

## Research



**Cite this article:** White TE, Umbers KDL. 2021

Meta-analytic evidence for quantitative

honesty in aposematic signals. *Proc. R. Soc. B*

**288:** 20210679.

<https://doi.org/10.1098/rspb.2021.0679>

Received: 22 March 2021

Accepted: 8 April 2021

**Subject Category:**

Evolution

**Subject Areas:**

evolution, behaviour, ecology

**Keywords:**

colour signal, anti-predator, aposematism, defence

**Author for correspondence:**

Thomas E. White

e-mail: [thomas.white@sydney.edu.au](mailto:thomas.white@sydney.edu.au)

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5401657>.

# Meta-analytic evidence for quantitative honesty in aposematic signals

Thomas E. White<sup>1</sup> and Kate D. L. Umbers<sup>2,3</sup>

<sup>1</sup>School of Life and Environmental Sciences, The University of Sydney, Sydney, New South Wales 2106, Australia

<sup>2</sup>School of Science, and <sup>3</sup>Hawkesbury Institute for the Environment, Western Sydney University, Penrith, New South Wales 2751, Australia

TEW, 0000-0002-3976-1734; KDLU, 0000-0002-9375-4527

The combined use of noxious chemical defences and conspicuous warning colours is a ubiquitous anti-predator strategy. That such signals advertise the presence of defences is inherent to their function, but their predicted potential for quantitative honesty—the positive scaling of signal salience with the strength of protection—is the subject of enduring debate. Here, we systematically synthesized the available evidence to test this prediction using meta-analysis. We found evidence for a positive correlation between warning colour expression and the extent of chemical defences across taxa. Notably, this relationship held at all scales; among individuals, populations and species, though substantial between-study heterogeneity remains unexplained. Consideration of the design of signals revealed that all visual features, from colour to contrast, were equally informative of the extent of prey defence. Our results affirm a central prediction of honesty-based models of signal function and narrow the scope of possible mechanisms shaping the evolution of aposematism. They suggest diverse pathways to the encoding and exchange of information, while highlighting the need for deeper knowledge of the ecology of chemical defences to enrich our understanding of this widespread anti-predator adaptation.

## 1. Introduction

Predation has driven the evolution of striking adaptations for defence among prey [1–3]. Few strategies are as widespread as aposematism which, in its best-studied form, is characterized by the coupling of chemical defences with conspicuous colour patterns as aversive ‘warnings’ to predators [4]. That such colours are qualitatively honest in signalling the presence of defences is an inherent feature of aposematism, and their evolutionary stability is a predictable consequence of the broad alignment of interests between signallers and receivers [5,6]. In general terms, predators avoid the costs of a noxious meal while prey reduce the risk of injury or death, and the system is resistant to cheating since weakly defended prey are less able to bear the costs of predation [7,8]. While the qualitative honesty of aposematic signals is well established, their potential for quantitative honesty—defined by a positive, linear correlation between colour signal expression and the quality of defences—is less clear.

The question of quantitative honesty in aposematism has received extensive theoretical attention (table 1). The earliest support arose from the consideration of aposematic signals as handicaps (*sensu* [16]), which suggested that positive correlations may emerge assuming either a (unspecified) mechanistic link between chemical defences and signal expression, or more general predation costs which are differentially borne across signallers [17]. While those strictly Zahavian explanations have been criticized for biological implausibility [7], later models both tangential to and outside of the framework identified plausible routes to quantitative honesty (e.g. [9,11,12,18]). Blount *et al.* [9] presented arguments for energy, or energy turnover, as a shared and limited resource in the synthesis of both chemical defences and pigments. This was extended by

**Table 1.** Summary of the predicted associations between the expression of aposematic signals and quality of chemical defences. Scales and signs refer to the level at which quantitative associations may manifest and the direction (+ or −) or absence (none) of any linear correlation, respectively, with 'quantitative honesty' represented by positive correlations at each scale. References detail theoretical treatments supporting the noted mechanisms, and asterisks (\*) indicate scales at which the plausibility of a mechanism was extrapolated (in the original study) rather than explicitly examined.

scale	sign	mechanism(s)	references
among individuals	+	go-slow prey sampling; handicap via resource trade-off; automimicry; exaptation	Blount <i>et al.</i> [9], Guilford [10]; Speed & Franks [11]; Holen & Svenningsen [12]; Lee <i>et al.</i> [13]
	−	high innate defence efficacy; non-limiting resources	Blount <i>et al.</i> [9]; Lee <i>et al.</i> [13]
among populations	+	go-slow prey sampling*; handicap via resource trade-off*; Batesian mimicry; correlated marginal costs; automimicry	Blount <i>et al.</i> [9]; Franks <i>et al.</i> [14], Guilford [10], Speed & Ruxton [15]; Speed & Franks [11]; Holen & Svenningsen [12]
	−	fecundity-linked display costs; non-limiting resources*	Blount <i>et al.</i> [9]; Speed & Ruxton [15]
among species	+	go-slow prey sampling*; handicap via resource trade-off*; Batesian mimicry; correlated marginal costs; automimicry*	Blount <i>et al.</i> [9]; Franks <i>et al.</i> [14], Guilford [10], Speed & Ruxton [15]; Speed & Franks [11]; Holen & Svenningsen [12]
	−	fecundity-linked display costs; non-limiting resources*	Blount <i>et al.</i> [9]; Speed & Ruxton [15]
none	none	low or variable marginal costs; no shared signal/defence resources	Guilford & Dawkins [7]; Speed & Ruxton [15]

