Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Causes of variation in problem-solving performance in wild Western Australian Magpies (Gymnorhina tibicen dorsalis)

Scout O Vos

scout.vos@research.uwa.edu.au

The University of Western Australia https://orcid.org/0009-0007-0838-4503

Elizabeth M Speechley

The University of Western Australia https://orcid.org/0000-0003-0690-955X

Edward C Galluccio

The University of Western Australia https://orcid.org/0000-0002-9870-2355

Amanda R Ridley

The University of Western Australia https://orcid.org/0000-0001-5886-0992

Research Article

Keywords: problem-solving, innovation, cognition, magpies, individual variation

Posted Date: July 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10197540/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Causes of variation in problem-solving performance in wild Western Australian Magpies (*Gymnorhina tibicen dorsalis)*

Author list and order: Scout O. Vos[1] (0009-0007-0838-4503)**, Elizabeth M. Speechley**[1] (0000-0003-0690-955X)**, Edward C. Galluccio**[1] (0000-0002-9870-2355)**, Amanda R. Ridley**[1] (0000-0001-5886-0992).

Author affiliations: [1]The University of Western Australia, Centre for Evolutionary Biology, School of Biological Science (M092), Crawley, WA, 6009, Australia

Corresponding author: Scout O. Vos (scout.vos@research.uwa.edu.au)

Key words: problem-solving, innovation, cognition, magpies, individual variation

Abstract

Problem-solving ability is considered a reliable measure of innovation in wild animals and may therefore be an important indicator of the ability to adapt to ecological change. Both problem-solving ability and cognitive ability have been suggested as factors that influence the ability of animals to adapt to urban environments, yet whether these two factors are related to each other is unclear. We tested problem-solving ability in free-living Western Australian magpies (*Gymnorhina tibicen dorsalis)* located in an urban area by implementing a six-in-a-row threshold of consecutive solves to constitute successful demonstration of the ability to solve a task beyond chance performance. Individuals were presented with two puzzles (a lid-flip and lid-pierce puzzle) in a randomised order, and we examined factors affecting (a) the time taken to first solve the task and (b) to solve the task six times in a row. We also compared problem-solving ability to associative learning score (representing a measure of cognition). Our results revealed that the factors that affected the time to first solve a puzzle were different to those that affected how long it took to solve a puzzle six times in a row, reinforcing the importance of testing problem-solving performance beyond chance levels. We also found that problem-solving ability was not related to associative learning performance, suggesting that the often-assumed positive link between problem-solving ability and cognition is not supported in this species.

Introduction

In dynamic and rapidly evolving ecosystems, an individual's condition, fitness and survival can be impacted by their ability to respond to novel stimuli or solve new problems within their environment (Lee & Thornton, 2021; Sih et al., 2011). Problem-solving ability, defined as the act of overcoming an obstacle to obtain an incentive (Rowell et al., 2021), is an experimental measure of innovative behaviour and adaptive capacity, and is considered to be increasingly relevant in urban environments where novel resources are more frequently exploited. Urban-adapted species (those that can establish and persist in urban environments over multiple generations) tend to exhibit diverse behavioural adjustments, including shifts in times of behavioural occurrences (e.g., changes in optimal foraging times) and novel foraging strategies to exploit anthropogenic resources (Gaynor et al., 2024; Griffin & Diquelou, 2015; Kark et al., 2007; Lowry et al., 2013; Møller, 2009; Petelle et al., 2023; Sol et al., 2005; Townsend et al., 2025). Accordingly, identifying the mechanisms

underlying innovative behaviours is vital for predicting adaptive behavioural changes to urbanisation and environmental change (Alberti et al., 2017; Sol et al., 2020).

Animal innovations can range broadly from the ability to access novel food types, through to the expression of new behaviours in response to novel stimuli or obstacles (Tebbich et al., 2016). To date, problem-solving ability has been demonstrated in a variety of captive vertebrate taxa including the red panda (*Ailurus fulgens),* brushtail possum (*Trichosurus vulpecula*), and several parrot and corvid species (Benson-Amram et al., 2016; Harris et al., 2023; Lambert et al., 2019; Taylor, 2014). Relatively few studies have examined problem-solving ability in wild populations and thus our understanding of how and why problem-solving abilities vary among individuals remains limited (Morand-Ferron et al., 2016; Roth et al., 2010).

Cognitive abilities are crucial for survival as they facilitate the identification of resources, avoidance of predators, interactions with rivals and mates, and can influence reproductive success (Morand-Ferron, 2017; Szabo et al., 2022; Wooster et al., 2024). Cognition in animals is the mental process that encapsulates the way an individual processes, retains and uses environmental data to shape behavioural responses (Shettleworth, 2001). Cognitive ecology provides a framework to link cognition to ecological pressures and examine how it shapes fitness and adaptability (Morand-Ferron, 2017; Shettleworth, 2009). However, it remains contentious as to whether problem-solving ability reflects an independent domain of cognition or if it is interlinked with other cognitive measures. Problem-solving may involve multiple processes including associative learning, inhibitory control and persistence (Rowell et al., 2021). Performance varies across taxa, intraspecific populations and between individuals (Rowell et al., 2021). Capturing this variation can be methodologically challenging, as many studies are often unable to obtain large enough sample sizes in wild populations, and many existing studies rely on single exposure to problem-solving tasks (see Benson-Amram et al. (2016); Cole and Quinn (2011); Morand-Ferron and Quinn (2011); Stanton et al. (2024) for examples of single exposure to problem-solving tasks), rather than measuring the above chance frequency of solving a task. These methodological constraints limit generalisation of findings and hinder the identification of intra- and inter-specific variation in problem-solving ability. Additionally recent research has found that problem-solving ability is unrelated to other cognitive measures including associative and reversal learning, suggesting that these cognitive tests may capture independent processes or context-dependent behaviours (Audet et al., 2024). This therefore raises vital questions about how problem-solving measures should be implemented, measured and interpreted and whether current methods of cognitive measures can aptly explain variation in problem-solving ability.

Western Australian magpies (*Gymnorhina tibicen dorsalis*) (hereafter; magpies) are a subspecies of the Australian magpie*,* and provide an ideal system to explore the relationship between problem-solving ability and one's cognitive ability. Magpies are commonly observed in urban environments, suggesting that they are able to adapt well to novel conditions (Blair, 1996), and cognitive tests have been previously successfully conducted on this species (Ashton et al., 2018; Blackburn et al., 2022; Speechley et al., 2024), thus allowing for direct comparison between cognition and problem-solving ability at the individual level. Here we tested whether magpies exhibit a consistent ability to solve simple problem-solving tasks using a repeated-measures success

criterion. We examined which factors best explain variance in performance, and whether cognitive performance (proxied by associative learning score) is related to problem-solving ability.

Methodology

Study Species and Population

The Western Australian magpie is a large (250 - 370 g) group-living passerine bird (Speechley et al., 2024). Group sizes typically vary between 2 - 12 adults (median group size = 5). The study population (15 groups) was located within urban parklands across two residential suburbs in the greater Perth metropolitan area (Western Australia) - Guildford (31.8994° S, 115.9717° E) and Crawley (31.9752° S, 115.8213° E). Individuals are well-habituated to researcher presence and readily engage with cognitive tasks (Ashton et al., 2018; Speechley et al., 2024). The sex of adult magpies (>2 years old) is readily determined by sex-specific plumage colour differences (Ashton et al., 2018). Most individuals are ringed making them identifiable by their unique ring-colour combinations or differences in plumage and/or scarring patterns.

Problem-solving tests

All testing was conducted in the morning (6 - 11 am) when the birds are most active (Edwards et al., 2015). At least one male and one female adult per group was tested, however two additional individuals beyond this sample size were included to increase statistical power (n = 34, 17 males, 17 females from 15 different groups). Target individuals for testing were chosen at random from the list of birds present in each social group prior to in-situ testing to mitigate observer bias and avoid choosing only the bolder individuals.

Testing trials began with a training phase. This initial phase allowed individuals to habituate to the presence of the array and to learn that a food reward (small amount of mozzarella cheese, < 1 g) was associated with it (Boogert et al., 2011; Mirville et al., 2016). In the post-training phase, there were two arrays presented in randomised order to each individual. The first array was the lid-flip task (Figure 1, Panel A and B). To solve the puzzle, individuals must flip a lid off a clear cylinder to access a food reward visible inside. The second was the lid-pierce task (Figure 1, Panel C and D). Individuals must pierce through a barrier of butcher's paper at the top of a container to remove the food reward inside, which is visible from the side of the container. Both task bases were made of wood, 15 cm wide and 20 cm long, with a thickness of 1 cm. Cylinders were made of firm transparent plastic (height = 5 cm, diameter = 4 cm, depth = 2 cm), affixed to the wooden base using a non-toxic and odourless adhesive. For the lid-flip task, the detachable lid had a 5.5 cm total diameter which encompassed both the lid, and the rim (extending 1cm from the lid) attached. For the lid-pierce task, butchers paper was placed over the rim of the cylinder and affixed using non-toxic and odourless adhesive patches. It was secured by a thickened elastic band to ensure birds could not remove the paper from the sides and requires penetrative access through the top of the cylinder. The array designs were adopted from the standard set of problem-solving arrays for passerine bird species presented by Audet et al., (2024).



