A meta-analysis of the group-size effect on vigilance in mammals

Guy Beauchamp, a,b* Zhongqiu Li, b* Cong Yu, b Peter A. Bednekoff, c and Daniel T. Blumstein d,e

aIndependent Researcher, Canada, bLab of Animal Behavior and Conservation, School of Life Sciences, Nanjing University, 210023, Nanjing, China, cDepartment of Biology, Eastern Michigan University, Ypsilanti, MI 48197, USA, and dDepartment of Ecology & Evolutionary Biology, University of California, 621 Young Drive South, Los Angeles, CA 90095–1606, USA

Received 26 February 2021; revised 16 April 2021; editorial decision 29 April 2021; accepted 18 May 2021.

Group-size effects, whereby antipredator vigilance decreases as group size increases, are widely reported in mammals and birds but a meta-analysis has only been conducted in birds. We systematically reviewed the literature on mammalian group-size effects, estimated the effect sizes in each study, and conducted a phylogenetic meta-analysis. We obtained 296 effect sizes from 97 species belonging to 10 Orders and 26 Families. Overall, effect sizes indicated a moderate negative effect of group size ($r = -0.44$), but 43% of the effect sizes were compatible with a null effect of group size. There was significant heterogeneity in effect sizes. Weaker effect sizes occurred when vigilance was measured as a frequency or a duration rather than as a percentage of time spent vigilant, when measured in closed habitats, during the reproductive season, and in mixed-sex groups or during times when juveniles were absent. We infer a "file drawer problem" because there were relatively few studies with smaller sample sizes reporting small group-size effects. The results confirm the importance of group size in explaining variation in mammalian vigilance but also suggest which a substantial amount of variation remains unexplained. We suggest that future studies should aim to study mammalian group-size effects by quantifying the percentage of time allocated to vigilance rather than lower-power methods such as frequency or duration of vigilance.

Key words: antipredator behavior, group-size effect, mammals, meta-analysis, vigilance.

INTRODUCTION

Predation risk shapes the lives of many species of animals. Solitary species faced with such risk can only rely on their capabilities to evade capture. Camouflage, aposematic coloration, or external anatomical defenses in solitary animals can reduce the risk of encounter or the risk of capture by predators (Ruxton et al. 2004). By contrast, species which live in groups have additional means to increase their safety (Caro 2005). For instance, large groups are often more difficult to locate as opposed to a collection of individuals spread out in the habitat, reducing the rate of encounters with predators (Ioannou et al. 2011). If a predator detects the group, it can only approach undetected by evading the senses of every member of the group (Galton 1871). Warning about an impending attack can spread rapidly through a group after detection (Treherne and Foster 1981). Therefore, individuals which have not detected the predator by themselves can use this information and still escape before it is too late. Animals thus tend to initiate their flight at a longer, safer distance in larger groups (Stankowich and Blumstein 2005). Even when predators succeed in attacks, the risk per individual can be diluted for larger groups of potential prey (Bertram 1978). Indeed, if the predator can only capture one individual in the group after an attack, the odds of capture per individual are inversely proportional to group size. Because of all these potential benefits, prey animals which live in groups should experience greater safety (Krause and Ruxton 2002; Beauchamp 2014).

In the above sequence of steps in predator-prey interactions, vigilance is involved at the predator detection stage. Antipredator vigilance represents an allocation of time to detect predators often at the expense of foraging and other fitness-enhancing activities (Beauchamp 2015). Because of the greater safety of individuals in groups, group members could reduce their investment in vigilance at no increased risk to themselves (Pulliam 1973). The expected decrease in vigilance as group size increases is known as the group-size effect on vigilance. The prediction that vigilance decreases with group size can easily be investigated in the field by documenting how various measures of vigilance, such as the percentage of time devoted to vigilance, the number of vigilance bouts per unit time or the average duration of vigilance bouts, change as a function of group size.

Address correspondence to G. Beauchamp. E-mail: guygillesbeauchamp@gmail.com.