Holen & Svenningsen [12] as part of a broader set of evolutionarily stable strategy models, and affirmed the view that effective honesty may arise among individuals when resources are shared between the synthesis of defences and signals and, hence, trade-off during allocation. Plasticity [11], Batesian mimicry [14,19] and predator cognition [12] have been variously examined in alternative models, with several apparent routes to stable positive correlations between the level of defence and signal expression now identified (reviewed in [20]).

Despite mechanistic differences, the potential for quantitative honesty has found broad theoretical support (albeit not universal; table 1), and empirical efforts have chiefly focused on testing the existence and direction of predicted correlations. The results have been equivocal. Positive correlations, and hence quantitative honesty, are apparent among several species of coccinellid beetle, in which the levels of endogenously synthesized alkaloids—precoccinelline and/or coccinelline—covary with the conspicuousness of their bold, spotted signals [9,21,22]. By contrast, in the long-studied poison dart frogs, only negative correlations were identified between the potency of alkaloids and signal salience across species of *Ameregea* [23] and across populations of *Oophaga* [24]. An earlier study found no relationship at all [25], however, while more recent work in the group found moderate to strong positive correlations on the whole [26–28]. Analogous tests in wasps [29], moths [30], salamanders [31] and sea slugs [32] add to the accumulating evidence, though the central shared prediction of honesty-based models remains unresolved.

Inherent to the question of the existence of quantitative honesty is the scale at which signal/defence correlations should manifest. As above, much of the focus has been on the origins and maintenance of among-individual correlations, though similar potential exists at the scale of populations and even species [11,15,33]. Positive correlations among populations may arise, for example, if the production

costs of signals and defences scale in unison, or if the density of prey and the risk of predation vary across space [15,34]. If Batesian mimicry is at play then antagonistic coevolution may select for increasing conspicuousness among model animals across species or populations, to a similar end [14]. Holen & Svenningsen [12] too, although focused on among-individual honesty, note that the stable equilibria arising from 'go-slow' [10] and resource allocation trade-off [9] mechanisms likely extend to the scale of populations and species. By contrast, Leimar *et al.*'s [33] game-theoretical treatment showed that negative among-population correlations are possible when the strength of defences vary. Wider-ranging optimization models reveal similar flexibility in potential sign of cross-species correlations, with negative associations arising when marginal costs of defences or displays vary [15]. In empirical terms, this question has also received mixed support. Positive correlations across species [22,26,27] are met with negative counter-examples [23], while among populations each of positive [28], negative [24] and the absence of correlations [25] have been documented.

This expanse of empirical work on aposematic honesty is invaluable, but diversity in signal designs, chemical defences and the measurement of each is challenging for qualitative synthesis. Modern quantitative methods, however, provide robust tools for capitalizing on such variation, and so can offer substantive insight into longstanding questions [35,36]. To that end, we conducted a meta-analysis to examine whether and to what extent aposematic signals encode quantitatively honest information on chemical defences in prey. Specifically, we synthesized estimates of the correlation between the visual features of aposematic signals and chemical defences to test the presence and sign of the predicted covariation at three scales: among individuals, populations and species. The central shared prediction of honesty-based models is a positive correlation between the expression of signals and quality of chemical defences, with a negative or non-

existent relationship favouring an alternate set of candidate mechanisms governing the coevolution of aposematic signals and defences (table 1).

## 2. Methods

### (a) Literature search and study selection

We conducted a systematic search of the scientific literature using the *Web of Science* and *Scopus* databases for publications up to July 2020. Following initial scoping searches, we used the search string ((aposemati\* OR warning colo\* OR warning signal) AND (honest\* OR toxic\* OR chem\* OR handicap OR quantitative OR unpalatab\* OR unprofit\*)). In the *Web of Science* database, we also refined by the categories ecology, evolutionary biology, environmental sciences, zoology, biology and behavioural sciences. Our search returned 735 studies following the removal of duplicates, of which 24 were ultimately appropriate for quantitative synthesis, including three studies identified through references list searches (see electronic supplementary material, figure S1 for PRISMA diagram, and electronic supplementary material for further search details).

We included all publicly available studies (peer-reviewed, preprint or theses) that empirically examined the scaling relationship between a continuous measure of the quality, efficacy and/or quantity of chemical defences (internal or secreted), and aposematic signal appearance in organisms, or which contained sufficient data for such relationships to be estimated directly (electronic supplementary material, table S1). Included studies could be experimental or observational and could test the relationship at any scale: among individuals, populations or species. Among individuals, such correlations were typically estimated from individual-level measures of signal appearance and chemical defence (further detailed below) across a sample of prey organisms from a single population. This general procedure is true of studies at higher scales too, albeit with population- or species-level averages of signal appearance and chemical defence underlying the estimated correlation(s). We excluded commentaries and reviews, effects collected from known Batesian mimics, or measures derived from artificial prey. We contacted authors to request necessary data where it was not available ( $n = 1$ ) and excluded these studies if it could not be obtained.