Figure 1. Top and side views of the two problem-solving arrays used in the study: the lid-flip (Panel a, b) and lid-pierce (Panel c, d) puzzles.

Prior to commencing trials, for both puzzles, the array was baited out of sight of the focal individual. The array was then placed 5 m away from the testing individual. Once the array was placed down, the observer moved backwards (to a distance of 5 m) to allow the focal individual to attempt to solve the array with no observer interference. An individual was considered to have participated if it approached within 30cm of the array at least once. The array is repeatedly presented to the individual until the criteria of six consecutive solves was reached or the individual reached the cap of 100 failed trials.

The tested individual was kept in social isolation (approximately 10 m) from conspecifics across all attempts to avoid any social learning (*sensu* (Ashton et al., 2018)). This social isolation was achieved without interference: the focal individual was not moved or contained to a particular place. Rather, the array was only presented to the focal individual when it was foraging at least 10 m away from other group members under natural conditions. All individuals were given a 5-minute time per trial for completion. If the individual failed to solve the puzzle within this timeframe, the attempt was recorded as a failure, and the puzzle was re-presented until it was successfully solved. The number of presentations to solve the puzzle to the six-in-a-row threshold was recorded. Individuals that had not reached the criteria for consistent solving ability within 100 trials were considered to

have failed the puzzle task. Testing continued across consecutive days if an individual did not pass on the first day (n = 8, individuals requiring testing over multiple days) (Soravia et al., 2023).

During the 5-minute trials, we recorded time taken to approach the puzzle (seconds) as a proxy for neophobia, defined as the time from the initial deployment of the puzzle to the moment the individual was within 30 cm of the array; the time spent interacting with the puzzle (seconds), defined as the time after approach but prior to attempting to solve the puzzle (e.g., circling the array before contacting it); time taken to solve the puzzle (s) and total time (from placement of the puzzle to final solve time (s)).

Successful completion of both arrays was defined as successful removal of the reward from within the array. To pass, each puzzle must be solved six consecutive times - this criterion was chosen to ensure that success was not due to chance, adapted from previous studies that considered a learned association to be established after six consecutive successful trials (representing a binomial p-value of 0.016, significantly deviating from random binomial probability (Soravia et al., 2023)). Problem-solving performance was quantified by the speed with which it solved the array and by the number of attempts taken to solve the array (as a proxy for speed and efficiency of information processing and the value of trial-and-error based learning respectively).

Explanatory Factors

We quantified environmental and intrinsic variables that may have influenced problem-solving ability. Wind condition was recorded and categorised as high (>25 km/hour), medium (>10 km/hour - <25 km/hour), or low (0 km/hour - <10 km/hour) from the Bureau of Meteorology, (2025). We also recorded cloud cover (none, partial (>0% - < 50% cover)), or complete (>50% cover), from the Bureau of Meteorology (2025)). Temperature (°C, from Bureau of Meteorology (2025)) was also recorded during the testing period, as it is known to impact cognitive performance within magpie populations (Blackburn et al., 2022). For the eight individuals who were tested across multiple days, the average of environmental variables (including wind, cloud cover and temperature) was used.

Additional variables that were recorded included group size on the day of testing, and maximum group size (which included individuals absent during testing but normally present in the group, to represent the typical group size of the individual being tested). The weight (grams) of each individual was also recorded. Weight was collected by encouraging individuals to hop onto a top-pan scale (Ohaus Challenger, 1000 g ± 1 g) for a small food reward (mozzarella cheese, <1 g). Individuals were weighed within 5 days post-testing. Age was taken from the existing long-term database held at the University of Western Australia (UWA), encompassing information on all individuals within the study population since 2015.

Associative learning

Associative learning data, to compare with problem-solving performance, was collected between December-February in 2024 and June of 2025, using the methods described in Galluccio et al. (2026). Associative learning ability has been previously attributed to a general intelligence factor within magpies (Ashton et al., 2018;

Blackburn et al., 2024) as it is a domain-general trait that highlights whether animals can make an association between predictable cues, such as associating alarm calls from group members with the presence of danger (Morand-Ferron, 2017; Morand-Ferron et al., 2016; Speechley et al., 2024), thereby making it highly likely to be ecologically relevant (Speechley et al., 2024).Therefore, associative learning scores may be used as a proxy for determining individual cognition in this species.

The associative learning array was mounted on a wooden block (31 cm × 9 cm × 4 cm) with two wells of equidistant measures (3.2 cm diameter × 1.5 cm depth, 3.5 cm apart). PVC lids covered the wells and were fastened by elastic bands which were attached to the sides of the mount, allowing for lids to move when they are pecked. The two lids were painted with two different shades of the same colour to limit any pre-existing colour biases that may influence choice (Speechley et al., 2024). Both wells were rubbed with cheese prior to testing to reduce olfactory cues that could be used to locate the rewarded well. One colour shade was chosen at random to be the rewarded well, and the food reward would only be placed under that shade (the food reward is mozzarella cheese, <1 g). Across all trials the array was baited out of sight of the individual and deployed within 5 m of the focal individual. After the array was placed down the observer stepped away from the task (~5 m) to mitigate any observer interference. In the first approach the individual is allowed to investigate both wells until the reward is identified. In all subsequent trials only one well was allowed to be selected before the array was removed to reinforce a positive choice. Individuals that selected the correct well were allowed to consume the food reward. If the incorrect well was selected, the array was removed before they could locate the reward, thus imposing a cost on the incorrect selection. Associative learning was deemed to have occurred when an individual selected the rewarded colour shade in 10 out of 12 consecutive trials, which represented a significant deviation from random binomial probability (Ashton et al., 2018; Soravia et al., 2023). All individuals were tested in isolation to prevent any social learning (with conspecifics ~10 m away from the individual being tested) (Speechley et al., 2024).



Figure 2. Associative learning task showing colour cue pairing. Image credit: E. Galluccio.

Statistical Analysis

Statistical analysis was conducted in R version 4.5.1 (R Core Team, 2025). To analyse key drivers of variation in problem solving ability, generalised mixed linear models (GLMMs) were fitted using the *GlmmTMB* package (McGillycuddy et al., 2025). Solve time was used as the primary proxy for problem-solving in all models, with shorter solve time reflecting faster information processing during array completion. Model selection was conducted by comparing models fitted with predictor terms to a basic model which only included the response variable (solve time (s)) and random terms (*sensu* Harrison et al. (2018)). All continuous variables were standardised and centralised around zero prior to analysis to improve comparability of effect sizes between models. Prior to model selection, data was examined for violations of model assumptions including assessment of multicollinearity between predictors (using variance inflation factors, VIF) and residual distributions. Predictor variables where VIF > 3 were not included in the same model. Models were ranked based on the lowest Akaike's Information Criterion score (AICc, corrected for sample sizes <500) . Predictor terms with 95% confidence intervals (CIs) that did not intersect zero were included in the top model set. Candidate predictor variables were fitted in biologically informed model sets, rather than all possible combinations to reduce overfitting. Group and individual ID were included as random terms to account for non-independence of datapoints.

Trial 1 solve time

A subset of data from only the first solve (trial 1) from both the lid-flip and lid-pierce puzzles was analysed (n = 68) to determine the factors affecting time to first solve a puzzle. Individual ID was not included as a random term in the analysis as there was only a single datapoint per individual, per puzzle. A log transformation was applied to the response variable (solve time (s)), to achieve normality. As this dataset met assumptions of normality following the transformation and did not use repeated measures, linear mixed models (LMMs) were used for this analysis using the *lme4* package (Bates et al., 2015).

Time to solve the puzzle six times in a row

The solve time taken across all six trial was analysed under three different datasets. The first dataset included solve times for both arrays. This was to assess whether solve time significantly differed between the two puzzle arrays. The second dataset included only trials done on the lid-flip puzzle test and the third dataset included only trials from the lid-pierce puzzle array. The second and third datasets were included to understand if the variables influencing solve times differed between the two arrays. Solve time was positively skewed for all three datasets and was transformed using a log transformation. The dataset was modelled using a generalised mixed linear model (GLMMs) with gamma error distribution with a log link function. Basic models including only the response term (solve time (s)) and all random terms (group and individual ID) were first fitted to assess dispersion. Both group and individual ID were used as random effects in all models.