© The Author(s) 2021. Published by Oxford University Press on behalf of the International Society for Behavioral Ecology. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com
Hundreds of studies have examined this prediction in birds and mammals (Beauchamp 2015), the two taxa where interruptions during foraging are most often interpreted as an investment in antipredator vigilance. As the number of studies investigating the group-size effect on vigilance accumulated over time, researchers have examined support for the prediction by counting the number of studies reporting a statistically significant effect of group size on vigilance (Elgar 1989; Lima and Dill 1990; Quenette 1990). The use of nonparametric P values to evaluate support for a prediction is not recommended because nonparametric P values are sensitive to sample size. Indeed, different studies reporting the same correlation between vigilance and group size might pass the statistical significance level or not due to differences in sample size.

Meta-analysis provides a framework to evaluate support for a prediction by pooling information from all relevant studies to get an overall effect size (Lapsey and Wilson 2001), which in the case of vigilance studies would be the magnitude of the group-size effect on vigilance. Effect sizes from individual studies are also typically weighted by variables likely to affect their precision, such as sample size. Further analysis within the meta-analysis framework can also help to assess the effect of moderator variables on effect sizes. Moderator variables are ecological variables which can modulate the strength of the association with the dependent variable of interest. In the case of vigilance, examples include habitat type and reproductive season.

Despite a large number of studies available, surprisingly only one meta-analysis has focused on the group-size effect on vigilance (Beauchamp 2008). Using studies in birds, Beauchamp found slight to moderate effect sizes for the magnitude of the group-size effect on vigilance with the strongest effect for the percentage of time spent vigilant and the weakest for vigilance duration. However, the meta-analysis was based on a rather small number of studies which probably explains why few moderator variables could be identified. In addition, this meta-analysis ignored the fact that different estimates from the same species are unlikely to be independent and also that estimates from related species are more likely to be similar due to shared ancestry (Chamberlain et al. 2012).

The goal of our study was to perform a meta-analysis of the group-size effect on vigilance including the use of random effects to account for multiple measures per study, multiple measurements per species, and phylogenetic effects. To compare our results with estimates available in birds, we chose to focus on mammals. Birds and mammals overlap in the small body size range and can share some of the same predators. However, mammals are typically much larger, can live much longer, and typically face much larger terrestrial predators. Whether this has an impact on the magnitude of the group-size effect on vigilance is not known.

In addition, we sought to determine the effect of moderator variables on the magnitude of the group size on vigilance. Models of vigilance predict that the magnitude of the decrease in vigilance with group size should vary with predation risk although they do not agree on whether the largest decrease takes place when predation risk is high or low (Beauchamp 2019). In addition to predation risk, another key moderator variable is the reproductive season. Whereas vigilance plays a role in detecting predators, vigilance can also help to monitor rivals or competitors within the group (Treves 1998; Hirsch 2002; Favreau et al. 2010). Such social vigilance is expected to increase in larger groups thus mitigating the decrease in antipredator vigilance with group size (Beauchamp 2001). Weaker correlations with group size should thus occur during the reproductive season where mating competition is highest (Childress and Lang 2003; Cameron and Du Toit 2005; Li et al. 2012).

Vigilance studies have been conducted under a wide range of conditions. Some studies include foraging animals whereas others include animals drinking or resting. Some are conducted during the day or at night (Beauchamp 2007). Vigilance studies also include a wide range of species from the smallest to the largest and with widely different diets from carnivores to browser or grazer herbivores. Methodologically, some studies use a simple correlation approach to determine the association between vigilance and group size whereas others employ sophisticated models which include group size as well as potentially confounding variables in a multivariable framework. To determine the robustness of the group-size effect on vigilance, we, therefore, sought to determine whether vigilance effect sizes vary under these different conditions.

**METHODS**

**Data collection**

We searched the literature for studies reporting the effect of group size on vigilance in mammals. First, we consulted earlier reviews of vigilance in mammals (Elgar 1989; Quenette 1990; Beauchamp 2017, 2019), and then combined these results with a Web of Science online database search between 1978 and the end of 2020 using the following combinations of keywords: (alert or scan* or vigilan*) and mammal* and (group* or aggrega*). After eliminating duplicates and excluding articles from journals specialized in non-mammalian taxa, we examined each of the remaining articles (Figure 1).