### (b) Data extraction and effect size calculation

We used the correlation coefficient Pearson's  $r$ , transformed to Fisher's  $z$  (for its normalizing and variance-stabilizing qualities), as the effect size describing the relationship between aposematic signal coloration and prey defences for meta-analysis. These effects were extracted directly from the study text, converted from available test statistics or estimated from digitized figures using the R package 'metaDigitise' v.1.0 [37].

We also recorded information from each study, both theoretical and methodological, which we *a priori* hypothesized may moderate the strength of any relationship between colour and defences. This included the scale at which signal/defence correlations were estimated; among individuals, populations or species. We also examined the method used for quantifying chemical defences, since standard approaches fall into two categories. The first includes those studies that directly measure the concentration, quantity or diversity of specific compounds which invoke aversive responses in predators, such as alkaloids (e.g. [21,38,39]). The second general approach is to use bioassays in which proxy organisms are exposed to the relevant compounds or tissues containing them, such as skin extracts. Examples include exposing *Daphnia* to presumed toxins and recording mortality [22,40] or injecting mice with tissue extracts and recording irritation and/or mortality [23,28]. We, therefore,

coded each study dichotomously depending on which of the two general methods was used. Where necessary we also adjusted the directionality of measures so that they represented the same relationship (e.g. by converting *Daphnia* survival to mortality).

Rich methods are also available for quantifying the expression of colour signals, which vary in their intent [41,42]. Some centre on estimating the properties that describe absolute colour appearance; hue (the unique colour), saturation (spectral richness or purity), brightness (or 'lightness,' in human terms) and size (relative or absolute). While others attempt to capture the contextual conspicuousness of signals through measures of contrast. This may entail using visual models to first estimate the viewer-subjective appearance of colours, before measuring the difference between colours within a colour pattern (*internal* contrast), or the overall difference between colour patterns and natural viewing backgrounds (*external* contrast). In all cases, theory predicts positive covariation between signal measures and defence levels as a signature of quantitative 'honesty'. The one exception is signal 'hue', which is a directionless measure. That is, predictions of its scaling relationship with defences require mechanistic insight into colour and defence production, and knowledge of the environments in which it is viewed. We, therefore, took the absolute value of correlations between signal hue and defences throughout. A sensitivity analysis showed the only qualitative effect of this decision lay in the statistical significance of hue itself, which falls slightly below the threshold when considered on its original scale ( $r = 0.208$ , 95% CI =  $-0.04$ ,  $0.43$ ). Beyond signal hue, we coded all effects as quantifying one of saturation, brightness, size, internal contrast or external contrast.

