CYRIL RUOSO/NPL

**A common raven (*Corvus corax*) engages with an experimental task that is mounted above the ground.**

# How STRANGE are your study animals?

**Michael M. Webster & Christian Rutz**

A new framework for animal-behaviour research will help to avoid sampling bias — ten years on from the call to widen the pool of human participants in psychology studies beyond the WEIRD.

Ten years ago this week, researchers pointed out that many important findings in human experimental psychology cannot be generalized because study participants are predominantly drawn from a small, unrepresentative subset of the world's population: societies that are Western, Educated, Industrialized, Rich and Democratic (WEIRD)[1]. Mounting evidence suggests that there could be similar sampling problems in research on animals. Behavioural studies of a wide range of species — from insects to primates — could be affected, with researchers testing individuals that are not fully representative of the wider populations they seek to understand. For example, certain sampling protocols are likely to trap the boldest animals, potentially skewing experimental results[2].

It is high time for scientists who work on animal behaviour to identify, and mitigate, potential sampling biases. Simply gathering more data is not a solution, because researchers should always strive to minimize the number of experimental animals used. Instead, we propose a framework with a fitting acronym — STRANGE — that researchers can use to interrogate how unusual their study subjects are. This will aid the design of new studies, and enable potential biases to be declared and discussed when publishing completed work.

## WEIRD humans

In June 2010, human-evolutionary biologist Joe Henrich and his co-authors published a landmark paper challenging the widely held assumption that human behaviour varies little across populations[1]. They highlighted that the vast majority of research on human behaviour, and its cognitive basis, was conducted using participants from societies who are often outliers in broader comparisons. For example, WEIRD people are unusual in how they find their way around, what they consider to be fair, and their willingness to punish others[1].

Several research fields chart the behaviour of non-human animals, including comparative psychology, ethology, behavioural ecology, evolutionary biology and conservation science. They attempt to uncover the evolutionary origins, developmental pathways, adaptive function, underlying mechanisms and conservation relevance of a wide variety of behaviours. Their common goal is to identify general biological principles that apply within and across taxa. Many animal studies are susceptible to sampling biases.

## STRANGE framework

One well-known source of bias is the excessive reliance in some research fields on a few model organisms[3], such as fruit flies (*Drosophila* spp.) and laboratory mice (*Mus musculus*), which inevitably limits the generality of findings. Less obvious, but perhaps much more problematic, is the fact that many species exhibit substantial variation in behaviour, and only certain subsets of that diversity might (unintentionally) turn up in test samples. In primates, for example, individuals of particular 'personalities' can be more likely to take part in trials voluntarily[4] (see 'Beware of STRANGE'). Life-long participation in experiments can also significantly alter subjects' natural behaviour.

When developing the framework, our starting observation was that an animal's behaviour is shaped[5] by its genetic make-up, experience and social background. Ensuring that subjects are representative with regards to these three factors is paramount. Adding a few specific effects, which are well-documented yet often overlooked, provides the acronym STRANGE. It stands for: Social background; Trappability and self-selection; Rearing history; Acclimation and habituation; Natural changes in responsiveness; Genetic make-up; and Experience. STRANGE-related biases can affect both laboratory and field studies. They can influence which animals are sampled for testing, the extent to which they participate in experiments and, importantly, the behaviours that they exhibit during trials. This, in turn, can complicate comparisons between studies, hampering both the generalizability and the reproducibility of findings.

Our approach uses a simple test to evaluate the robustness of a completed or planned study. Researchers should ask: are my animal subjects strange when compared to the wider population for which I wish to make inferences, in any of the seven categories of the framework? Here, using selected examples, we showcase how sampling biases could affect the behaviours observed in animal studies (for further examples, see 'Beware of STRANGE' and Supplementary Information, Table S1).

**Social background.** This includes an animal's social status, the nature and frequency of its interactions with others, and its past opportunities to learn socially from other individuals or their products. For example, when pheasants (*Phasianus colchicus*) were given a spatial-discrimination task to learn which of two holes contained bait, they performed better when they had been housed in groups of five rather than in groups of three[6].

**Trappability and self-selection.** These are closely related processes. They mean that individuals with particular traits[2] are most likely to be caught or to participate voluntarily in experiments (see 'Beware of STRANGE'). In a classic study on pumpkinseed sunfish (*Lepomis gibbosus*), individuals collected using funnel traps (into which fish have to actively swim) were faster to start eating in the laboratory compared with those trapped using more indiscriminate nets[7].

Trappability effects are expected to be prominent in bio-logging studies in which animals are fitted with electronic tags for remote observation. For example, when researchers tag surfacing whales or seabirds in breeding colonies that are difficult to access, they might inadvertently obtain non-random samples. Self-selection biases are a well-known — but usually neglected — problem in laboratory and field studies of animal cognition[4].