Assessing learning trajectories

GLMM models with a binomial 'logit' link function were implemented using the glmer function from the lme4 package to assess if there was variation in learning trajectories both by puzzle type and between individuals as outlined by Harrington and Lambert (2025). Binary attempts to solve outcome (yes puzzle was solved by attempt, or no puzzle was not solved by attempt) was modelled as a function of the number of attempts taken to solve the puzzle, puzzle type (lid-flip or lid-pierce) and the interaction between the terms. Individual and group ID were included as random terms in all models to account for repeated measures and non-independence between groups.

A full model with all terms with interactions was compared to a reduced model (all terms, no interaction) and a null model using Akaike Information Criterion corrected for small sample sizes (AICc) (n = 1706, from 34 individuals from 14 different living groups across both lid-flip and lid-pierce puzzles). Population- and individual-level learning curves were plotted from probabilities extracted from above models.

Cognition & Problem-Solving ability

Associative learning score data was compared to problem-solving scores to ascertain whether there was a relationship between the two. Two proxies for problem-solving were used: (a) solve time (s) and (b) the number of attempts taken to solve the puzzle for each trial, which is a proxy for determining reliance on trial-and-error based solving of novel stimuli. Both solve time and attempts to solve were analysed using a gamma GLMM. For individuals that had two associative learning scores the average of the scores was used. Group and individual identity were included in all models as random terms. Analyses were restricted to individuals with both

problem-solving and associative learning scores (n = 162 scores from the three measures of problem-solving ability against associative learning score from 27 individuals from 14 groups). Two basic models were fitted with the two different problem-solving measures (attempts to solve the array (hereafter; attempts) and solve time (s)) as the response term to understand whether associative learning had a significant influence on problem-solving ability.

Results

Factors affecting initial problem-solving ability (time to first solve of puzzle)

Lid-flip puzzles were solved faster than lid-pierce puzzles based on the first solve time (Table 1). The number of attempts taken to solve the array was the strongest predictor of solve time (Table 1, Figure 3).

Problem-solving ability across all six trials

Solve time across all six trials was best explained by puzzle type, maximum group size and the time spent interacting with the puzzle (Table 2). Individuals who spent longer interacting with the puzzle also had longer solve times (Table 2, Figure 5). Lid-flip puzzles were solved faster and with a higher degree of consistency than lid-pierce puzzles (Figure 5). Most solve times were below 20 seconds for lid-flip puzzles and showed little variation across repeated attempts (mean = 3.35 ± 2.03 s, range = 12 s). Contrastingly, lid-pierce puzzles showed greater variability with several solve times exceeding 50 seconds (mean = 5.01 ± 2.5 s, range = 17s) see SM Figure 1). Larger group sizes were associated with longer solve times, (β = 0.221, 95% CI = 0.024-0.418; Table 2).

Learning trajectories across all six trials varied by individual (Figures 6-9) and by puzzle type (Figures 6-9). There was a difference in the number of individuals (from 34 individuals sampled) that exhibited learning across both arrays. 8 individuals showed they learned to solve the flip array, as opposed to the pierce puzzle which had 22 individuals show learning across all six trials (Table 3).

Table 1. Top model set for the factors influencing time taken to solve the puzzle for the first time (n = 68, 34 individuals from 15 groups). Group and individual ID were included as random terms. AICc and ΔAICc are provided for all models, with parameter estimates, standard error (SE) and 95% confidence intervals (CIs) reported for the best supporting model. Attempts = attempts taken to solve the array. The top candidate model with the lowest AICc in bold.

Terms	AICc	ΔAICc
Attempts	**177. 251**	**0**
Puzzle type + attempts	179.564	2.313

Puzzle type	219.338	42.087	
Puzzle type + Sex	221.274	44.023	
Puzzle type + Group size maximum	221.664	44.413	
Basic model (no predictors)	222.555	45.304	
Puzzle type + Sex + Group size maximum	223.764	46.513	

Parameter	Estimate	SE	CI
Attempts	0.388	0.042	0.303, 0.471

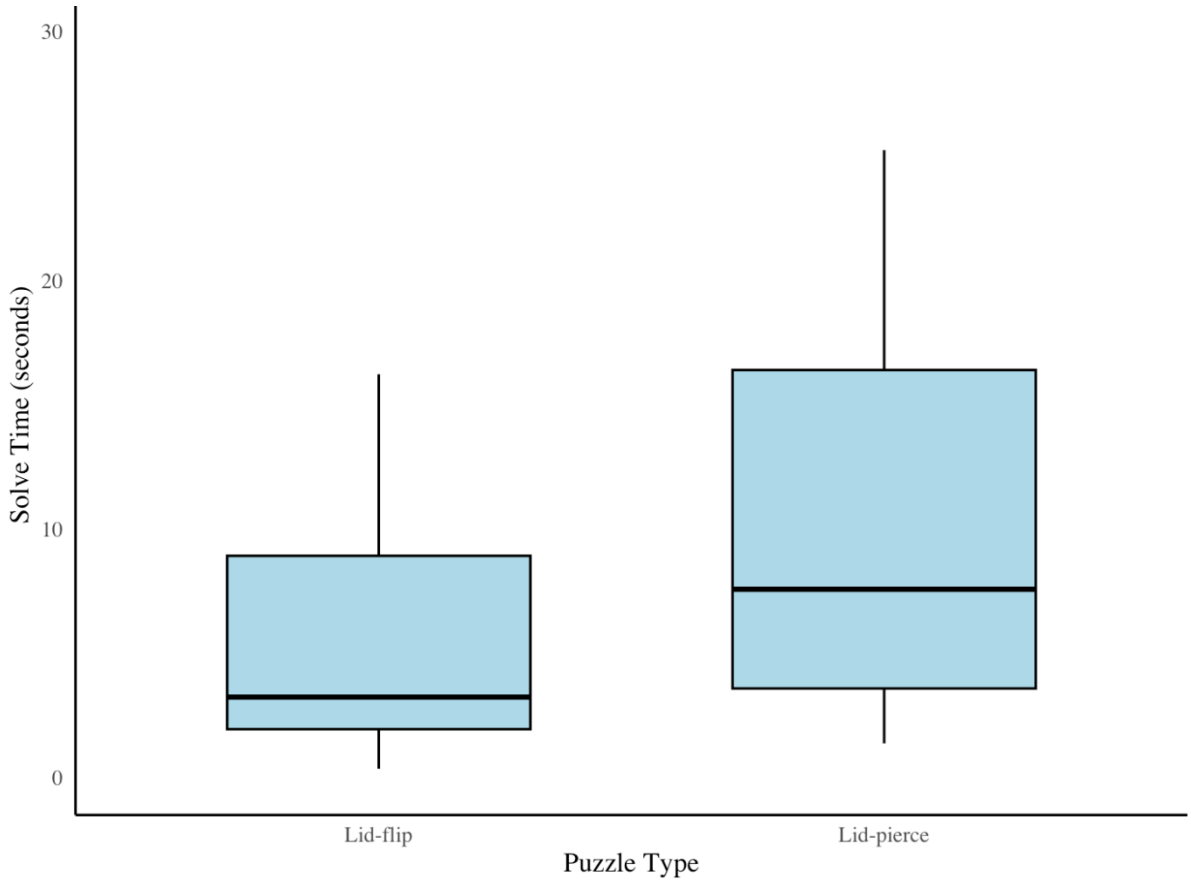


Figure 3. Predicted solve time (seconds) by puzzle type (lid-flip or lid-pierce) for the first successfully solved trial (trial 1) (n = 68 observations from 34 individuals across 15 groups).