### (c) Meta-analyses

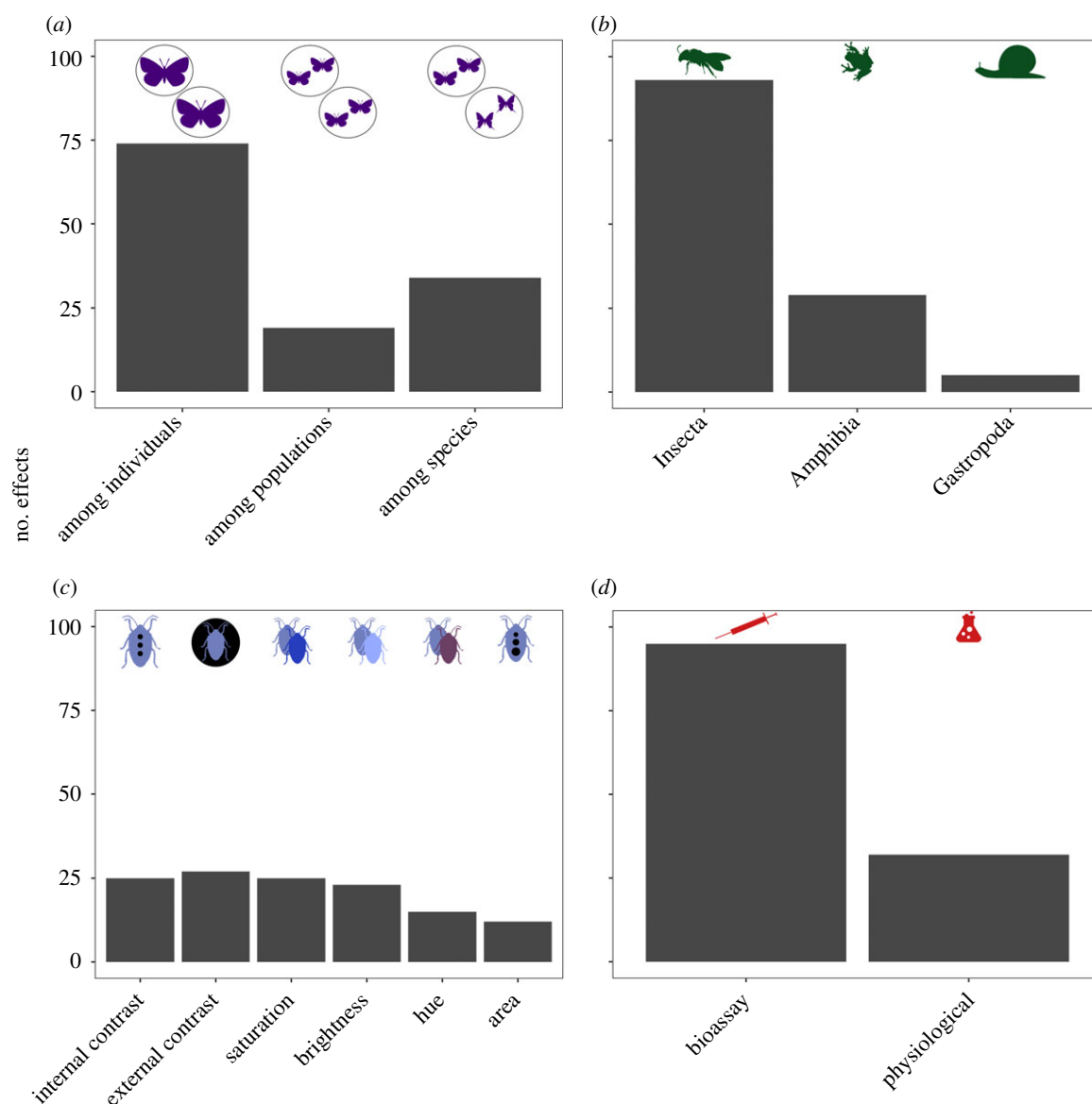
We ran both multi-level meta-analytic (intercept-only, MLM) and multi-level meta-regression (MLMR) models, using the package *metafor* v.2.1.0 [43] in R v.4.0.1 [44]. Almost all studies reported multiple effects so we included a study-level random effect in all models (following validation, see electronic supplementary material), as well as an observation-level effect to estimate residual variances which *metafor* does not report by default. Furthermore, multiple distinct measures—such as the hue, saturation and brightness of signals—are often collected from a single individual or group, and so are likely to be more correlated than those collected from independent groups. These correlations are almost never reported in original studies, and regularly go unconsidered in the meta-analysis, so we conservatively accounted for this possibility by fitting the off-diagonal covariances in the sampling covariance matrix assuming a correlation of  $r = 0.5$ . In all models, we consider effect size estimates whose 95% confidence intervals do not overlap zero to be statistically significant.

### (d) Publication bias

We formally tested for evidence of publication bias by visually inspecting a funnel plot of effect sizes versus standard errors for symmetry, and through an Egger's test on an intercept-only MLMA that included the full random effects structure described above [45].

## 3. Results

Our final analysis included 127 effect sizes from 24 studies [21–23,26–32,38–40,46–56]. Most studies reported multiple effects, with a mean of 6 and range of 1–27. Some 22 species were represented in total (considering only species-level studies), spanning three classes: insects, gastropods and amphibians (figure 1). With respect to the scale of the



**Figure 1.** The number of included effects by (a) scale, (b) taxa, (c) aspect of aposematic signal design and (d) measure of chemical defence. (Online version in colour.)

study, 74 effects (from seven studies) estimated signal/defence correlations among individuals, 19 (via four studies) were measures among populations and 34 (from 15 studies) were among species.

Across all studies, we identified a moderate positive correlation between aposematic colour signal expression and chemical defences (meta-analytic mean  $Z = 0.34$ , 95% CI = 0.19–0.47), as consistent with a prediction of quantitative honesty (figure 2). Effect size heterogeneity was moderate to high ( $I^2 = 0.70$ , 95% CI = 0.59–0.79) which is typical in ecology and evolutionary biology and suggests that variation among studies is not explicable by sampling variance alone. Notably, the positive correlation between signals and defences held across scales (table 2 for numerical results henceforth). That is, the correlation was moderate and positive irrespective of the scale at which the relationship was examined—among individuals, populations or species—though it is weakest, and our estimate least-precise, at the among-population scale in light of the relatively smaller sample size. All components of warning signals were similarly informative of prey defences, with each of signal hue, saturation, brightness and area, as well as internal and external contrast, scaling positively with the level of chemical defence.

These relationships also held irrespective of the way in which defences were quantified. Effects estimated from bioassay-based methods tended to be stronger than those using physiological measures of chemical quantity or diversity, but only slightly so. Finally, signal–defence correlations were similar in direction and magnitude among amphibians and insects though not gastropods (for whom the confidence interval included zero), which is unsurprising given the small sample of available effects for that group.