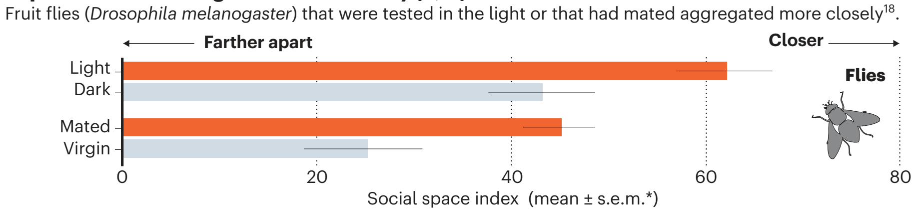
**Rearing history.** This describes an animal's developmental experiences, including the extent to which it has been exposed to a stimulating physical environment, to other animals and to humans. One study showed that captive-reared jumping spiders (*Phidippus audax*) were less active, less exploratory and had reduced interest in prey than were those collected from the wild[8]. Exposure to enrichment, social stimulation and exercise during development can affect brain development and, in turn, cognitive and motor performance[9]. This relationship is well established

## BEWARE OF STRANGE

Many factors can affect animal behaviour, as these examples show, with their STRANGE category acronyms in brackets (the studies shown explicitly set out to examine these effects). Problems arise when test samples are biased with regards to these categories, and when researchers do not account for this.

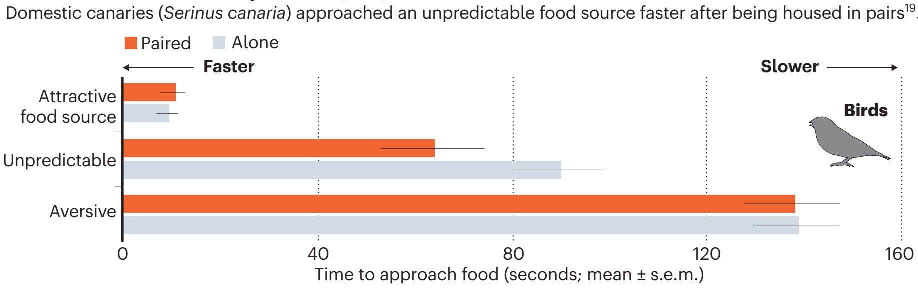
**Experience and timing influence sociality (E, N)**

Fruit flies (*Drosophila melanogaster*) that were tested in the light or that had mated aggregated more closely[18].



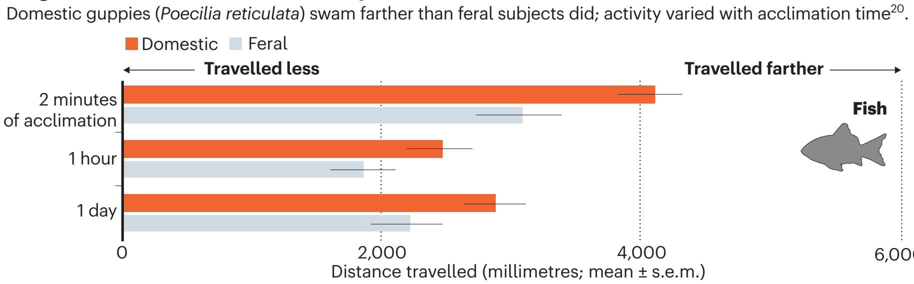

**Social contact induces 'optimism' (S, R)**

Domestic canaries (*Serinus canaria*) approached an unpredictable food source faster after being housed in pairs[19].



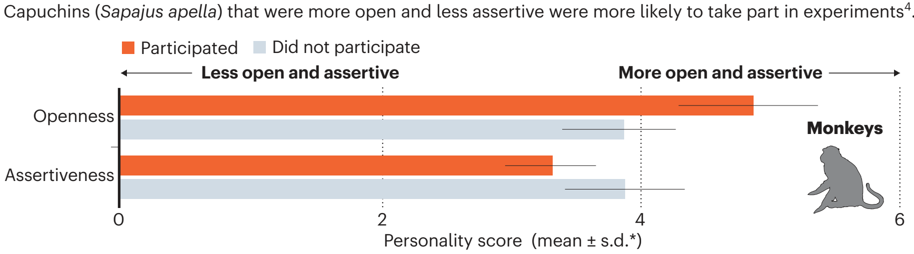

**Origin and acclimation affect activity (G, A)**

Domestic guppies (*Poecilia reticulata*) swam farther than feral subjects did; activity varied with acclimation time[20].




**Self-selection varies with personality (T)**

Capuchins (*Sapajus apella*) that were more open and less assertive were more likely to take part in experiments[4].




*s.d., standard deviation; s.e.m, standard error of the mean.

PETE OXFORD/NPL

**A hawksbill sea turtle (*Eretmochelys imbricata*) tagged with a satellite transmitter.**

for rats and mice[10].