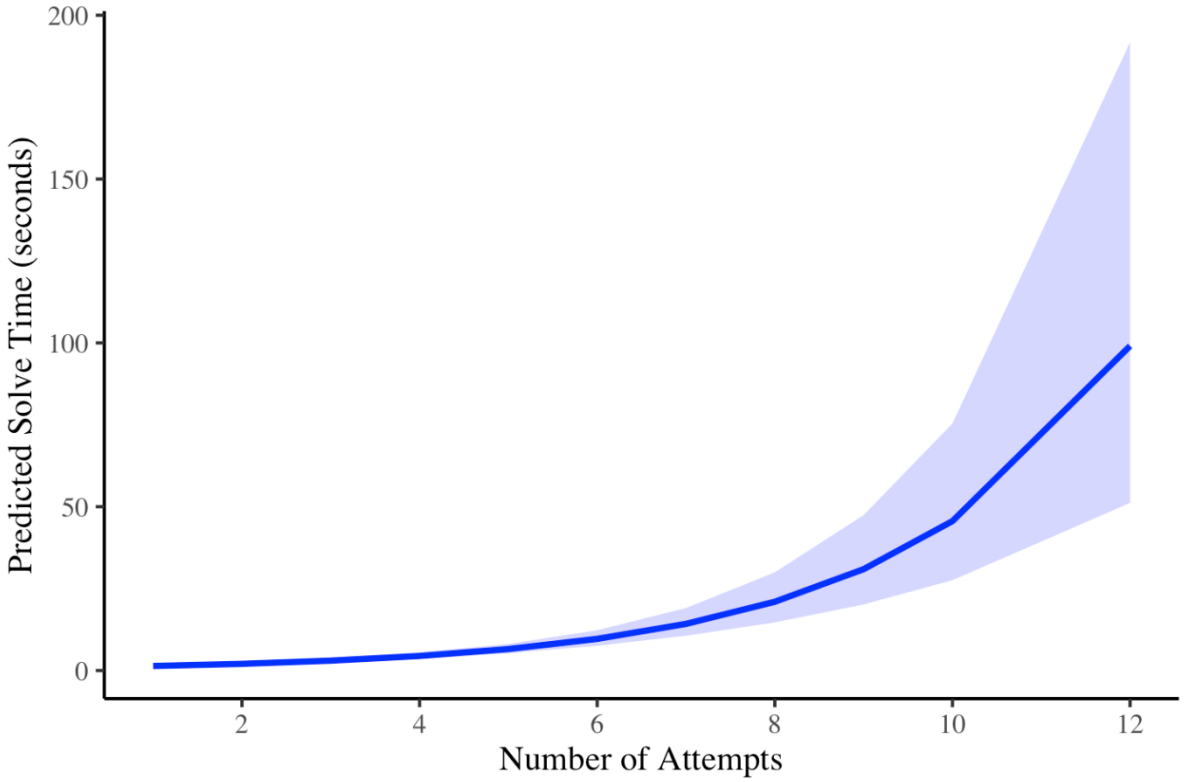


Figure 4. The relationship between the number of attempts and time to successfully solve the puzzle in the first trial (n = 68, 34 individuals from 15 groups). Figure generated from the data predictions presented in Table 1. The bold line represents the line of best fit; the shaded areas represent 95% confidence intervals.

Table 2. GLMM Top model set for the factors influencing solve time across all six trials analysed using a GLMM with a gamma distribution (n = 408 from 34 individuals across 15 groups). Group and individual ID were included as random terms in all models. Corrected AICc and ΔAICc are provided for all models, with parameter estimates, standard error (SE) and 95% confidence intervals (CIs) reported for the best supporting model. Best fit model with lowest AICc value in bold.

Predictor terms	AICc	ΔAICc	
Maximum group size + Time spent interacting + Puzzle Type	**2669.134**	**0**	
Time Interacting + Puzzle Type	2671.608	2.473	
Basic (no predictors)	2753.683	84.549	
Parameter	Estimate	SE	CI
Group size maximum	0.221	0.100	0.024, 0.418
Time spent interacting	0.055	0.021	0.015, 0.096
Puzzle type	0.994	0.106	0.785, 1.202

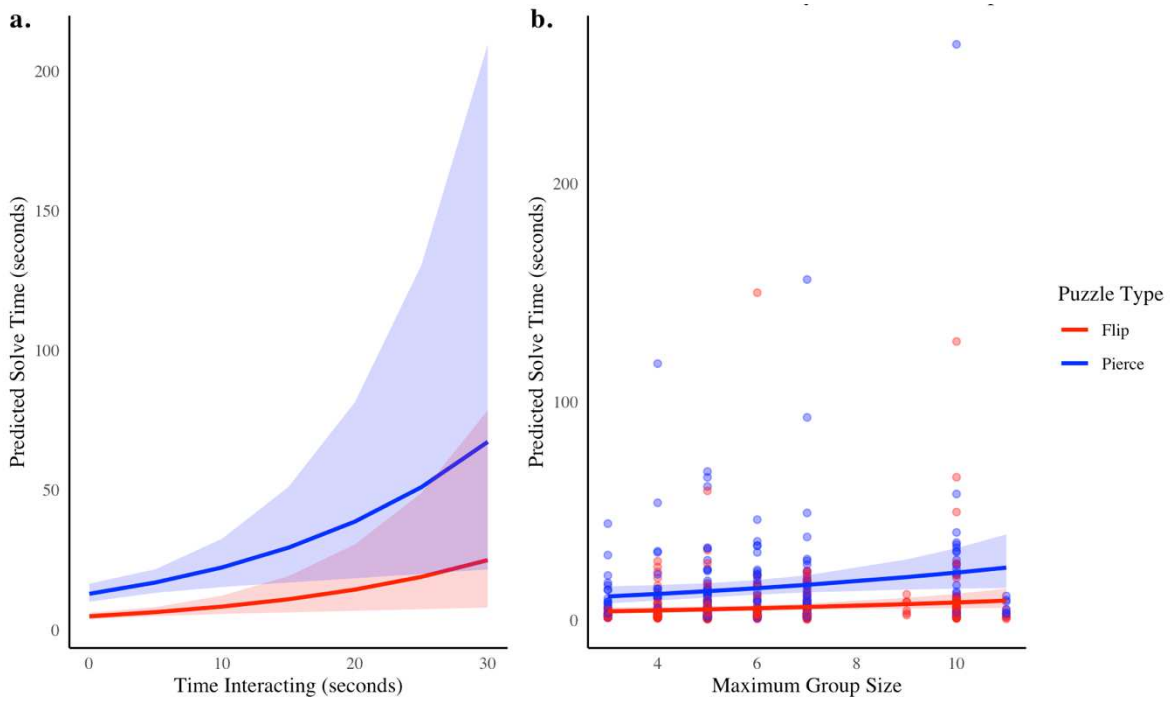


Figure 5. The relationship between solve time across all six trials and (a) the time an individual spent interacting with the array, (b) maximum group size and puzzle type (red = lid-flip, blue = lid-pierce). Figures are generated from the data predictions presented in Table 2. The solid line represents model predictions; the shaded areas represent 95% confidence intervals.

