Included
Grey/Grey	Included	Not included	Not included	Included	Included
Black/ Black	Included	Included	Included	Included	Included
Red/ Yellow	Included	Included	Included	Included	Included
Red/Blue	Included	Included	Included	Not included	Not included
Red/White	Included	Included	Included	Included	Included
Red/ Brown	Included	Included	Included	Included	Included
Orange/ Red	Included	Included	Included	Included	Included

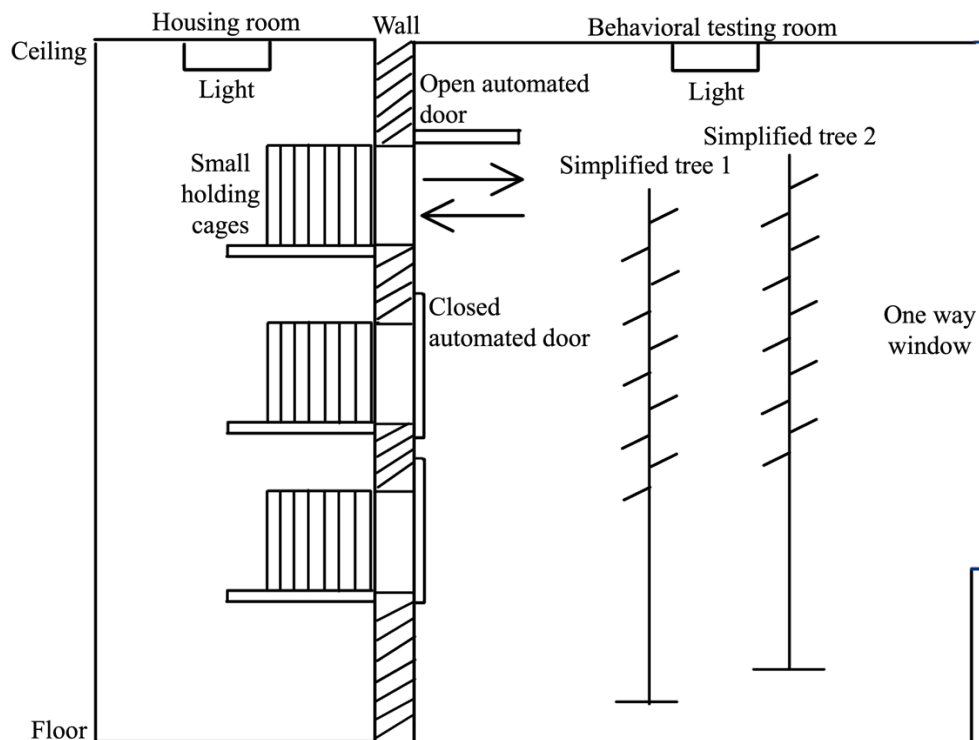


Fig. S2 Diagram of the experimental setup: side, eye-level view. On the left, the housing room with the small holding cages (30.5 x 30.5 x 40.5 cm) connected to automated doors that open into the behavioral testing room (right, 2.8 x 2.8 m). Diagram depicts simplified trees, the automated doors and a one-way window to observe the subjects without disturbance. The birds are moved from one room to the other via light manipulation. Lights in holding room are turned off to encourage birds to move into the behavioral room (lights on) and vice versa for returning them to their holding cages.