We used the following criteria for inclusion in the meta-analysis: (1) the study reported at least one measure of vigilance as a function of group size (% time spent vigilant, frequency of vigilance per unit time or average duration of vigilance bouts), and (2) the sample size was at least five (otherwise the sampling variance for estimates could not be calculated). The above three measures of vigilance are most reported in the literature. The percentage of time where at least one member of the group is vigilant, which is known as collective vigilance, is also a useful measure of vigilance (Pays et al. 2012; Iranzo et al. 2018), but it was not reported frequently enough to warrant inclusion.

We excluded studies reporting vigilance in the laboratory (e.g. Gosselin-Ildari and Koenig 2012), but retained those conducted with captive animals in outdoor enclosures potentially exposed to a greater variety of threats (e.g. Blumstein et al. 1999). We excluded effect sizes from studies with high levels of human disturbances, such as hunting (Benham et al. 2008), unless the results indicated that such disturbances did not affect the group-size effect on vigilance. We also excluded studies with sentinel systems in which individuals take turns to provide vigilance from a vantage point for the remaining group members (Rasa 1986). This particular form of vigilance is not common (Betchekoff 2015), and we preferred to focus on systems in which all individuals can contribute to overall vigilance. Also excluded were studies in which individuals from one species foraged in close association with nearby individuals from another species (e.g. Cords 1990; Payne and Jarman 1999), because it was not clear how members of other species contribute to the group-size effect on vigilance.

In most studies, group size represented the number of companions within a given radius around a focal individual or within a given distance from the nearest neighbor. We excluded studies where the measurement of group size combined several variables connected to group sizes such as the number of adults,
juveniles, and inter-individual distances (e.g. Blanchard et al. 2017). Similarly, we excluded studies where vigilance was calculated by combining variables such as vigilance on two or four legs and vigilance while foraging (Le et al. 2019). Most studies employed focal sampling to evaluate vigilance, but we also included studies which measured vigilance using the average percentage of individuals in a group vigilant over predetermined intervals.

We gathered the following information for each included study: type of vigilance measure (% time spent vigilant, vigilance frequency or vigilance duration), sample size (the number of individuals sampled or the number of groups sampled), diet (carnivore or herbivore and further distinguishing between browsers and grazers (Jarman 1974)), habitat type (open v. closed depending on the extent of visual obstruction), sex in the groups sampled (male, female or a mixture), activity (foraging, resting, drinking or licking), time of day when the samples were taken (diurnal, nocturnal or both), presence of juveniles, time of year (reproductive season, non-reproductive season or both), the occurrence of food supplementation, predation risk (low when predators were absent or in very low numbers), and multivariable statistical treatment (whether or not the statistical analysis took variables other than group size into account). Whereas we noted the minimum and maximum group size in the sampled groups, it was not possible to obtain this information in 51 cases so that this information was not analyzed.

In many cases, researchers provided estimates of the group-size effect on vigilance in subgroups of individuals or at different times of the year. All these estimates were considered separately but were considered nested within the study id for statistical purposes (see below for details). Finally, we gathered mean body mass for each species from a published source (Smith et al. 2003).

### Statistical analysis

Each estimate of the group-size effect on vigilance was transformed into the correlation coefficient $r$ using standard transformations (Polanin and Snilstveit 2016). Common estimates of the association between vigilance and group size included non-parametric and parametric correlation coefficients, R-square values from linear and non-linear regression models (which we adjusted before taking the square root as suggested (Nakagawa and Cuthill 2007)), $t$-tests from beta estimates, F-tests from linear models, and odds ratios from binomial models. $R$ coefficients were transformed into $Z$ scores using Fisher’s transformation (Sokal and Rohlf 1995). The inverse of the

### Figure 1

Flow diagram describing the sample size at different stages of the meta-analysis of the group-size effect on vigilance in mammals.
The variance of Z-scores was used to weigh each Z score estimate. The overall average Z score was back-transformed to an r-value for presentation purposes.

Before proceeding further, we review some basics for interpreting effect sizes. In essence, each effect size is the amount of variation in vigilance which can be explained by variation in group size. Effect sizes will be smaller when measurement error related to group size is greater, as this error introduces extraneous variation. Reducing the variation in group size can also reduce effect size by reducing the amount of explained variation in vigilance more than the amount of unexplained variation.