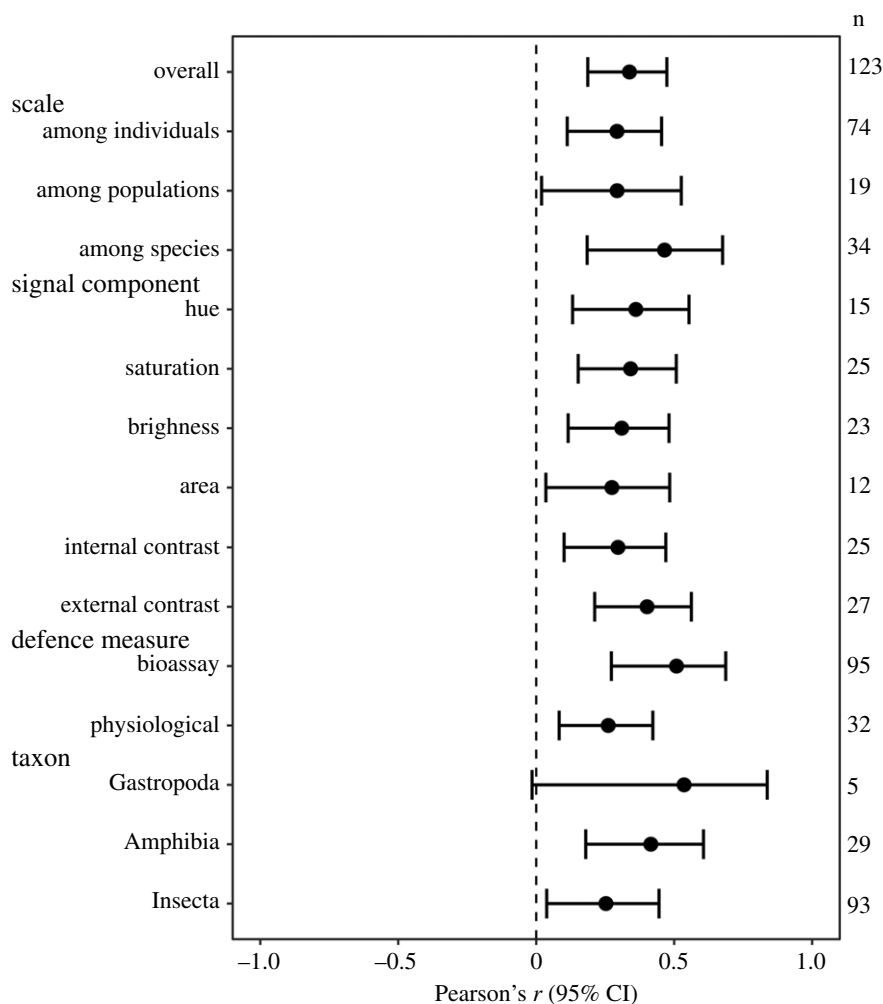
#### (a) Publication bias

Visual inspection of the funnel plot suggested no identifiable asymmetry and hence no evidence of publication bias (electronic supplementary material, figure S2), which was supported by an Egger's test (estimate 95% CI = −0.25–0.07,  $t_{125} = -0.03$ ,  $p = 0.98$ ).

## 4. Discussion

Conspicuous warning colours are known to advertise the presence of defences among prey [57,58]. The question of *quantitative honesty*—the positive scaling of signal expression





**Figure 2.** Forest plot denoting the meta-analytic mean effect, and moderators of the correlation between aposematic signal expression and chemical defences in prey. Shown are Pearson's correlations backtransformed from Fisher's  $z$ , with 95% confidence intervals about the mean. Sample sizes are displayed on the right. 'Scale' indicates the level at which correlations were estimated, while 'signal component' captures the aspect of aposematic signal appearance against which the level of defence was correlated. 'Defence measure' differentiates effects estimated via the use of bioassays versus direct measures of the quantity and/or diversity of defensive chemicals.

and defences—has remained contentious, however, with contrasting theoretical (table 1) and equivocal empirical [28,40,52] support to date. Here, we report meta-analytic evidence for a positive correlation between aposematic signal structure and prey defences, as consistent with a prediction of quantitative honesty. Notably, this correlation holds at all scales; among individuals, populations and species, though significant between-study heterogeneity remains unexplained. Both absolute signal features, such as saturation and brightness, and contextual measures of conspicuousness, were salient channels of information on the extent of defences, and the relationship held irrespective of the measure of chemical defence used.

These results contradict early theory [33] and support the central shared prediction of a positive signal–defence correlation among honesty-based models [9,12]. The strength and consistency of effects across scales (figure 2) argues for generality in the underlying mechanism, for which resource allocation trade-offs and 'go-slow' prey sampling are promising candidates (table 1; [9,12]). The former predicts the emergence of positive correlations across scales given a key assumption—competition for resources within individuals—is met [9,12]. The exact form this competition takes remains unknown, though antioxidant-mediated trade-offs

have been proposed [9], in an argument which echoes similar work on pigment-based sexual signals [59,60]. Coccinellid beetles present a promising example, among whom the maternal allocation of both pigments and precoccinelline may underlie adult signal honesty through the metabolic conversion of the latter *in ovo* [38,47]. Resource-dependent variation in signal/defence correlations [38] and the negative scaling of predation rates with conspicuousness [22] strengthen the case and offer the firmest line of evidence for resource mediated trade-offs as a driver of signal honesty, as broadly consistent with our results.