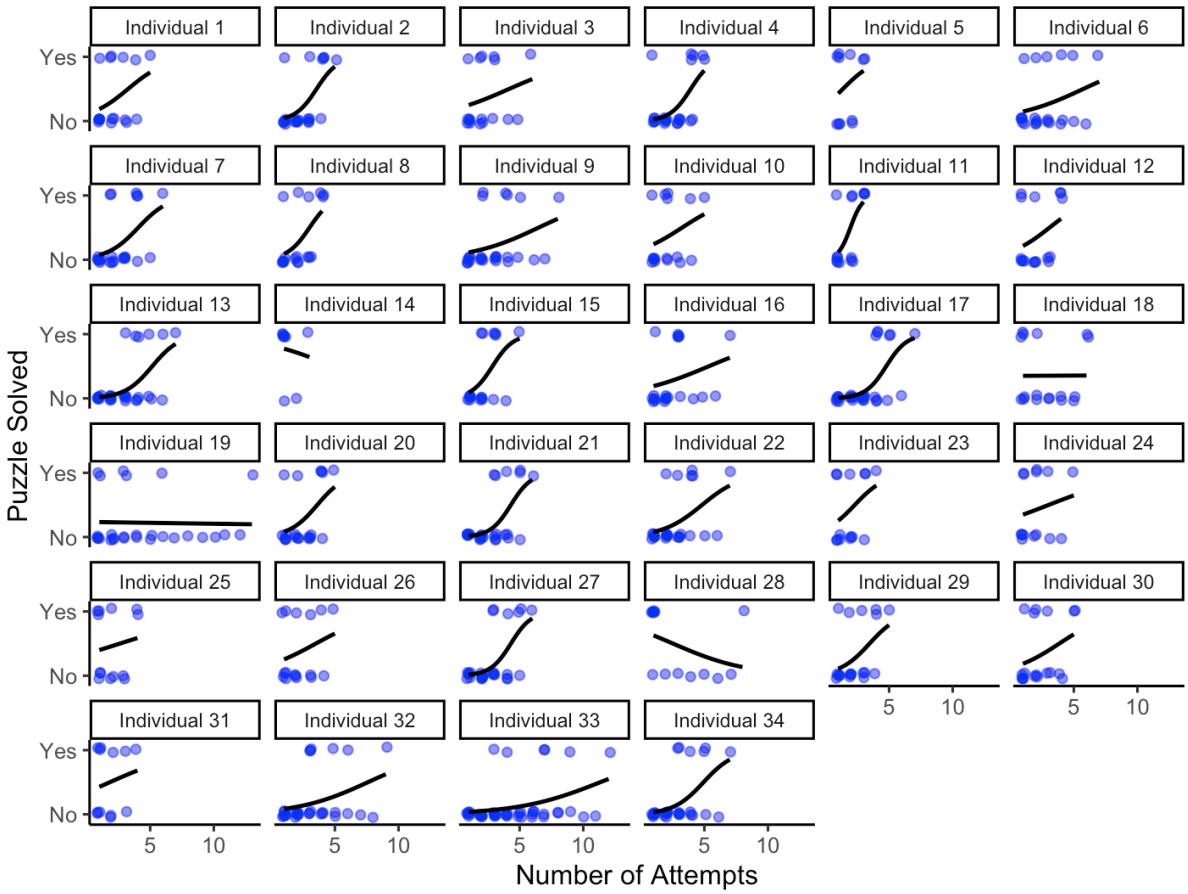


Figure 6. Learning curve trajectories showing the success or failure (yes/no) of the attempts taken to solve the lid-flip puzzle against the number of attempts taken to reach successful completion of the puzzle for all six trials (n = 683 across six trials from 34 individuals across 15 living groups). The black line represents an individual-specific learning curve, with each panel indicating a separate individual.

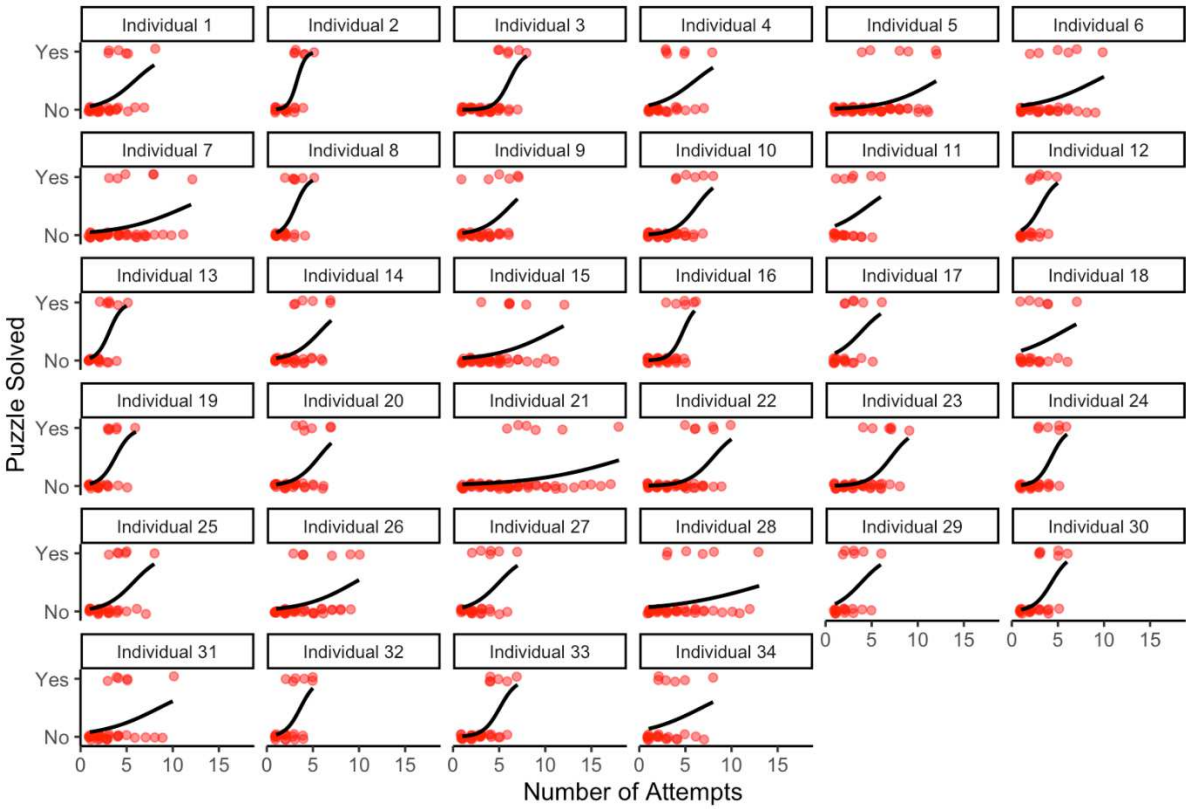


Figure 7. Learning curve trajectories showing the success or failure (yes/no) of the attempts taken to solve the lid-pierce puzzle against the number of attempts taken to reach successful completion of the puzzle for all six trials. (n = 1023 across six trials from 34 individuals across 15 living groups). The black line represents an individual-specific learning curve, with each panel indicating a separate individual.

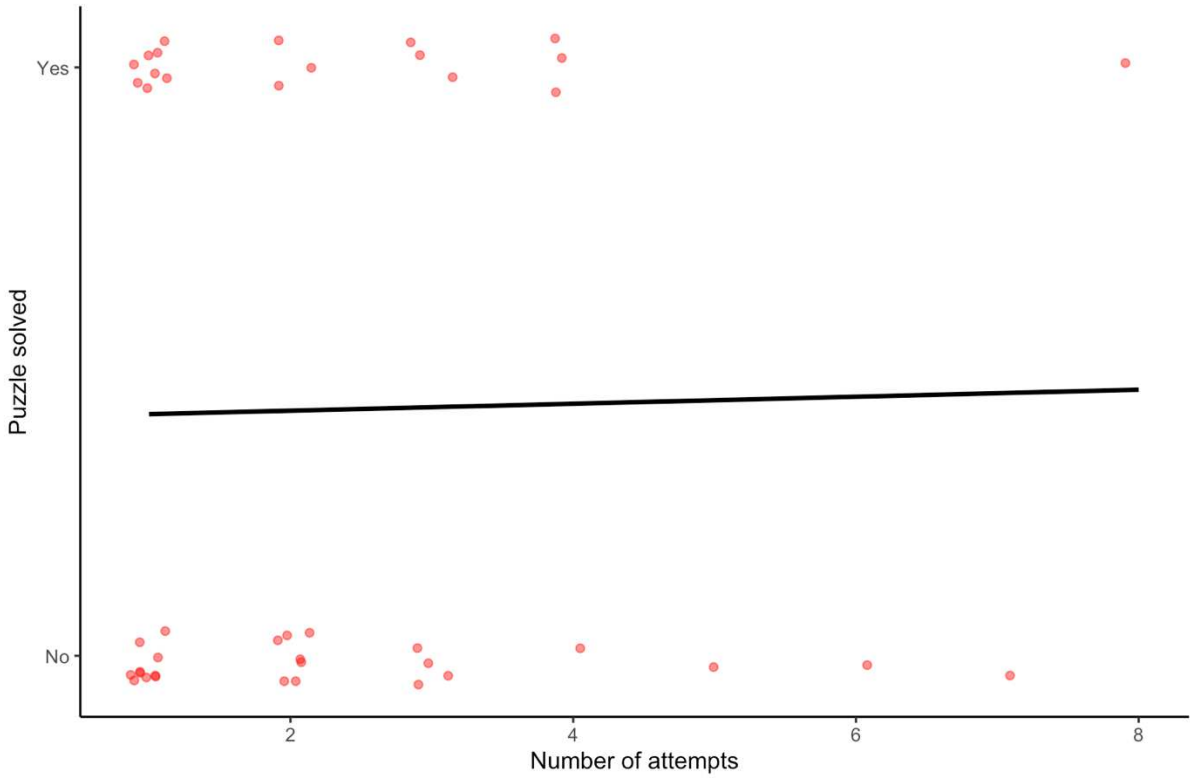


Figure 8. Representative learning curve trajectories showing the success or failure (yes/no) of the attempts taken to solve the lid-flip puzzle against the number of attempts taken to reach successful completion of the puzzle for all six trials from three representative individuals three representative individuals (the individuals that scored the lowest number of attempts, the median number of attempts and the highest number of attempts recorded for the array) (n = 43 across six trials from 3 representative individuals across 3 living groups). The black line represents an individual-specific learning curve, with each panel indicating a separate individual.