We used a phylogenetic meta-analysis framework with the metafor R package (Viechtbauer 2010) to obtain weighted estimates of Z scores. We used species as a random factor to account for multiple estimates for the same species. We also used study id as a random effect to account for multiple measurements within the same study. To account for phylogenetic relatedness, we included the variance-covariance distance matrix between species in each model. The distance matrix was obtained from a 50% majority consensus tree constructed from a set of 1000 phylogenetic trees of the included species using the latest mammalian phylogeny available (Upham et al. 2019). We obtained the consensus tree and branch lengths from the ape R package (Paradis et al. 2004).

Heterogeneity of the Z scores among studies was assessed using the intercept-only model along with the random effects. We then examined the effect of moderator variables on variation in Z scores among studies. The independent variables included vigilance measure type, body mass in \( \log_{10} \) scale, diet (carnivore, herbivore grazer or non-grazer), habitat type (open v. closed), sex (male, female, or a mixture), activity (foraging v. non-foraging), time of day (diurnal v. non-diurnal), juveniles (present or not), time of year (reproductive season, non-reproductive season or both), food supplementation (present or not), predation risk (low or not), and multivariable treatment (present or not). Some categories for the independent variables were collapsed due to the small sample size for some of the categories. We treated the different measures of vigilance as a moderator variable because this allowed us to make a direct comparison of their associated effect sizes. For publication bias, we produced funnel plots of Z scores against their standard errors looking for evidence which small scores are more likely in studies with small standard errors (thus large sample sizes).

### RESULTS

We obtained 296 effect sizes from 1978 through 2020 from 152 different studies (Figure 1). We included 97 species from 10 Orders and 26 Families. The most common Order was Artiodactyla with 181 effect sizes (61%) and the most common Family was Bovidae with 118 effect sizes (40%). Species ranged in size from 0.1 kg to nearly 4000 kg. Vigilance was recorded mostly during feeding (281 effect sizes or about 95%) with some effect sizes collected during drinking, licking, or resting. Predation risk was deemed low in studies on predator-free islands or outdoor enclosures (26 effect sizes or about 9%). The number of effect sizes was 198 for the percentage of time spent vigilant, 74 for vigilance frequency, and 24 for vigilance duration.

The weighted mean value of \( r \) was −0.44 (95% CI: −0.66, −0.22). A total of 128 effect sizes (43%) were compatible with a null effect of group size or a positive effect. The breakdown by the measure of vigilance revealed 82 of 198 non-significant or positive effect sizes for time spent vigilant (41%), 34 of 74 non-significant or positive effect sizes for frequency of vigilance (47%), and 12 of 24 non-significant or positive effect sizes for vigilance duration (50%).

The intercept-only model revealed statistically significant heterogeneity among Z scores (\( Q_{(295)} = 6494.6, P < 0.0001 \)). Moderators included in the model explained a significant amount of variation in Z scores (\( Q_{(16)} = 166.9, P < 0.0001 \)). In particular, the magnitude of the group-size effect on vigilance was weaker when the measurement of vigilance was frequency or duration rather than the percentage of time spent vigilant (Table 1). In addition, the magnitude was weaker in closed rather than open habitats.

