nature research

Corresponding author(s): Professor Christian Rutz

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

We attempted to trap a large, representative sample of our study population.

2. Data exclusions

Describe any data exclusions.

Of 41 crows trapped, we immediately released 6 birds because of their breeding or health status, and 1 bird escaped on the day of capture. The remaining 34 crows were pre-tested, to confirm their hook-tool-making abilities; of these: 3 were released before we finalised the experimental protocol; 9 could not be confirmed to manufacture hooked tools during the pre-testing period; and 5 were reluctant to engage with the apparatus during their first experimental trial, and were therefore released. The remaining 17 hook-tool proficient crows (8 birds in 2012; 9 in 2013) participated in treatments 1a and 1b, and 8 of these (2012 field season only) also experienced treatments 2a and 2b.

3. Replication

Describe whether the experimental findings were reliably reproduced.

Findings were replicated reliably across crows (17 subjects participated in treatments 1a and 1b, and 8 of these also experienced treatments 2a and 2b), with subjects having access in each trial to six replicates of the three different foraging tasks (for sample sizes, see Fig. 1e). For consistency, all videos of crow behaviour were scored in randomised order by the same hypothesis-naive observer (co-author Caitlin Higgott), and only these scores were used for final analyses. Rescoring of a subsample by a second observer (first author James St Clair) showed very high agreement: in 33 extractions from two trials, scorers differed only by a mean of 0.4 s, which is 1.8% of the median extraction duration of 21.8 s, with a maximum disagreement of 1.6 s (Pearson correlation, $r = 0.9997$, $t = 246.1$, d.f. = 1,31, $p < 0.0001$; Supplementary Fig. 1). The effect size of all experimental results (see Fig. 1d) by far exceeded the small measurement errors of our video-scoring methodology.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

We used a within-subject design where crows experienced multiple experimental treatments (either 1a and 1b; or 1a, 1b, 2a and 2b). Within subjects, the order of treatments was randomised. The allocation of prey types to holes in task logs was randomised between subjects, but kept constant across treatments within subjects. In treatments 2a and 2b, tools were set out on the tool-presentation log in random order. Videos of crow behaviour were scored in randomised order.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

All videos of crow behaviour were scored by the same hypothesis-naive observer (co-author Caitlin Higgott).

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☒	☐ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☒	☐ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

- Therneau, T.M. coxme: Mixed effects Cox models. R package version 2.2-5. (2012); http://cran.r-project.org/package=coxme
- Therneau, T.M. survival: Survival analysis. R package version 2.38. (2012); http://cran.r-project.org/package=survival

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

No unique materials were used.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used.

b. Describe the method of cell line authentication used.

No eukaryotic cell lines were used.

c. Report whether the cell lines were tested for mycoplasma contamination.

No eukaryotic cell lines were used.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

No eukaryotic cell lines were used.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

We tested wild-caught, temporarily-captive New Caledonian crows (Corvus moneduloides). Our experimental sample consisted of 13 females (3 adults; 6 immatures; 4 juveniles) and 4 males (1 adult; 3 immatures), while the sample of excluded crows consisted of 11 females (5 adults; 2 immatures; 4 juveniles) and 13 males (7 adults; 4 immatures; 2 juveniles) (sexing for all-but-one subject based on morphology).

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

Our study did not involve human research participants.