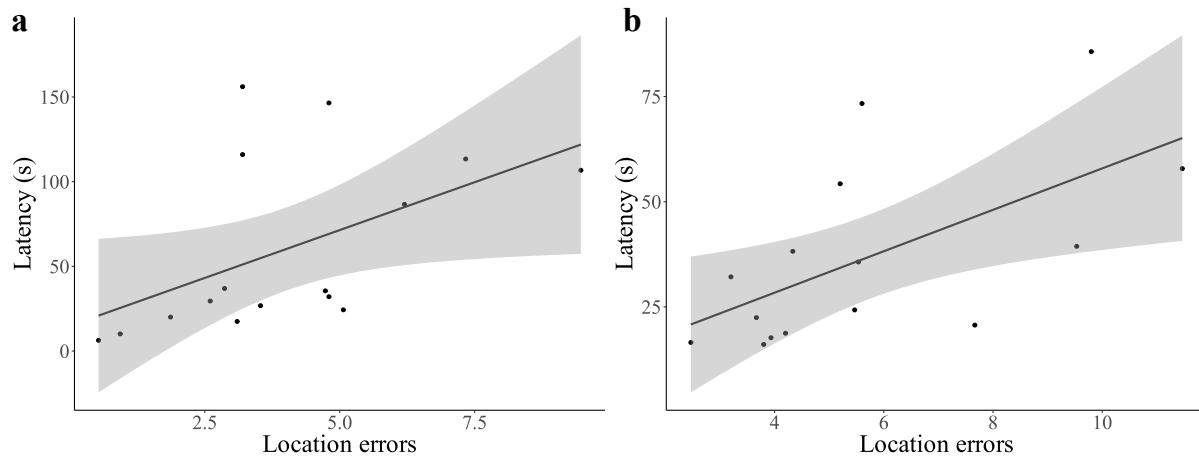


Fig. S3 Relationship between the mean number of location errors and the mean latency in the (a) spatial learning task and the (b) reversal learning task.

Table S2. Summary of the principal component analysis of plumage variables.

	Standard deviation	Proportion of variance	Cumulative Proportion
PC1	0.170	0.423	0.423
PC2	0.132	0.256	0.679
PC3	0.094	0.129	0.809
PC4	0.088	0.113	0.921

Table S3. Loadings for all variables principal component analysis of plumage variables.

	PC1	PC2	PC3	PC4
Mean brightness Bib	0.982	0.018	0.096	0.107
UV Chroma Bib	0.009	0.013	0.007	0.005
Mean brightness Mantle	0.056	-0.014	-0.479	0.261
UV Chroma Mantle	0.005	-0.019	-0.034	0.052

Mean brightness Cap	-0.117	-0.175	0.033	0.063
UV Chroma Cap	0.011	0.018	0.025	0.012
Mean brightness Nape	-0.134	0.164	0.354	0.906
UV Chroma Nape	-0.007	0.007	0.021	0.001
Mean brightness Breast	0.017	0.219	-0.786	0.232
UV Chroma Breast	-0.009	0.008	0.021	-0.020
Mean brightness Cheek	-0.021	0.945	0.116	-0.196
UV Chroma Cheek	0.013	-0.004	0.012	0.002

Table S4. F statistic and *p* value of the non-significant models applied in the statistical analysis of the plumage variables.

Model	F statistic	<i>p</i> value
B2 Bib ~ mean errors spatial learning + mean latency spatial learning	$F_{2,12} = 0.84$	0.46
B2 Breast ~ mean errors spatial learning + mean latency spatial learning	$F_{2,12} = 0.62$	0.55
B2 Nape ~ mean errors spatial learning + mean latency spatial learning	$F_{2,12} = 0.85$	0.45
S1U White patches ~ mean errors spatial learning + mean latency spatial learning	$F_{2,12} = 2.56$	0.12
B2 Dark patches ~ mean errors spatial learning + mean latency spatial learning	$F_{2,12} = 0.25$	0.78

S1U Dark patches ~ mean errors spatial learning + mean latency spatial learning	$F_{2,12} = 2.01$	0.18
Contrast Bib-Breast ~ mean errors spatial learning + mean latency spatial learning	$F_{2,12} = 0.71$	0.51
B2 Bib ~ visits to the old location	$F_{1,12} = 0.41$	0.53
B2 Breast ~ visits to the old location	$F_{1,12} = 5.40$	0.04
B2 Cheek ~ visits to the old location	$F_{1,12} = 0.52$	0.48
B2 Nape ~ visits to the old location	$F_{1,12} = 0.13$	0.73
B2 Bib ~ mean errors reversal learning + mean latency reversal learning	$F_{2,11} = 0.65$	0.54
B2 Cheek ~ mean errors reversal learning + mean latency reversal learning	$F_{2,11} = 0.66$	0.54
B2 Breast ~ mean errors reversal learning + mean latency reversal learning	$F_{2,11} = 0.97$	0.41
B2 Nape ~ mean errors reversal learning + mean latency reversal learning	$F_{2,11} = 1.11$	0.36
B2 White patches ~ mean errors reversal learning + mean latency reversal learning	$F_{2,11} = 0.13$	0.88
B2 Dark patches ~ mean errors reversal learning + mean latency reversal learning	$F_{2,11} = 0.93$	0.43
S1U White patches ~ mean errors reversal learning + mean latency reversal learning	$F_{2,11} = 3.74$	0.06