### Table 1

**Effect of moderator variables on weighted estimates of Z scores for the relationship between vigilance and group size in mammals (296 effect sizes from 152 different studies representing 97 species)**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Beta (SE)</th>
<th>P value</th>
<th>Direction of effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body mass in log scale</td>
<td>0.024 (0.048)</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Measurement type: Vigilance frequency v. time spent vigilant</td>
<td>0.031 (0.013)</td>
<td>&lt;0.0001</td>
<td>Weaker correlation with vigilance frequency</td>
</tr>
<tr>
<td>Measurement type: Vigilance duration v. time spent vigilant</td>
<td>0.15 (0.019)</td>
<td>&lt;0.0001</td>
<td>Weaker correlation with vigilance duration</td>
</tr>
<tr>
<td>Sex: male v. female</td>
<td>0.019 (0.012)</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Sex: mixed sex v. female</td>
<td>0.15 (0.047)</td>
<td>0.002</td>
<td>Weaker correlation in mixed sex groups</td>
</tr>
<tr>
<td>Diet: carnivore v. grazer</td>
<td>−0.18 (0.17)</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>Diet: non-grazer v. grazer</td>
<td>0.089 (0.079)</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>Habitat: open v. closed</td>
<td>−0.16 (0.041)</td>
<td>&lt;0.0001</td>
<td>Weaker correlation in closed habitats</td>
</tr>
<tr>
<td>Activity: non-feeding v. feeding</td>
<td>0.15 (0.097)</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Time of day: diurnal v. non-diurnal</td>
<td>−0.11 (0.12)</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>Juveniles present v. absent</td>
<td>−0.051 (0.022)</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Food supplementation present v. absent</td>
<td>0.095 (0.11)</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Time of year: reproductive season v. non-reproduction season</td>
<td>0.35 (0.053)</td>
<td>&lt;0.0001</td>
<td>Weaker correlation during reproductive season</td>
</tr>
<tr>
<td>Time of year: mixed v. non-reproduction season</td>
<td>0.15 (0.083)</td>
<td>0.078</td>
<td></td>
</tr>
<tr>
<td>Predation risk low v. normal</td>
<td>−0.0087 (0.0062)</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>Multivariable treatment present v. absent</td>
<td>0.082 (0.06)</td>
<td>0.17</td>
<td></td>
</tr>
</tbody>
</table>
during the reproductive rather than the non-reproductive season, in mixed-sex rather than female groups, and groups without juveniles (Table 1). Inspection of the funnel plot revealed that small estimates for Z scores tended to cluster in the area where the standard error was small indicating a relative lack of studies reporting small group-size effects when the sample size was small and the standard errors large (Figure 2). In other words, the reported group size effect was smaller than the global average for published studies with larger sample sizes and larger than the global average for published studies with smaller sample sizes.

DISCUSSION

We detected a moderate (Cohen 1988) effect of group size on vigilance in mammals. Nevertheless, a large percentage of the effect sizes (43%) were compatible with a null or positive effect of group size. Our analyses identified methodological as well as ecological factors which affect the magnitude of the group-size effect on vigilance in mammals, which might explain why vigilance varied with group size to a different extent in different studies.

Methodologically, group-size effects were stronger on average when vigilance was quantified by the percentage of time allocated to vigilance compared with the frequency of vigilance or its average duration. A weaker correlation with group size for the duration of vigilance has also been documented in birds (Beauchamp et al. 2008). These results are consistent with the idea that variation in vigilance mostly reflects changes in the duration of intervals between vigilance bouts and to a lesser extent changes in the duration of vigilance bouts. Percentage of time vigilant captures the overall change in the duration and frequency of vigilance bouts. By contrast, the average duration of vigilance bouts ignores how frequently they occur, and the frequency of vigilance bouts varies with the amount that intervals between vigilance bouts are adjusted above and beyond the duration of a vigilance bout. (If an animal only adjusted the duration of vigilance bouts while keeping the duration of intervals between vigilance bouts the same, vigilance frequency would change in the opposite direction). We recommend that future studies in mammals investigate the group-size effect on vigilance by conducting focal animal samples and documenting the percentage of time allocated to vigilance. Changes in vigilance may secondarily be described as due to changes in the duration and frequency of vigilance bouts.

Weaker effect sizes in closed habitats might involve statistical considerations as well as biological phenomena. One statistical issue is that group size may be measured less accurately in closed habitats where it is difficult to see all animals and determine which companions can be considered group mates for a focal animal. In addition, the range of group sizes may be smaller in closed habitats. Large, well-defined groups are usually seen in open habitats (Jarman 1974; Underwood 1982), although we note that even in open habitats quantifying group size may present difficulties (Blumstein and Daniel 2003). Smaller variation in group size and greater measurement error can on their own reduce effect sizes. Beyond any statistical effects, however, the basics of predator-prey interactions—attack, detection, information spread, and escape—can differ between closed and open habitats. It is likely that predators can approach closer before detection, collective detection is less effective, and options for coordinated defense or escape are more limited in closed habitats. Thus, the benefits of large groups are potentially reduced in closed habitats perhaps explaining the weaker effect of group size. Disentangling the contribution of all these factors remains a challenge for future studies.