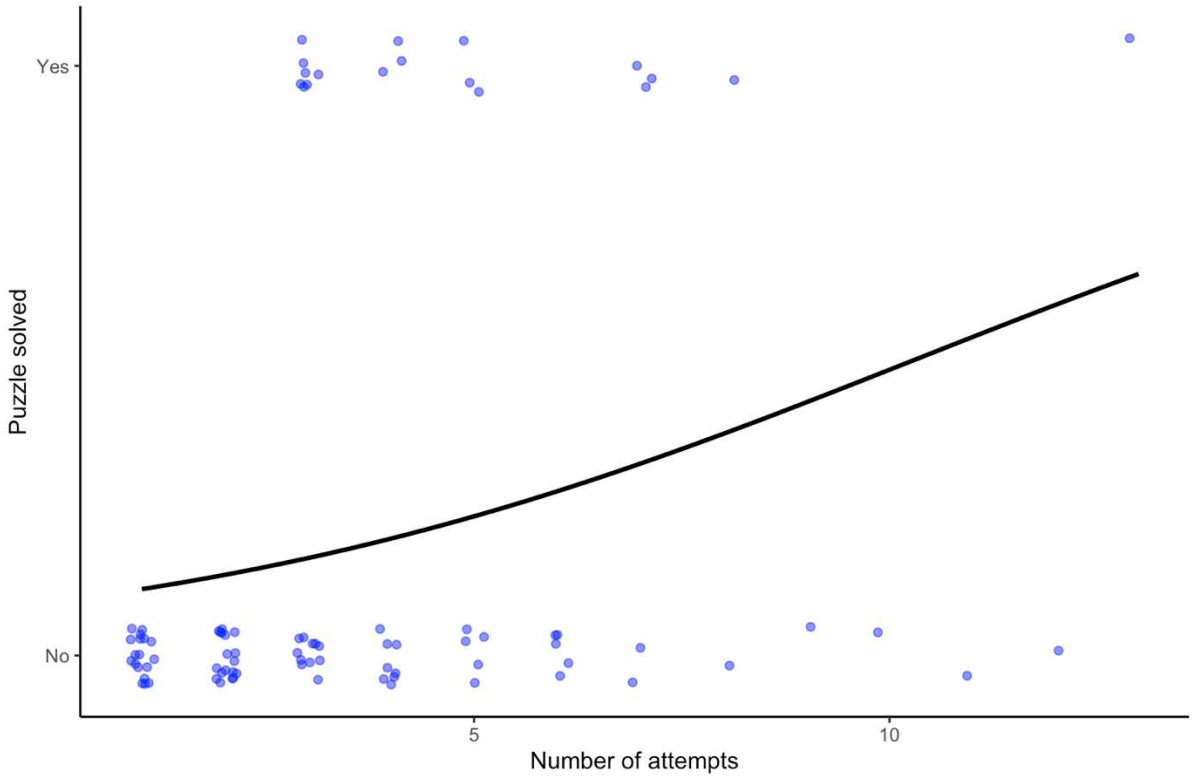


Figure 9. Representative learning curve trajectories showing the success or failure (yes/no) of the attempts taken to solve the lid-pierce puzzle against the number of attempts taken to reach successful completion of the puzzle for all six trials from three representative individuals (who scored the lowest number of attempts, the median number of attempts and the highest number of attempts recorded for the array) (n = 90 across six trials from 3 individuals across 3 living groups). The black line represents an individual-specific learning curve, with each panel indicating a separate individual.

Table 3. Summary of individual learning outcomes for both the lid-flip and lid-pierce puzzle array, showing individuals that did learn vs did not learn and the total number of individuals sampled (n = 1023 observations across six trials from 34 individuals across 15 living groups). Individuals were classified as learners if the slope of the binomial regression was positive and p = < 0.05

Puzzle Type	Total number of learners	Total number of non-learners	Total number of individuals
Lid-flip	8	26	34
Lid-pierce	22	12	34

Associative learning score vs problem-solving ability

For the subset of individuals who had both associative learning scores and problem-solving scores, there was no significant relationship between these two terms (Table 4).

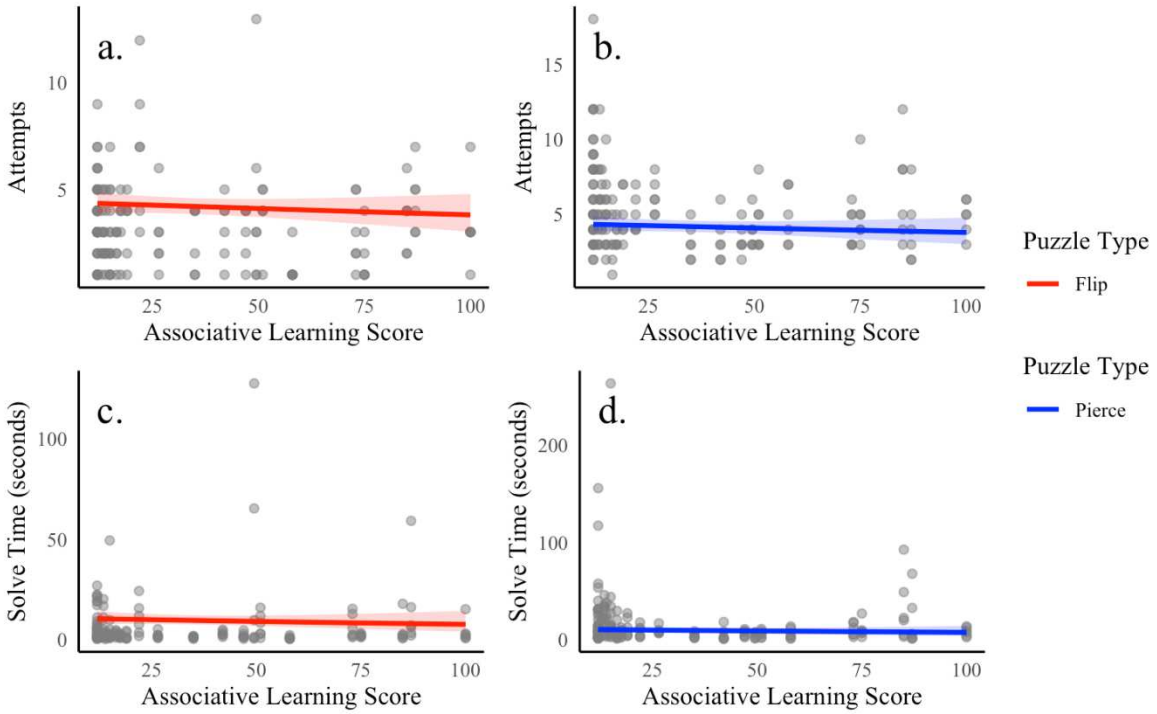


Figure 10. Scatterplot of the relationship between associative learning score and problem-solving ability. (a) & (b) represent the number of attempts to solve the puzzle, for the two different puzzle types, (c) & (d) represent solve time to represent speed of information processing. Figures are generated from the model predictions presented in Table 4 (n = 324 from 27 individuals across 14 different groups). Attempts represent the number of attempts taken to solve the array.

Table. 4. GLMM top model set for observing the relationship between associative learning score and problem-solving ability measures (attempts to solve (the array) and solve time across all trials)) (n = 324 from 27 individuals across 14 groups). Corrected AICc and ΔAICc are provided for all models, with parameter estimates, standard error (SE) and 95% confidence intervals (CIs) reported for the best supporting model. Best fit model with lowest AICc value in bold.

Model	AICc	ΔAICc
Basic model (no predictors)	**1402.961**	**0.000**
Associative Learning x attempts to solve	1404.143	1.182
Basic model (no predictors)	**2157.915**	**0.000**

Associative Learning x solve time	2159.349	1.434	
Parameter	Estimate	SE	CI
Associative Learning x attempts to solve	-0.002	0.002	-0.005, 0.002
Associative Learning x solve time (s)	-0.004	0.004	-0.012, 0.005

Discussion

We found that wild Western Australian magpies can solve simple problem-solving tasks, consistent with findings on other bird species (Audet et al., 2024). In addition, we found that they were able to solve the task multiple times, not just once (with one solve not allowing determination of whether solve was by chance for some individuals). Thus, we were able to show problem-solving ability above chance levels for the first time in a wild animal (see Cole et al. (2012); Cole and Quinn (2011); Petelle et al. (2023); Stanton et al. (2024) for cases of problem-solving in wild animals, but without a consecutive threshold of success of problem-solving per animal). In addition, we were able to determine causes of individual variation in problem-solving ability which differed between the initial solve (trial 1 only) and six consecutive solves (trials 1 through 6). We also found that an individual's associative learning score (considered a proxy for general cognitive ability, Ashton et al., 2018, Blackburn et al., 2024) was not a significant predictor of any measure of problem-solving ability.