S1U Dark patches ~ mean errors reversal learning + mean latency reversal learning	$F_{2,11} = 3.41$	0.07
B2 White patches ~ visits to the old location	$F_{1,12} = 2.98$	0.11
B2 Dark Patches ~ visits to the old location	$F_{1,12} = 0.28$	0.60
S1U White patches ~ visits to the old location	$F_{1,12} = 2.61$	0.13
S1U Dark Patches ~ visits to the old location	$F_{1,12} = 0.09$	0.77
Contrast Bib-Cheek ~ visits to the old location	$F_{1,12} = 0.65$	0.43
Contrast Cheek-Cap ~ visits to the old location	$F_{1,12} = 0.61$	0.45
Contrast Cheek-Nape ~ visits to the old location	$F_{1,12} = 0.59$	0.46
Contrast Cheek-Mantle ~ visits to the old location	$F_{1,12} = 0.13$	0.72
Contrast Bib-Breast ~ visits to the old location	$F_{1,12} = 6.07$	0.03
Contrast Bib-Breast ~ mean errors reversal learning + mean latency reversal learning	$F_{2,11} = 0.67$	0.53
Contrast Bib-Cheek ~ mean errors reversal learning + mean latency reversal learning	$F_{2,11} = 0.86$	0.45
Contrast Cheek-Cap ~ mean errors reversal learning + mean latency reversal learning	$F_{2,11} = 0.60$	0.57
Contrast Cheek-Nape ~ mean errors reversal learning + mean latency reversal learning	$F_{2,11} = 0.50$	0.62
Contrast Cheek-Mantle ~ mean errors reversal learning + mean latency reversal learning	$F_{2,11} = 0.99$	0.40
B2 Bib ~ mean errors long-term memory + mean latency long-term memory	$F_{2,11} = 0.37$	0.70

B2 Cheek ~ mean errors long-term memory + mean latency long-term memory	$F_{2,11} = 0.56$	0.58
B2 Breast ~ mean errors long-term memory + mean latency long-term memory	$F_{2,11} = 3.70$	0.06
B2 Nape ~ mean errors long-term memory + mean latency long-term memory	$F_{2,11} = 0.50$	0.62
B2 White patches ~ mean errors long-term memory + mean latency long-term memory	$F_{2,11} = 0.79$	0.48
B2 Dark patches ~ mean errors long-term memory + mean latency long-term memory	$F_{2,11} = 0.52$	0.61
S1U White patches ~ mean errors long-term memory + mean latency long-term memory	$F_{2,11} = 0.07$	0.93
S1U Dark patches ~ mean errors long-term memory + mean latency long-term memory	$F_{2,11} = 0.89$	0.44
Contrast Bib-Breast ~ mean errors long-term memory + mean latency long-term memory	$F_{2,11} = 2.55$	0.12
Contrast Bib-Cheek ~ mean errors long-term memory + mean latency long-term memory	$F_{2,11} = 0.57$	0.58
Contrast Cheek-Cap ~ mean errors long-term memory + mean latency long-term memory	$F_{2,11} = 0.71$	0.51
Contrast Cheek-Nape ~ mean errors long-term memory + mean latency long-term memory	$F_{2,11} = 0.57$	0.58

Contrast Cheek-Mantle ~ mean errors long-term memory + mean latency long-term memory	$F_{2,11} = 0.51$	0.61
--	-------------------	------

Table S5. Summary of the principal component analysis for the mean of the song variables.

	Standard deviation	Proportion of variance	Cumulative Proportion
PC1	0.227	0.547	0.547
PC2	0.163	0.283	0.831
PC3	0.088	0.083	0.913

Table S6. Loadings for all variables of the principal component analysis for the mean of the song variables.