**Acclimation and habituation.** These refer to behavioural changes over time following handling, tagging or exposure to new testing situations. For instance, green turtles (*Chelonia mydas*) spent less time swimming and more time feeding the day after they were tagged with video cameras than they did on the day of tagging. This probably reflects the turtles' recovery from the tagging procedure and their habituation to the cameras[11].

**Natural changes in responsiveness.** These sometimes follow daily, reproductive or seasonal cycles, or the transition from one life stage to another. This means that the timing of experiments is often crucial, as highlighted by a study that found that honeybees (*Apis mellifera*) learn more effectively in the morning than at other times of the day[12].

**Genetic make-up.** This can have profound effects on behaviour. For example, the experience of losing territorial fights early in life has different effects in wild-type brown rats (*Rattus norvegicus*) than in individuals from a common genetic strain. The laboratory rats end up spending less time investigating intruders as adults[13]. There can also be marked differences in behaviour between wild populations, and between males and females, as has been shown in an experimental study on anti-predator behaviour in Trinidadian guppies (*Poecilia reticulata*), in which females spent more time shoaling than did males, for instance[14].

**Experience.** This encompasses opportunities for individual learning, such as participation in earlier experiments. (This means there can be overlap with some of the other framework categories.) After male chaffinches (*Fringilla coelebs*) had been lured by a playback of a rival's song, captured, handled and released again, they sang fewer territorial songs in response to simulated territory intrusions than did birds that had not been captured[15]. Long-lived animals can accumulate complex experimental histories in research laboratories, which in our view must be better documented and accounted for.

**"Concerns arise whenever samples of study subjects are unwittingly biased."**

## Crucial considerations

Some general points about the framework are worth highlighting. First, there are important conceptual differences between WEIRD and STRANGE. The former identifies attributes of a particular demographic group; the latter refers to a suite of factors that can affect behaviour. The seven categories of STRANGE are not problematic by themselves. In fact, they are often the focus of well-designed research projects, such as those we mention here, or are confounding factors that have been explicitly controlled for. Concerns arise whenever samples of study subjects are unwittingly biased with regards to any of these categories, and when researchers overlook that fact.

Second, there is overlap and strong interdependence between some STRANGE categories. For example, the origin of an animal – whether it was wild-caught or captive-bred – will often simultaneously affect its genetic make-up, social background, experience and rearing history. Third, we designed the categories to be broad enough to accommodate future extensions. Finally, although STRANGE refers to samples of animal subjects, effects can be moderated by study protocols. For instance, depending on the species, testing with or without others present can significantly affect an animal's willingness to participate, as often observed in fish, birds and primates.

What will our critics say? Some might note that several of the effects we discuss – such as self-selection and experience biases – have been highlighted as problematic in the past. We feel that, because research practice seems to have improved little in response to specific warnings, it is time to introduce a memorable framework that integrates all factors that could affect the generalizability of animal-behaviour studies.

Others might point out that most of the examples we mention simply illustrate drivers of behavioural variation, and not sampling bias. As we will explain, systematic studies are urgently required to assess our contention that the potential for bias is both widespread and routinely ignored.

## 3D reproducibility

It is possible to identify and mitigate STRANGE-related biases at little or no extra cost, with a simple '3D' approach: design, declare, discuss (see also Supplementary information, Box S1).

**Design.** There are many opportunities for researchers to make their test samples more representative (see Box S1, step 1). For example, we recommend that projects that rely on trapping wild subjects consider using a variety of trap types or bait preparations. They can also sample across multiple populations to reduce systematic biases. Similarly, in studies in which animals effectively select themselves for participation, we encourage researchers to think about ways of altering the testing environment or the task itself, to encourage more inclusive participation. Anecdotal evidence reveals that some crows, for instance, are hesitant to approach experimental tasks that are placed on the floor, but will readily engage with them when they are mounted just one metre above the ground.

**Declare.** When submitting a manuscript, researchers should supply – and journals should ask for – an objective evaluation of a system's 'STRANGEness' so that editors and reviewers can gauge the scope for bias. Although reporting standards have significantly improved over the past few years, a surprising number of journals – including many specializing in animal behaviour and cognition – still lack robust reporting policies. We urge journals to insist that all behavioural studies report detailed subject-attribute data as set out in the ARRIVE guidelines (Animal Research: Reporting of *In Vivo* Experiments)[16], with some additions (such as a full declaration of the attributes of non-participating subjects; see Box S1, step 2a).

Our STRANGE framework can then be used together with this information to evaluate the scope for sampling biases (Box S1, step 2b), and to describe which precautionary steps were taken to avoid bias, if any (Box S1, step 2c). These declarations enhance transparency, provide a valuable resource for systematic reviews and formal meta-analyses, and will hopefully encourage better planning of future studies.