There were differences in predictors of variation in solve time between the initial solve (trial 1) and the six consecutively solved trials. Trial 1 solve time was best predicted by the number of attempts to solve the puzzle, with an increase in attempts resulting in an increased solve time. The ability to successfully solve the puzzle six times consecutively was best predicted by group size, the time spent interacting with the array and puzzle type. The difference in predictors across single-solve and consecutive-solve thresholds suggests that single solve measures of problem-solving ability may not capture all factors affecting problem-solving ability. The single trial threshold predictors (attempts taken to solve) likely reflect exploratory components of problem-solving, with the repeated exposure to the array likely capturing broader influences including social context, persistence and learning efficiency. This was reflected in findings from the learning curve analysis which showed that despite the lid-flip puzzle being solved faster overall and requiring fewer attempts to solve the puzzle, fewer individuals exhibited significant learning trajectories across repeated trials relative to the lid-pierce puzzle. The addition of the six consecutive solves threshold demonstrates learned task association (Soravia et al., 2023) and provides a method for identifying multiple drivers of variation in performance. This is corroborated by findings from Audet et al. (2024) which suggest that problem-solving cannot be reduced to a single cognitive domain but may instead emerge as the result of interactions between factors such as experience, learning capabilities and environmental context.

The absence of a significant relationship between associative learning and problem-solving ability suggests that problem-solving ability in Western Australian magpies may rely on a broader set of cognitive domains rather than associative learning alone despite previous studies suggesting a relationship (Auersperg et al., 2012;

Benson-Amram et al., 2016; Griffin & Guez, 2014). This is elucidated further through the differing effects of group size across cognitive measures. Whilst previous studies have reported positive correlations between group size and cognitive performance in this species (Ashton et al., 2018; Blackburn et al., 2022; Speechley et al., 2024), larger group sizes in this study were associated with longer solve times across repeated trials. Increased group size may also elevate vigilance requirements owing to the need to observe not only threats and predators but also conspecifics, especially in open or urban environments where human disturbances are frequent (Griffin & Guez, 2015). Problem-solving tasks may therefore be more susceptible to attentional diversion and social interference as opposed to associative learning tasks, alluding to the notion that problem-solving may instead involve interactions among multiple cognitive domains including associative learning and inhibitory control, behavioural flexibility and memory, rather than having a single associative predictor (Morand-Ferron, 2017; Shettleworth, 2001). Although associative learning had been proposed as a baseline mechanism underlying stimulus-reward associations (Dickinson, 2012) and has been linked to problem-solving in species such as great tits (*Parsus major*) (Cauchard et al., 2024), other studies have failed to detect a relationship between problem-solving and cognition (Audet et al., 2024). As only simple tasks were used in this study, more difficult tasks may help clarify cognition-problem-solving relationships and identify the conditions under which performance becomes limited.

Differences in the time taken to solve the lid-flip and lid-pierce puzzles likely reflect variation in how each array aligned with species-typical foraging strategies and vigilance behaviours. Birds typically avoid pressing their faces directly against objects whilst foraging, as doing so can reduce both the visual field and hearing whilst heightening vigilance behaviours, often resulting in task abandonment (Martin, 2007; Tyrrell & Fernández-Juricic, 2017). Occlusion effects such as these are consistent with findings in parrots (Psittacidae) which showed increased vigilance and reduced interaction time with food sources when novel stimuli were present (Mettke-Hofmann et al., 2002). The lid-pierce puzzle, which required individuals to place their faces close to an opaque barrier, may therefore have conflicted more with natural foraging tendencies then the lid-flip puzzle. Aligning tasks with species-specific and ecologically relevant behaviours is important, as task design can influence performance independently of the cognitive trait being measured.

Conclusion

This study provides insight into the causes of individual-level variation in problem-solving ability in a wild animal. We show that solving once, compared to a threshold of six consecutive solves, have different predictors, indicating that single-solve thresholds may both overestimate performance by not ensuring above-chance frequency of success and underestimate the complexity of factors influencing performance. Lastly, we found that problem-solving ability was not correlated with cognition, suggesting that it may emerge from interactions among social, intrinsic and environmental factors rather than a single cognitive measure.

Acknowledgements

This project was fully funded through the University of Western Australia School of Biological Sciences. We would also like to thank Meredith Thayer who volunteered her time to assist with fieldwork and data collection.

Statement of Ethics

This study was carried out under an approved ethics license from the University of Western Australia Animal Ethics Committee (ethics approval number: 2024/ET000171). All applicable international, national, and institutional guidelines for the care and use of animals was followed.

Competing interests

The authors declare no conflicts of interest.

References

Alberti, M., Correa, C., Marzluff, J. M., Hendry, A. P., Palkovacs, E. P., Gotanda, K. M., Hunt, V. M., Apgar, T. M., & Zhou, Y. (2017). Global urban signatures of phenotypic change in animal and plant populations. *Proceedings of the National Academy of Sciences*, *114*(34), 8951-8956. https://doi.org/10.1073/pnas.1606034114

Ashton, B. J., Ridley, A. R., Edwards, E. K., & Thornton, A. (2018). Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature*, *554*(7692), 364-367. https://doi.org/10.1038/nature25503

Audet, J.-N., Couture, M., Lefebvre, L., & Jarvis, E. D. (2024). Problem-solving skills are predicted by technical innovations in the wild and brain size in passerines. *Nature Ecology & Evolution*, *8*(4), 806-816. https://doi.org/10.1038/s41559-024-02342-7

Auersperg, A. M., Gajdon, G. K., & von Bayern, A. M. (2012). A new approach to comparing problem solving, flexibility and innovation. *Communicative & Integrative Biology*, *5*(2), 140-145. https://doi.org/10.4161/cib.18787

Australian Bureau of Meteorology. (2025). *Climate data* BoM. https://www.bom.gov.au/climate/data/

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1 - 48. https://doi.org/10.18637/jss.v067.i01

Benson-Amram, S., Dantzer, B., Stricker, G., Swanson, E. M., & Holekamp, K. E. (2016). Brain size predicts problem-solving ability in mammalian carnivores. *Proceedings of the National Academy of Sciences*, *113*(9), 2532-2537. https://doi.org/10.1073/pnas.1505913113

Blackburn, G., Ashton, B. J., Thornton, A., Hunter, H., Woodiss-Field, S., & Ridley, A. R. (2024). Investigating the relationship between physical cognitive tasks and a social cognitive task in a wild bird. *Animal Cognition*, *27*(1), 52. https://doi.org/10.1007/s10071-024-01892-4

Blackburn, G., Broom, E., Ashton, B. J., Thornton, A., & Ridley, A. R. (2022). Heat stress inhibits cognitive performance in wild Western Australian magpies, Cracticus tibicen dorsalis. *Animal Behaviour*, *188*, 1-11. https://doi.org/10.1016/j.anbehav.2022.03.016

Blair, R. B. (1996). Land use and avian species diversity along an urban gradient. *Ecological applications*, *6*(2), 506-519. https://doi.org/10.2307/2269387

Boogert, N. J., Anderson, R. C., Peters, S., Searcy, W. A., & Nowicki, S. (2011). Song repertoire size in male song sparrows correlates with detour reaching, but not with other cognitive measures. *Animal Behaviour*, *81*(6), 1209-1216. https://doi.org/10.1016/j.anbehav.2011.03.004

Cauchard, L., Bize, P., & Doligez, B. (2024). How to solve novel problems: the role of associative learning in problem-solving performance in wild great tits Parus major. *Animal Cognition*, *27*(1), 32. https://doi.org/10.1007/s10071-024-01872-8

Cole, E. F., Morand-Ferron, J., Hinks, A. E., & Quinn, J. L. (2012). Cognitive ability influences reproductive life history variation in the wild. *Current Biology*, *22*(19), 1808-1812. https://doi.org/10.1016/j.cub.2012.07.051

Cole, E. F., & Quinn, J. L. (2011). Personality and problem-solving performance explain competitive ability in the wild. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1731), 1168-1175. https://doi.org/10.1098/rspb.2011.1539

Dickinson, A. (2012). Associative learning and animal cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1603), 2733-2742. https://doi.org/10.1098/rstb.2012.0220

Edwards, E. K., Mitchell, N. J., & Ridley, A. R. (2015). The impact of high temperatures on foraging behaviour and body condition in the Western Australian Magpie Cracticus tibicen dorsalis. *Ostrich*, *86*(1-2), 137-144. https://doi.org/10.2989/00306525.2015.1034219

Galluccio, E., Ashton, B. J., Dutour, M., Thornton, A., & Ridley, A. R. (2026). Heat stress impacts the relationship between cognition and anti-predator behaviour in wild magpies. *Behavioral Ecology*. https://doi.org/10.1093/beheco/arag046