Even with the benefit of larger sample size and updated methods, the group size effect in mammals from this study is much the same as that reported in birds ($r = -0.42$; Beauchamp 2008). This illustrates the robust importance of foraging with others for a broad range of species that differ in body size, diet, and general ecology. Yet, because 43% of the studies were compatible with a null or positive effect of group size, we are left to understand why so many studies failed to detect the predicted relationship. As noted earlier, small effect sizes could be explained by how vigilance was quantified (estimates of the percentage of time allocated to vigilance led to the largest effect sizes). Small effect sizes could also be explained by how group size was quantified. For instance, when group size was quantified as the number of residents in the colony in golden marmots (Marmota caudata), little variation in vigilance was explained by this measure of group size (Blumstein 1996). Additionally, estimates of the group-size effect can differ depending on whether group size is quantified as foraging aggregation size or the number of companions within some set distance (e.g. Blumstein et al. 2001). Finally, small effect sizes could reflect how group size was statistically analyzed. Some studies pool estimates of time allocation or other measures of vigilance at each group size and analyze this relationship. Such aggregation reduces variation and makes it easier to detect a group-size effect. Other studies use raw data and fit complex models, which typically explain less variation. We note that studies using a multivariable approach reported similar or weaker group-size effects than univariable models. We advise caution when comparing group-size effects from different studies using different definitions of vigilance or group size and different statistical procedures.

We assume which group-size effects are protective and represent an antipredator strategy. However, group-size effects on vigilance could result from increased competition in larger groups.

Figure 2
Funnel plot of the magnitude of the group-size effect on time spent vigilant in mammals ($n = 296$). Fisher’s Z scores are shown versus the standard error of the estimates produced by an intercept-only model including random effects. The diagonal dashed lines show the 95% confidence intervals and the vertical dashed lines show the 95% confidence intervals around the mean estimate.
(Clark and Mangel 1986). Models predict a weaker correlation between vigilance and group size when food competition intensity increases (Beauchamp and Ruxton 2003), which might explain why some studies failed to report a stronger association. To assess this hypothesis in the meta-analysis framework, we need to incorporate proxies of competition intensity. Along the same line, vigilance can be directed at rivals in the group in addition to predators. As the need to monitor potentially threatening neighbors typically increases with group size (Treves 1998; Hirsch 2002; Favreau et al. 2010), the addition of social vigilance will decrease the strength of the association between vigilance and group size. We found support for this prediction because the correlation between vigilance and group size weakened during the more competitive breeding season and in mixed-sex groups. Other biologically important differences between the breeding and non-breeding seasons also probably play a role. For instance, large groups are probably less stable during the breeding season, which means that the potential range of group sizes is more limited. In smaller groups, dynamics between group members may reflect factors unrelated to collective detection and risk dilution. The presence of juveniles is also a consideration as it typically leads to higher vigilance (see Beauchamp 2015 for a review), and as shown here to an increase in the magnitude of the group-size effect. All these factors can have an impact on the relationship between vigilance and group size.

Unlike with birds where no publication bias was detected, the mammalian group-size effect data suggest a publication bias commonly referred to as the file drawer problem (Rosenthal 1979). Reported effect sizes were bigger for smaller studies, suggesting either that results compatible with a null effect of group size are not submitted for publication, are rejected when submitted, or that authors only present effect sizes when the results are statistically significant (e.g. Crosmary et al. 2012; van der Meer et al. 2012; Monclús et al. 2015). We advise researchers to provide effect sizes even when the results are not statistically significant, and reviewers and editors to still consider negative results. We hope that the open science movement towards more transparent publication practices and pre-publishing protocols may result in all types of results being published in the future regardless of the statistical outcome. It will be through a thorough analysis of those results that we will be better able to understand both the drivers of group-size effects in mammals as well as factors which constrain them.

**FUNDING**

Z.L. received support from the Tibet Major Science and Technology Project (XZ201901-GA-06) and D.T.B. received support from the National Science Foundation.

Data availability: Analyses reported in this article can be reproduced using the data provided by Beauchamp et al. (2021).

Handling editor: Amanda Ridley

**REFERENCES**