**"We encourage animal-behaviour researchers to routinely ask how unusual their samples are."**

**Discuss.** It is essential to detail any potential issues prominently in the main body of research papers (Box S1, step 3). This would force authors to explicitly link their findings to the studied sample, rather than to the population or species as a whole[17].

As well as supporting all steps of this 3D approach, STRANGE provides a convenient memory aid: we encourage animal-behaviour researchers to routinely ask how unusual their samples are. (For further questions about specific categories, see Table S1.)

## Next steps

There is a large body of literature demonstrating how the categories in our STRANGE framework can affect the behaviours researchers observe. On the basis of this evidence, our personal research experience and our discussions with many colleagues, we suspect that STRANGE-related problems are widespread.

We now urgently need retrospective analyses of published work that quantify how often test samples are biased with regards to these categories, when researchers failed to account for – or declare – these confounding factors.

Human experimental psychology has made great strides towards addressing sampling biases by improving reporting standards. For example, the Association for Psychological Science recommends that authors identify the participant population, explain their selection and consider how generalizable their findings are. However, despite these efforts, problems are surprisingly persistent, with many published studies still providing insufficient detail about their participants[17].

Animal-behaviour scientists have a lot to learn from the WEIRD debate. We hope that our STRANGE framework will help to improve how animal-behaviour research is conducted, reported and interpreted.

**Sampling biases can affect research on model organisms, such as fruit flies.**

SOLVIN ZANKL/NPL

## The authors

**Michael M. Webster** is a lecturer in the School of Biology, University of St Andrews, UK.
**Christian Rutz** is a professor in the School of Biology, University of St Andrews, UK, and worked on this article while he was the 2019–2020 Grass Fellow at the Radcliffe Institute for Advanced Study, Harvard University, Cambridge, Massachusetts, USA.
e-mail: christian.rutz@st-andrews.ac.uk

1. Henrich, J., Heine, S. J. & Norenzayan, A. *Behav. Brain Sci.* **33**, 61–83 (2010).
2. Biro, P. A. & Dingemanse, N. J. *Trends Ecol. Evol.* **24**, 66–67 (2009).
3. Alfred, J. & Baldwin, I. T. *eLife* **4**, e06956 (2015).
4. Morton, F. B., Lee, P. C. & Buchanan-Smith, H. M. *Anim. Cogn.* **16**, 677–684 (2013).
5. Bateson, P. & Laland, K. N. *Trends Ecol. Evol.* **28**, 712–718 (2013).
6. Langley, E. J., van Horik, J. O., Whiteside, M. A. & Madden, J. R. *Anim. Behav.* **142**, 87–93 (2018).
7. Wilson, D. S., Coleman, K., Clark, A. B. & Biederman, L. *J. Comp. Psychol.* **107**, 250–260 (1993).
8. Carducci, J. P. & Jakob, E. M. *Anim. Behav.* **59**, 39–46 (2000).
9. Thornton, A. & Lukas, D. *Phil. Trans. R. Soc. B* **367**, 2773–2783 (2012).
10. van Praag, H., Kempermann, G. & Gage, F. H. *Nature Rev. Neurosci.* **1**, 191–198 (2000).
11. Thomson, J. A. & Heithaus, M. R. *J. Exp. Mar. Biol. Ecol.* **450**, 15–20 (2014).
12. Lehmann, M., Gustav, D. & Galizia, C. G. *Behav. Ecol. Sociobiol.* **65**, 205–215 (2011).
13. Vidal, J., Buwalda, B. & Koolhaas, J. M. *Behav. Processes* **88**, 76–80 (2011).
14. Magurran, A. E. *Phil. Trans. R. Soc. B* **353**, 275–286 (1998).
15. Budka, M., Matyjasiak, P., Typiak, J., Okołowski, M. & Zagalska-Neubauer, M. *J. Ornithol.* **160**, 673–684 (2019).
16. Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M. & Altman, D. G. *PLoS Biol.* **8**, e1000412 (2010).
17. Rad, M. S., Martingano, A. J. & Ginges, J. *Proc. Natl Acad. Sci. USA* **115**, 11401–11405 (2018).
18. Simon, A. F. *et al. Genes Brain Behav.* **11**, 243–252 (2012).
19. Lalot, M., Ung, D., Péron, F., d'Ettorre, P. & Bovet, D. *Behav. Processes* **134**, 70–77 (2017).
20. O'Neill, S. J., Williamson, J. E., Tosetto, L. & Brown, C. *Behav. Ecol. Sociobiol.* **72**, 166 (2018).

Supplementary information accompanies this article; see go.nature.com/2uefq6g.