The 'go-slow' sampling of prey, by contrast, emphasizes predator cognition as a driving force in the evolution of aposematism. The central insight is that if conspicuous prey tend to be better defended, and predators are able to discriminate among prey based on the basis of chemical defences and colour, then predators should be more wary when sampling conspicuous prey [10,12]. The process of taste-rejection will thus lead to differential prey survival as a function of conspicuousness. And if the costs of increased conspicuousness are unequally borne across well- and poorly defended prey, then positive signal/defence correlations will result [10–12,18]. Dedicated tests of the hypothesis do not yet exist, though key underlying

**Table 2.** Full parameter estimates from MLM and MLMR models of the mediators of the correlation between aposematic colour signal expression and chemical defences. Shown are sample sizes, mean Pearson's  $r$  (backtransformed from Fisher's  $z$ ), lower and upper 95% confidence intervals and heterogeneity ( $I^2$ ).

	$r$	lower CI	upper CI	$n$	$I^2$
overall					0.70
intercept-only	0.34	0.19	0.47	123	
scale					0.70
among individuals	0.29	0.11	0.45	74	
among populations	0.29	0.02	0.53	19	
among species	0.47	0.18	0.68	34	
signal component					0.70
hue	0.36	0.13	0.55	15	
saturation	0.34	0.15	0.51	25	
brightness	0.31	0.12	0.48	23	
area	0.27	0.04	0.48	12	
internal contrast	0.30	0.10	0.47	25	
external contrast	0.40	0.21	0.56	27	
defence measure					0.69
bioassay	0.51	0.27	0.69	95	
physiological	0.26	0.08	0.42	32	
taxon					0.70
Amphibia	0.42	0.18	0.61	29	
Gastropoda	0.54	−0.02	0.84	5	
Insecta	0.25	0.04	0.45	93	

assumptions have found in-principle support. Common terrestrial [61–64] and marine [65] generalist predators are capable of taste-rejecting prey on the basis of their chemical defences, as well as discriminating among the visual features of signals [66–68]. Conspicuousness and defences have also been shown to interact to increase the probability of prey rejection [69,70], in part through a learned wariness toward conspicuous prey [71]. The taxonomic and ecological breadth of this work suggests ‘go-slow’ sampling as an appealing catch-all explanation, in keeping with the generality of our findings (table 2). Nonetheless, the empirical case for predator cognition driving quantitative honesty is underdeveloped and is an exciting area for attention.

A striking feature of extant work is that all visual properties of aposematic signals, both contextual (e.g. contrast) and absolute (e.g. hue, saturation), are, on average, reliable channels of information on prey defences. This is reflective of the diverse demands on signal efficacy, which include the need for salience within noisy viewing environments, and the government of cognitive processes such as generalization, categorization and memorization in viewers [72–75]. The former is defined by conspicuousness, both internal and external pattern contrast, which relates the structure of signals to their detectability amidst the desaturated hues of natural vistas [22,28,76,77]. The reliable memorization and categorization of stimuli, by contrast, is tied to the features of colour, brightness (albeit less often) and their spatial arrangement [76,78,79]. This functional partitioning of contextual and absolute signal features is similarly borne out in combined tests. Birds, for example, can learn to remember and discriminate among prey on the basis of colour, but

are consistently drawn to greater contrast irrespective of prior experience [74,80].

It is clear, then, that no signal feature is singularly informative of prey defences to predators. Note too, however, that this may partly be due to correlations among signal elements which are inherent to colour production. The chromaticity and brightness of patches tend to covary with contextual measures of external conspicuousness, since natural scenes are often dominated by the dull browns, greys and greens of earth and foliage [81]. Correlations exist among absolute features too, owing to the physics of light absorption by pigments and scattering by nano-structures [82]. The strength of these effects is likely limited, however, as suggested by the aforementioned evidence of selection on and attention to distinct signal components by viewers. Prey behaviour also works to degrade these links, as individuals actively modify their contextual conspicuousness independently of the ‘fixed’ visual features of their signals (e.g. [83,84]). More generally, our finding that all aspects of aposematic signals may reliably inform viewers of prey defences suggests there are myriad routes to effective advertisement. When considered alongside the stability of correlations at the scale of populations and species (which span a breadth of signal designs), these results also touch on the standing question of why aposematic signals are often variable. That is, there appears to be few fundamental design constraints on signals, at least with respect to their potential for honesty (figure 2), which is consistent with the emerging appreciation of diversity in aposematic signals as the norm, rather than a ‘paradox’ [85].

One note of caution is warranted when interpreting the strength of effects in biological, rather than statistical, terms. The visual appearance of signals and efficacy of chemical defences, and hence the honesty of advertisements, are ultimately subjective. Measures of each must, therefore, be calibrated against the perceptual experience of ecologically relevant receivers. With respect to signals, this is increasingly possible through the use of methods for capturing subjective visual experiences (reviewed in [41,86,87]) which continue to be applied and tested in the context of aposematism [22,53,67,78]. Progress in understanding the ecology of chemical defences as they relate to prey sampling and survival has been more gradual (though unabating, e.g. [88–90]), and some fundamental questions remain. The extent to which *Daphnia* spp. mortality or time-to-recovery in mice—common measures of prey ‘toxicity’—map to the responses of predators and, consequently, prey fitness, is untested. Similarly, our ability to infer predator perception and prey survival from the measured concentration of defensive chemicals is limited by the sheer complexity of the physiological and ecological pathways linking these processes (reviewed in [91]). Continued developments in these fields will be immensely valuable, then, not least for further grounding the signal–defence correlations here identified (figure 2) in ecological and evolutionary reality.