Gaynor, K. M., Abrahms, B., Manlove, K. R., Oestreich, W. K., & Smith, J. A. (2024). Anthropogenic impacts at the interface of animal spatial and social behaviour. *Philos Trans R Soc Lond B Biol Sci*, *379*(1912), 20220527. https://doi.org/10.1098/rstb.2022.0527

Griffin, A. S., & Diquelou, M. C. (2015). Innovative problem solving in birds: a cross-species comparison of two highly successful passerines. *Animal Behaviour*, *100*, 84-94. https://doi.org/10.1016/j.anbehav.2014.11.012

Griffin, A. S., & Guez, D. (2014). Innovation and problem solving: a review of common mechanisms. *Behavioural Processes*, *109*, 121-134. https://doi.org/10.1016/j.beproc.2014.08.027

Griffin, A. S., & Guez, D. (2015). Innovative problem solving in nonhuman animals: the effects of group size revisited. *Behavioral Ecology*, *26*(3), 722-734. https://doi.org/10.1093/beheco/aru238

Harrington, K. J., & Lambert, M. L. (2025). Beyond criterion: cognitive flexibility in wild striated caracaras. *Biology letters*, *21*(11). https://doi.org/10.1098/rsbl.2025.0495

Harris, H., Wat, K. K. Y., Banks, P. B., Greenville, A., & McArthur, C. (2023). Grow up, be persistent, and stay focused: keys for solving foraging problems by free-ranging possums. *Behavioral Ecology*, *34*(5), 790-803. https://doi.org/10.1093/beheco/arad054

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794. https://doi.org/10.7717/peerj.4794

Kark, S., Iwaniuk, A., Schalimtzek, A., & Banker, E. (2007). Living in the city: can anyone become an ‘urban exploiter'? *Journal of Biogeography*, *34*(4), 638-651. https://doi.org/10.1111/j.1365-2699.2006.01638.x

Lambert, Megan L., Jacobs, I., Osvath, M., & von Bayern, Auguste M. P. (2019). Birds of a feather? Parrot and corvid cognition compared. *Behaviour*, *156*(5-8), 505-594. https://doi.org/10.1163/1568539X-00003527

Lee, V. E., & Thornton, A. (2021). Animal cognition in an urbanised world. *Front Ecol Evol*, *9*. https://doi.org/10.3389/fevo.2021.633947

Lowry, H., Lill, A., & Wong, B. B. M. (2013). Behavioural responses of wildlife to urban environments. *Biological Reviews*, *88*(3), 537-549. https://doi.org/10.1111/brv.12012

Martin, G. R. (2007). Visual fields and their functions in birds. *Journal of Ornithology*, *148*(2), 547-562. https://doi.org/10.1007/s10336-007-0213-6

McGillycuddy, M., Popovic, G., Bolker, B. M., & Warton, D. I. (2025). Parsimoniously Fitting Large Multivariate Random Effects in glmmTMB. *Journal of Statistical Software*, *112*(1), 1 - 19. https://doi.org/10.18637/jss.v112.i01

Mettke-Hofmann, C., Winkler, H., & Leisler, B. (2002). The significance of ecological factors for exploration and neophobia in parrots. *Ethology, v.108, 249-272 (2002)*, *108*. https://doi.org/10.1046/j.1439-0310.2002.00773.x

Mirville, M. O., Kelley, J. L., & Ridley, A. R. (2016). Group size and associative learning in the Australian magpie (Cracticus tibicen dorsalis). *Behavioral Ecology and Sociobiology*, *70*(3), 417-427. https://doi.org/10.1007/s00265-016-2062-x

Møller, A. P. (2009). Successful city dwellers: a comparative study of the ecological characteristics of urban birds in the Western Palearctic. *Oecologia*, *159*(4), 849-858. https://doi.org/10.1007/s00442-008-1259-8

Morand-Ferron, J. (2017). Why learn? The adaptive value of associative learning in wild populations. *Current Opinion in Behavioral Sciences*, *16*, 73-79. https://doi.org/10.1016/j.cobeha.2017.03.008

Morand-Ferron, J., Cole, E. F., & Quinn, J. L. (2016). Studying the evolutionary ecology of cognition in the wild: a review of practical and conceptual challenges. *Biological Reviews*, *91*(2), 367-389. https://doi.org/10.1111/brv.12174

Morand-Ferron, J., & Quinn, J. L. (2011). Larger groups of passerines are more efficient problem solvers in the wild. *Proceedings of the National Academy of Sciences*, *108*(38), 15898-15903. https://doi.org/10.1073/pnas.1111560108

Petelle, M. B., Jacobs, P. J., & le Roux, A. (2023). Innovative problem-solving in a small, wild canid. *Animal Cognition*, *26*(2), 405-413. https://doi.org/10.1007/s10071-022-01678-6

R Core Team. (2025). R: A Language and Environment for Statistical Computing. In: R Foundation for Statistical Computing.

Roth, T. C., LaDage, L. D., & Pravosudov, V. V. (2010). Learning capabilities enhanced in harsh environments: a common garden approach. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1697), 3187-3193. https://doi.org/10.1098/rspb.2010.0630

Rowell, M. K., Pillay, N., & Rymer, T. L. (2021). Problem solving in animals: proposal for an ontogenetic perspective. *Animals (Basel)*, *11*(3). https://doi.org/10.3390/ani11030866

Shettleworth, S. J. (2001). Animal cognition and animal behaviour. *Animal Behaviour*, *61*(2), 277-286. https://doi.org/10.1006/anbe.2000.1606

Shettleworth, S. J. (2009). Cognition, evolution, and behavior. *Oxford University Press*. https://doi.org/10.1093/oso/9780195319842.001.0001

Sih, A., Ferrari, M. C., & Harris, D. J. (2011). Evolution and behavioural responses to human-induced rapid environmental change. *Evol Appl*, *4*(2), 367-387. https://doi.org/10.1111/j.1752-4571.2010.00166.x

Sol, D., Duncan, R. P., Blackburn, T. M., Cassey, P., & Lefebvre, L. (2005). Big brains, enhanced cognition, and response of birds to novel environments. *Proceedings of the National Academy of Sciences*, *102*(15), 5460-5465. https://doi.org/10.1073/pnas.0408145102

Sol, D., Lapiedra, O., & Ducatez, S. (2020). Cognition and adaptation to urban environments. *Urban Evolutionary Biology*. https://doi.org/10.3389/fevo.2022.953494

Soravia, C., Ashton, B. J., Thornton, A., & Ridley, A. R. (2023). High temperatures are associated with reduced cognitive performance in wild southern pied babblers. *Proc Biol Sci*, *290*(2011), 20231077. https://doi.org/10.1098/rspb.2023.1077

Speechley, E. M., Ashton, B. J., Thornton, A., Simmons, L. W., & Ridley, A. R. (2024). Heritability of cognitive performance in wild Western Australian magpies. *Royal Society Open Science*, *11*(3), 231399. https://doi.org/doi:10.1098/rsos.231399

Stanton, L. A., Cooley-Ackermann, C., Davis, E. C., Fanelli, R. E., & Benson-Amram, S. (2024). Wild raccoons demonstrate flexibility and individuality in innovative problem-solving. *Proceedings of the Royal Society B: Biological Sciences*, *291*(2027), 20240911. https://doi.org/10.1098/rspb.2024.0911

Szabo, B., Valencia-Aguilar, A., Damas-Moreira, I., & Ringler, E. (2022). Wild cognition – linking form and function of cognitive abilities within a natural context. *Current Opinion in Behavioral Sciences*, *44*, 101115. https://doi.org/10.1016/j.cobeha.2022.101115

Taylor, A. H. (2014). Corvid cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(3), 361-372. https://doi.org/10.1002/wcs.1286

Tebbich, S., Griffin, A. S., Peschl, M. F., & Sterelny, K. (2016). From mechanisms to function: an integrated framework of animal innovation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1690), 20150195. https://doi.org/10.1098/rstb.2015.0195

Townsend, A. K., Williams, K. E. G., & Nannas, N. J. (2025). Inbreeding and cognitive impairment in animals. *Behavioral Ecology*, *36*(1), arae101. https://doi.org/10.1093/beheco/arae101

Tyrrell, L. P., & Fernández-Juricic, E. (2017). Avian binocular vision: It's not just about what birds can see, it's also about what they can't. *PLOS ONE*, *12*(3), e0173235. https://doi.org/10.1371/journal.pone.0173235

Wooster, E. I. F., Gaynor, K. M., Carthey, A. J. R., Wallach, A. D., Stanton, L. A., Ramp, D., & Lundgren, E. J. (2024). Animal cognition and culture mediate predator–prey interactions. *Trends in Ecology & Evolution*, *39*(1), 52-64. https://doi.org/10.1016/j.tree.2023.09.012