Setting the agenda in research

# Comment

## Supplementary information to:

# How STRANGE are your study animals?

A Comment published in *Nature* **582**, 337–340 (2020)
https://doi.org/10.1038/d41586-020-01751-5

**Michael M. Webster & Christian Rutz**

**Table S1 | Further examples for the seven categories of the STRANGE framework (for details, see notes beneath the table)**

| STRANGE categories | Suggested questions | Selected examples |
|---|---|---|
| **Social background** includes an animal's social status, the nature and frequency of its social interactions, and its past access to social-learning opportunities. | – *Does the test sample of subjects have an unusual social background?*<br>– *What is the social rank of the subjects? Could this affect participation in experiments?*<br>– *Are subjects housed alone or in groups? If in groups, of what size?*<br>– *Are subjects tested alone or in the presence of other animals? Are testing conditions adapted for non-participating subjects?*<br>– *What social experiences did subjects have prior to testing (e.g., of aggression, courtship, or mating)?*<br>– *Could subjects have previously acquired information via social learning that affects their test performance?* | – Social rank affected innovation in chimpanzees[21].<br>– Dominance rank positively correlated with cognitive performance in starlings[22].<br>– Rearing density affected social information use and shoaling in guppies[23].<br>– Mating experience influenced courtship and mate competition in fruit flies[24].<br>– Social dominance interacted with social rearing condition to shape boldness and aggressiveness in skinks[25].<br>– Spatial discrimination ability positively correlated with social rank in male pheasants[26].<br>– Natural group size was positively correlated with cognitive performance in Australian magpies[27–29]. |
| **Trappability and self-selection** are closely related processes that result, respectively, in individuals with certain characteristics (such as particular 'personality' types) being more likely to be trapped, or to participate voluntarily in experiments. Trappability effects are expected to be prominent in bio-logging studies where subjects are fitted with electronic tags for remote observation, while self-selection biases are a well-known – but usually neglected – problem in laboratory and field studies of animal cognition. | – *If animals are collected using traps, could this introduce sampling bias (e.g., by targeting bolder, more active, or hungrier individuals)?*<br>– *Are different trapping methods used to avoid bias (e.g., different trap types, bait preparations, or trap placement strategies)?*<br>– *If a self-selecting experimental design is used, do all potential subjects participate? What are the attributes of the non-participating subjects?*<br>– *Can you rule out systematic bias in participation (e.g., by social rank, or personality type)?*<br>– *Are test conditions adjusted to allow participation of otherwise excluded subjects (e.g., by amending the set-up or testing environment)?* | – Trappability of badgers varied between study sites, age and season[30].<br>– More exploratory and risk-taking flycatchers were more likely to enter traps[31].<br>– Bolder agamas entered traps sooner than shyer ones[32].<br>– Faster growing trout were more likely to be captured in nets[33].<br>– Faster exploring great tits were more likely to enter camera-equipped nest boxes[34].<br>– Pheasant chicks' self-selection in experiments varied with sex, condition, personality and experience[35].<br>– Self-selected participation in experiments was correlated with personality traits in squirrel |

| | | |
|---|---|---|
| | | monkeys[36].<br>– Sex, condition, and trap type affected trappability and trap-happiness of lampreys[37]. |
| **Rearing history** describes an animal's developmental experiences, including the extent to which it has been exposed to a stimulating physical environment, other animals, and humans. Exposure to enrichment, social stimulation and exercise during development can affect brain development, and in turn, cognitive and motor performance. | – *Does the test sample of subjects have an unusual rearing history?*<br>– *What is known about the origin of the subjects? Are they collected from the wild or captive-bred?*<br>– *If captive-bred, were they raised by their parents (in species with parental care), by unrelated conspecifics, or by humans?*<br>– *To what extent are subjects habituated to humans and testing environments?*<br>– *To what extent have subjects experienced physical enrichment?*<br>– *Are subjects housed alone or in groups?*<br>– *If housed in groups, to what extent are these similar in size and composition to the groups these animals live in in nature?*<br>– *Are differences in subjects' rearing history accounted for?* | – Female fruit flies reared alone were more aggressive[38].<br>– Male fruit flies from enriched environments had greater mating success[39].<br>– Enrichment enhanced spatial memory in mice[40].<br>– Hand-reared cranes were less vigilant than those reared by parents[41].<br>– Environmental variability promoted behavioural flexibility in cod[42].<br>– Environmental enrichment reduced habituation and problem-solving times in rattlesnakes[43].<br>– Enculturation affected tool-use performance in chimpanzees[44].<br>– Environmental enrichment was associated with 'optimistic' response biases in starlings[45].<br>– Environmentally-enriched salmon took fewer risks[46].<br>– Environmentally-enriched zebrafish were more aggressive[47]. |
| **Acclimation and habituation** can result in behavioural changes over time, following handling, tagging, or exposure to novel testing situations. | – *Is the test sample of subjects unusual with regards to acclimation and habituation?*<br>– *Could the behaviour of subjects be affected by the presence of a human observer?*<br>– *Could the behaviour of subjects be affected by the presence of experimental equipment?*<br>– *Do subjects have sufficient time to acclimate to captivity, and is acclimation time standardized?*<br>– *Do all subjects acclimate to experimental conditions at the same rate?* | – Habituation to human observers reduced defensive behaviour over a period of days in Magellanic penguins[48].<br>– Ravens habituated to different modes of gaze following at different rates[49].<br>– Human observer presence reduced feeding and other behaviours in unhabituated baboons and macaques[50].<br>– Behaviour of damselfish took two days to stabilize after being brought into captivity[51].<br>– Reef fish gradually acclimated to the presence of |