Aposematism is an exemplar of anti-predator defence, and its study has driven general advances in evolutionary ecology [1]. That warning signals appear quantitatively honest resolves a key problem in our understanding of their

function and evolution, though much remains to be learned. Knowledge of the physiological and/or ecological processes linking signal expression to defences is paramount, for which work on signal honesty, condition-dependence and oxidative stress in sexual systems may prove informative [92–95]. With respect to signal design, the targets of selection are only understood in general terms [96]. The relative contributions of chromatic (colour) and achromatic (brightness) conspicuousness to signal salience is unresolved [76,80], for example, as is the importance of within-pattern versus against-background contrast for viewer detection and cognition [67,80]. These are exciting areas for progress on a captivating feature of the natural world.

**Data accessibility.** All data and code underlying our analyses are available on GitHub ([https://github.com/EaSElab-18/ms\\_metawarn](https://github.com/EaSElab-18/ms_metawarn)) and have been persistently archived via Zenodo [97].

**Authors' contributions.** T.E.W.: conceptualization, data curation, formal analysis, methodology, project administration, visualization, writing original draft, review and editing; K.D.L.U.: conceptualization, formal analysis, methodology, writing original draft, review and editing. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

**Competing interests.** We declare we have no competing interests.

**Funding.** We received no funding for this study.

**Acknowledgements.** We appreciate the efforts of three anonymous reviewers and the associate editor, whose insights substantially improved the study. T.E.W. thanks Elizabeth Mulvenna and Cormac White for their enduring support. K.D.L.U. thanks Tim and Charlie Kahlke for support and grounding.

## References

- Ruxton GD, Sherratt TN, Speed MP. 2004. Avoiding attack: the evolutionary ecology of crypsis, warning signals and mimicry. Oxford, UK: Oxford University Press.
- Stevens M, Merilaita S. 2011. *Animal camouflage: mechanisms and function*. Cambridge, UK: Cambridge University Press.
- Umbers KDL, De Bona S, White TE, Lehtonen J, Mappes J, Endler JA. 2017 Deimatism: a neglected component of antipredator defence. *Biol. Lett.* **13**, 20160936. (doi:10.1098/rsbl.2016.0936)
- Rojas B, Valkonen J, Nokelainen O. 2015 Aposematism. *Curr. Biol.* **25**, R350–R351. (doi:10.1016/j.cub.2015.02.015)
- Maynard Smith J, Harper D. 2003 *Animal signals*. Oxford, UK: Oxford University Press.
- Searcy WA, Nowicki S. 2005 *The evolution of animal communication: reliability and deception in signaling systems*. Princeton, NJ: Princeton University Press.
- Guilford T, Dawkins MS. 1993 Are warning colors handicaps? *Evolution* **47**, 400–416. (doi:10.1111/j.1558-5646.1993.tb02102.x)
- Sherratt TN. 2002 The coevolution of warning signals. *Proc. Natl Acad. Sci. USA* **269**, 741–746. (doi:10.1098/rspb.2001.1944)
- Blount JD, Speed MP, Ruxton GD, Stephens PA. 2009 Warning displays may function as honest signals of toxicity. *Proc. R. Soc. B* **276**, 871–877. (doi:10.1098/rspb.2008.1407)
- Guilford T. 1994 ‘Go-slow’ signalling and the problem of automimicry. *J. Theor. Biol.* **170**, 311–316. (doi:10.1006/jtbi.1994.1192)
- Speed MP, Franks DW. 2014 Antagonistic evolution in an aposematic predator–prey signaling system. *Evolution* **68**, 2996–3007. (doi:10.1111/evo.12498)
- Holen ØH, Svenningsson TO. 2012 Aposematism and the handicap principle. *Am. Nat.* **180**, 629–641. (doi:10.1086/667890)
- Lee TJ, Speed MP, Stephens PA. 2011 Honest signaling and the uses of prey coloration. *Am. Nat.* **178**, E1–E9. (doi:10.1086/659997)
- Franks DW, Ruxton GD, Sherratt TN. 2009 Warning signals evolve to disengage batesian mimics. *Evolution* **63**, 256–267. (doi:10.1111/j.1558-5646.2008.00509.x)
- Speed MP, Ruxton GD. 2007 How bright and how nasty: explaining diversity in warning signal strength. *Evolution* **61**, 623–635. (doi:10.1111/j.1558-5646.2007.00054.x)
- Zahavi A. 1991 On the definition of sexual selection, Fisher's model, and the evolution of waste and of signals in general. *Anim. Behav.* **42**, 501–503. (doi:10.1016/S0003-3472(05)80052-1)
- Grafen A. 1990 Biological signals as handicaps. *J. Theor. Biol.* **144**, 517–546. (doi:10.1016/S0022-5193(05)80088-8)
- Speed MP, Ruxton GD, Blount JD, Stephens PA. 2010 Diversification of honest signals in a predator–prey system: honest signals in a predator–prey system. *Ecol. Lett.* **13**, 744–753. (doi:10.1111/j.1461-0248.2010.01469.x)
- Sherratt TN, Franks DW. 2005 Do unprofitable prey evolve traits that profitable prey find difficult to exploit? *Proc. R. Soc. B* **272**, 2441–2447. (doi:10.1098/rspb.2005.3229)
- Summers K, Speed MP, Blount JD, Stuckert AMM. 2015 Are aposematic signals honest? A review. *J. Evol. Biol.* **28**, 1583–1599. (doi:10.1111/jeb.12676)
- Bezzerides AL, McGraw KJ, Parker RS, Hussein J. 2007 Elytra color as a signal of chemical defense in the Asian ladybird beetle *Harmonia axyridis*. *Behav. Ecol. Sociobiol.* **61**, 1401–1408. (doi:10.1007/s00265-007-0371-9)
- Arenas LM, Walter D, Stevens M. 2015 Signal honesty and predation risk among a closely related group of aposematic species. *Sci. Rep.* **5**, 11021. (doi:10.1038/srep11021)
- Darst CR, Cummings ME, Cannatella DC. 2006 A mechanism for diversity in warning signals: conspicuousness versus toxicity in poison frogs. *Proc. Natl Acad. Sci. USA* **103**, 5852–5857. (doi:10.1073/pnas.0600625103)
- Wang JJ. 2011 Inversely related aposematic traits: reduced conspicuousness evolves with increased toxicity in a polymorphic poison-dart frog. *Evolution*