	PC1	PC2	PC3
Delta frequency <i>fee</i> note	-0.881	-0.128	-0.401
Delta latency <i>fee</i> note	0.069	-0.003	-0.040
Frequency 5 % <i>fee</i> note	-0.162	-0.104	0.435
Frequency 95 % <i>fee</i> note	-0.185	-0.088	0.388
Peak frequency <i>fee</i> note	-0.169	-0.093	0.429
Delta frequency <i>bee</i> note	-0.236	0.936	0.198
Delta latency <i>bee</i> note	0.005	-0.029	-0.085
Frequency 5 % <i>bee</i> note	-0.109	-0.190	0.282
Frequency 95 % <i>bee</i> note	-0.121	-0.128	0.275
Peak frequency <i>bee</i> note	-0.118	-0.151	0.266
Song length	0.059	-0.016	-0.056

<i>fee</i> note length	0.128	0.009	-0.055
<i>bee</i> note length	-0.105	-0.055	-0.159
Glissando ratio	-0.047	-0.002	-0.061
Interval ratio	-0.023	0.007	0.079

Table S7. Summary of the principal component analysis for the coefficient of variation of the song variables.

	Standard deviation	Proportion of variance	Cumulative Proportion
PC1	0.099	0.676	0.676
PC2	0.041	0.118	0.795
PC3	0.037	0.096	0.891

Table S8. Loadings for all variables of the principal component analysis for the coefficient of variation of the song variables.

	PC1	PC2	PC3
Delta frequency <i>fee</i> note	0.612	-0.518	-0.463
Delta latency <i>fee</i> note	0.422	-0.261	0.444
Frequency 5 % <i>fee</i> note	-0.084	-0.081	0.224
Frequency 95 % <i>fee</i> note	-0.001	-0.080	0.204
Peak frequency <i>fee</i> note	-0.083	-0.104	0.241
Delta frequency <i>bee</i> note	0.326	0.333	0.250
Delta latency <i>bee</i> note	0.263	0.627	-0.277
Frequency 5 % <i>bee</i> note	-0.066	-0.117	0.179

Frequency 95 % <i>bee</i> note	-0.037	-0.086	0.208
Peak frequency <i>bee</i> note	-0.042	-0.118	0.199
Song length	0.261	0.114	-0.024
<i>fee</i> note length	0.341	0.075	0.335
<i>bee</i> note length	0.243	0.284	0.224
Glissando ratio	0.088	-0.027	-0.093
Interval ratio	0.004	0.002	0.106

Table S9. F statistic and *p* value of the non-significant models applied in the statistical analysis of the song variables.

Model	F statistic	<i>p</i> value
PC1 CV song variables ~ mean errors spatial learning + mean latency spatial learning	$F_{2,11} = 1.48$	0.27
PC2 CV song variables ~ mean errors spatial learning + mean latency spatial learning	$F_{2,11} = 0.07$	0.93
PC3 CV song variables ~ mean errors spatial learning + mean latency spatial learning	$F_{2,11} = 0.33$	0.73
CV Interval ratio ~ mean errors spatial learning + mean latency spatial learning	$F_{2,11} = 0.19$	0.83
CV glissando ratio ~ mean errors spatial learning + mean latency spatial learning	$F_{2,11} = 1.07$	0.38
PC1 CV song variables ~ mean errors reversal learning + mean latency reversal learning	$F_{2,10} = 0.06$	0.94

PC2 CV song variables ~ mean errors reversal learning + mean latency reversal learning	$F_{2,10} = 1.75$	0.22
PC3 CV song variables ~ mean errors reversal learning + mean latency reversal learning	$F_{2,10} = 0.26$	0.78
PC1 CV song variables ~ visits to the old location	$F_{1,11} = 0.74$	0.41
PC2 CV song variables ~ visits to the old location	$F_{1,11} = 0.32$	0.58
PC3 CV song variables ~ visits to the old location	$F_{1,11} = 0.41$	0.54
CV Interval ratio ~ mean errors reversal learning + mean latency reversal learning	$F_{2,10} = 0.61$	0.56
CV glissando ratio ~ mean errors reversal learning + mean latency reversal learning	$F_{2,10} = 2.19$	0.16
PC1 CV song variables ~ mean errors long-term memory + mean latency long-term memory	$F_{2,10} = 0.16$	0.85
PC2 CV song variables ~ mean errors long-term memory + mean latency long-term memory	$F_{2,10} = 0.72$	0.51
PC3 CV song variables ~ mean errors long-term memory + mean latency long-term memory	$F_{2,10} = 0.92$	0.43
CV Interval ratio ~ mean errors long-term memory + mean latency long-term memory	$F_{2,10} = 1.21$	0.34
CV glissando ratio ~ mean errors long-term memory + mean latency long-term memory	$F_{2,10} = 2.21$	0.16