| | | |
|---|---|---|
| | | cameras, but not to human observers[52]. |
| **Natural changes in responsiveness** sometimes follow daily, reproductive or seasonal cycles, or the transition from one life stage to another. This means that the timing of experiments is often critical. | – *Is the test sample of subjects unusual with regards to any natural changes in responsiveness?*<br>– *Is the study species known to exhibit diel, or seasonal variation in behaviour?*<br>– *Is the timing of experiments standardized to account for possible effects due to time of day, photoperiod, or season?*<br>– *Are all subjects of the same developmental stage, age and reproductive state? If not, how may this affect the behaviours observed?* | – White suckers had circadian activity patterns, and these were more stable in shoals[53].<br>– Young female guppies were more sensitive to model age when copying mate choice[54].<br>– Gravid female garter snakes were less active and sheltered in different landscape features[55].<br>– Cognitive performance and anxiety responses varied with estrous cycle in wild-type female mice[56].<br>– Reproductive state affected the strength of response to social information in sticklebacks[57].<br>– Zebra finches exhibited circadian rhythmicity of activity and singing[58].<br>– Enrichment affected memory in cricket nymphs, but not adults[59]. |
| **Genetic make-up** can have profound effects on behaviour. There can be marked behavioural differences in genetic make-up between wild populations, and between wild and laboratory populations, often hampering attempts at broader generalization. Sex differences in behaviour are well documented. | – *Is the test sample of subjects unusual with regards to its genetic make-up?*<br>– *Is the test sample of subjects sex-biased?*<br>– *Are the subjects from a specific genetic line?*<br>– *Is the line chosen suitable for examining the behaviour of interest? In other words, can artificial selection (or lack of natural selection) have affected behavioural competency or test performance?*<br>– *If the subjects are wild-type, is the source population known?*<br>– *Are inferences explicitly linked to the genetic line or study population investigated?* | – Fear responses differed between divergent strains of Japanese quail[60].<br>– Genetic strains of mice differed in exploratory behaviour and cognitive performance[61].<br>– Population-level variation in male fruit fly courtship songs had a genetic basis[62].<br>– Males and females differed in their responses to predator cues in a strain of mice[63].<br>– Differences in boldness and anti-predator behaviour between wild and domestic zebrafish had a genetic basis[64].<br>– Natural selection via predation drove differences in shoaling behaviour between guppy populations[65].<br>– Different strains of the parasite *Toxoplasma gondii* had differing effects on host behaviour[66]. |

| | | |
|---|---|---|
| | | – Behavioural syndromes varied between natural populations of the delicate skink[67]. |
| **Experience** encompasses opportunities for individual learning, such as participation in earlier experiments. Long-lived animals can accumulate complex experimental histories in research laboratories, which must be documented and accounted for. | – *Does the test sample of subjects have unusual experience?*<br>– *Have subjects participated in similar earlier experiments that may affect test performance?*<br>– *Have subjects participated in different experiments that may affect test performance?*<br>– *Have subjects experienced husbandry procedures that may affect their behaviour?*<br>– *Have subjects accrued experiences in the wild that may affect test performance?*<br>– *Have subjects had previous opportunities to learn that may overshadow learning in the present study?*<br>– *Are differences in subjects' experience accounted for?* | – Experience of clustered versus dispersed food shaped social-foraging behaviour of pollock[68].<br>– Experience of predator cues affected anti-predator behaviour of freshwater snails[69].<br>– Repeated disturbance of laboratory-housed Poeciliid fish increased their boldness[70].<br>– Tadpoles expressed time-of-day specific anti-predator behaviour based on experience[71].<br>– Learning one solution to an experimental task inhibited learning of alternative solutions in chimpanzees[72].<br>– Magpies recognized individual humans and behaved differently towards them[73].<br>– Defensive behaviour of gopher tortoises varied with experience of human disturbance[74].<br>– Exposure to novel prey reduced wariness towards different novel prey in great and coal tits[75].<br>– Experience of courtship and mating altered personality traits in sticklebacks[76]. |

The STRANGE framework collates a suite of factors that can affect animal behaviour; the acronym stands for: Social background, Trappability and self-selection, Rearing history, Acclimation and habituation, Natural changes in responsiveness, Genetic make-up, and Experience. As noted in the main text, these factors are often the focus of well-designed research projects, like the ones listed in this table, or are confounds that have been explicitly controlled for. But problems arise whenever samples of study subjects are biased with regards to one or several of the seven categories, and researchers do not account for this. Such unexplained variation can significantly impact the interpretation of experimental outcomes, limit the generalizability of findings, complicate comparisons between studies, and hamper reproducibility. We therefore recommend that researchers routinely ask themselves: *Are my animal subjects unusual – or strange – when compared to the wider population for which I wish to make inferences, in any of the seven categories of the STRANGE framework?* In this table, we suggest a non-exhaustive set of additional, category-specific questions researchers may find useful when trying to mitigate, or detect, STRANGE-related biases when designing their experiments or interpreting their findings (for a step-by-step guide to using the STRANGE framework, see Box S1). Note that there is overlap and strong interdependence between some STRANGE categories, and that for some of the examples listed here, several may apply. Examples were chosen to cover a broad range of taxa and study contexts, and are listed chronologically by publication date.

**Box S1 | The 3D approach to using the STRANGE framework – recommendations for researchers and journals**

**(1) DESIGN**

Consult the ARRIVE guidelines[16] and STRANGE framework (this article) when planning studies.

**(2) DECLARE**

Include the following text in journal author guidelines or reporting summaries:

*(a) Provide detailed information – as applicable – on the origin (incl. trapping method), sex, age/developmental stage, mass/body condition, social status, personality type, housing conditions (incl. social contacts and enrichment), past opportunities for individual and social learning, experimental history, and testing protocols (incl. social context), for:*

- *the final sample of subjects contributing data to the study; and*
- *the subjects that were part of the original sample, but did not contribute data (describe reasons for exclusion).*

*(b) Evaluate scope for sampling biases based on the declarations made under (a), especially with regards to subjects' origin, self-selection behaviour, and prior experience.*

*(c) Describe what efforts (if any) were undertaken to mitigate potential sampling biases, especially with regards to sourcing representative subjects (such as using a variety of trapping methods), or adjusting experimental protocols to suit non- or slowly-engaging individuals.*

**(3) DISCUSS**

Summarize the declarations in step (2) in two brief statements in the main text of research articles: one in the Methods section evaluating the STRANGEness of the test sample, and another in the Discussion section explaining how potential biases may limit the generalizability of the reported findings.

## Supplementary References


21. Reader, S. M. & Laland, K. N. *Int. J. Primatol.* **22**, 787–805 (2001).
22. Boogert, N. J., Reader, S. M. & Laland, K. N. *Anim. Behav.* **72**, 1229–1239 (2006).
23. Chapman, B. B., Ward, A. J. W. & Krause, J. *Anim. Behav.* **76**, 923–929 (2008).
24. Saleem, S., Ruggles, P. H., Abbott, W. K. & Carney, G. E. *PLoS ONE* **9**, e96639 (2014).
25. Riley, J. L., Noble, D. W., Byrne, R. W. & Whiting, M. J. *R. Soc. Open Sci.* **4**, 161082 (2017).
26. Langley, E. J., van Horik, J. O., Whiteside, M. A. & Madden, J. R. *R. Soc. Open Sci.* **5**, 171475 (2018).
27. Ashton, B. J., Ridley, A. R., Edwards, E. K. & Thornton, A. *Nature* **554**, 364–367 (2018).
28. Ashton, B. J., Ridley, A. R. & Thornton, A. *Learn. Behav.* **47**, 277–279 (2019).
29. Smulders, T. V. *Learn. Behav.* **47**, 275–276 (2019).
30. Tuyttens, F. A. M. *et al. J. Appl. Ecol.* **36**, 1051–1062 (2019).
31. Garamszegi, L. Z., Eens, M. & Török, J. *Anim. Behav.* **77**, 803–812 (2009).
32. Carter, A. J., Heinsohn, R., Goldizen, A. W. & Biro, P. A. *Anim. Behav.* **83**, 1051–1058 (2012).
33. Biro, P. A. *Oecologia* **171**, 339–345 (2013).
34. Stuber, E. F. *et al. Behav. Ecol.* **24**, 1092–1098 (2013).
35. van Horik, J. O., Langley, E. J., Whiteside, M. A & Madden, J. R. *Behav. Processes* **134**, 22–30 (2017).
36. Polgár, Z., Wood, L. & Haskell, M. J. *Am. J. Primatol.* **79**, e22639 (2017).
37. Reinhardt, U. G & Hrodey, P. J. *Fishes* **4**, 34 (2019).
38. Ueda, A. & Kidokoro, Y. *Physiol. Entomol.* **27**, 21–28 (2002).
39. Dukas, R. & Mooers, A. Ø. *Anim. Behav.* **66**, 741–749 (2003).
40. Frick, K. M. & Fernandez, S. M. *Neurobiol. Aging* **24**, 615–626 (2003).
41. Kreger, M. D., Estevez, I., Hatfield, J. S. & Gee, G. F. *Appl. Anim. Behav. Sci.* **89**, 243–261 (2004).
42. Braithwaite, V. A. & Salvanes, A. G. *Proc. R. Soc. B* **272**, 1107–1113 (2005).
43. Almli, L. M. & Burghardt, G. M. *J. Appl. Anim. Welf. Sci.* **9**, 85–109 (2006).
44. Furlong, E. E., Boose, K. J. & Boysen, S. T. *Anim. Cogn.* **11**, 83–97 (2008).
45. Matheson, S. M., Asher, L. & Bateson, M. *Appl. Anim. Behav. Science* **109**, 374–383 (2008).
46. Roberts, L. J., Taylor, J. & de Leaniz, C. G. *Biol. Conserv.* **144**, 1972–1979 (2011).
47. Woodward, M. A., Winder, L. A. & Watt, P. J. *Fishes* **4**, 22 (2019).
48. Walker, B. G., Dee Boersma, P. & Wingfield, J. C. *Conserv. Biol.* **20**, 146–154 (2006).
49. Schloegl, C., Kotrschal, K. & Bugnyar, T. *Anim. Behav.* **74**, 769–778 (2007).
50. Iredale, S. K., Nevill, C. H. & Lutz, C. K. *Appl. Anim. Behav. Sci.* **122**, 53–57 (2010).
51. Biro, P. A. *Anim. Behav.* **83**, 1295–1300 (2012).
52. Nanninga, G. B., Côté, I. M., Beldade, R. & Mills, S. C. *Ethology* **123**, 705–711 (2017).
53. Kavaliers, M. *Can. J. Zool.* **58**, 1399–1403 (1980).
54. Dugatkin, L. A. & Godin, J. G. J. *Behav. Ecol.* **4**, 289–292 (1993).
55. Charland, M. B. & Gregory, P. T. *J. Zool.* **236**, 543–561 (1995).
56. Walf, A. A., Koonce, C., Manley, K. & Frye, C. A. *Behav. Brain Res.* **196**, 254–260 (2009).
57. Webster, M. M. & Laland, K. N. *Proc. R. Soc. B* **278**, 619–627 (2011).
58. Wang, G., Harpole, C. E., Trivedi, A. K. & Cassone, V. M. *J. Biol. Rhythms* **27**, 145–155 (2012).
59. Mallory, H. S., Howard, A. F. & Weiss, M. R. *PloS ONE* **11**, e0152245 (2016).
60. Mignon-Grasteau, S. *et al. Behav. Processes* **61**, 69–75 (2003).
61. Brooks, S. P., Pask, T., Jones, L. & Dunnett, S. B. *Genes Brain Behav.* **4**, 307–317 (2005).
62. Etges, W. J., Over, K. F., De Oliveira, C. C. & Ritchie, M. G. *Anim. Behav.* **71**, 1205–1214 (2006).
63. Adamec, R., Head, D., Blundell, J., Burton, P. & Berton, O. *Physiol. Behav.* **88**, 12–29 (2006).
64. Wright, D., Nakamichi, R., Krause, J. & Butlin, R. K. *Behav. Genet.* **36**, 271–283 (2006).
65. Huizinga, M., Ghalambor, C. K. & Reznick, D. N. *J. Evol. Biol.* **22**, 1860–1866 (2009).
66. Kannan, G. *et al. Folia Parasitol.* **57**, 151–155 (2010).
67. Michelangeli, M., Chapple, D. G., Goulet, C. T., Bertram, M. G. & Wong, B. B. *Behav. Ecol.* **30**, 393–401 (2019).
68. Ryer, C. H. & Olla, B. L. *Anim. Behav.* **49**, 411–418 (1995).
69. Turner, A. M., Turner, S. E. & Lappi, H. M. *Anim. Behav.* **72**, 1443–1450 (2006).
70. Brown, C., Burgess, F. & Braithwaite, V. A. *Behav. Ecol. Sociobiol.* **62**, 237–243 (2017).
71. Ferrari, M. C. & Chivers, D. P. *Anim. Behav.* **78**, 11–16 (2009).
72. Hrubesch, C., Preuschoft, S. & van Schaik, C. *Anim. Cogn.* **12**, 209–216 (2009).
73. Lee, W. Y., Choe, J. C. & Jablonski, P. G. *Anim. Cogn.* **14**, 817–825 (2011).
74. Bateman, P. W., Fleming, P. A., Jones, B. C. & Rothermel, B. B. *Behaviou*r **151**, 1267–1280 (2014).
75. Adamová-Ježová, D., Hospodková, E., Fuchsová, L., Štys, P. & Exnerová, A. *Behav. Processes* **131**, 24–31 (2016).
76. Monestier, C. & Bell, A. M. *Proc. R. Soc. B* **287**, 20192936 (2020).