- 65, 1637–1649. (doi:10.1111/j.1558-5646.2011.01257.x)
25. Daly JW, Myers CW. 1967 Toxicity of Panamanian poison frogs (*Dendrobates*): some biological and chemical aspects. *Science* **156**, 970–973. (doi:10.1126/science.156.3777.970)
  26. Summers K, Clough ME. 2001 The evolution of coloration and toxicity in the poison frog family (*Dendrobatidae*). *Proc. Natl Acad. Sci. USA* **98**, 6227–6232. (doi:10.1073/pnas.101134898)
  27. Santos JC, Cannatella DC. 2011 Phenotypic integration emerges from aposematism and scale in poison frogs. *Proc. Natl Acad. Sci. USA* **108**, 6175–6180. (doi:10.1073/pnas.1010952108)
  28. Maan ME, Cummings ME. 2012 Poison frog colors are honest signals of toxicity, particularly for bird predators. *Am. Nat.* **179**, E1–E14. (doi:10.1086/663197)
  29. Vidal-Cordero J, Moreno-Rueda G, López-Orta A, Marfil-Daza C, Ros-Santaella JL, Ortiz-Sánchez F. 2012 Brighter-colored paper wasps (*Polistes dominula*) have larger poison glands. *Front. Zool.* **9**, 20. (doi:10.1186/1742-9994-9-20)
  30. Briolat ES, Zagrobelyny M, Olsen CE, Blount JD, Stevens M. 2019b No evidence of quantitative signal honesty across species of aposematic burnet moths (Lepidoptera: Zygaenidae). *J. Evol. Biol.* **32**, 31–48. (doi:10.1111/jeb.13389)
  31. Sanchez E, Pröhl H, Lüddecke T, Schulz S, Steinfartz S, Vences M. 2019 The conspicuous postmetamorphic coloration of fire salamanders, but not their toxicity, is affected by larval background albedo. *J. Exp. Zool. B: Mol. Dev. Evol.* **332**, 26–35.
  32. Cortesi F, Cheney KL. 2010 Conspicuousness is correlated with toxicity in marine opisthobranchs: aposematic signals in opisthobranchs. *J. Evol. Biol.* **23**, 1509–1518. (doi:10.1111/j.1420-9101.2010.02018.x)
  33. Leimar O, Enquist M, Sillen-Tullberg B. 1986 Evolutionary stability of aposematic coloration and prey unprofitability: a theoretical analysis. *Am. Nat.* **128**, 469–490. (doi:10.1086/284581)
  34. Ruxton GD, Speed MP, Broom M. 2009 Identifying the ecological conditions that select for intermediate levels of aposematic signalling. *Evol. Ecol.* **23**, 491–501. (doi:10.1007/s10682-008-9247-3)
  35. Koricheva J, Gurevitch J, Mengersen K. 2013 *Handbook of meta-analysis in ecology and evolution*. Princeton, NJ: Princeton University Press.
  36. Gurevitch J, Koricheva J, Nakagawa S, Stewart G. 2018 Meta-analysis and the science of research synthesis. *Nature* **555**, 175–182. (doi:10.1038/nature25753)
  37. Pick JL, Nakagawa S, Noble DW. 2019 Reproducible, flexible and high-throughput data extraction from primary literature: the metaDigitise R package. *Methods Ecol. Evol.* **10**, 426–431. (doi:10.1111/2041-210X.13118)
  38. Blount JD *et al.* 2012 How the ladybird got its spots: effects of resource limitation on the honesty of aposematic signals: warning signal honesty in a ladybird. *Funct. Ecol.* **26**, 334–342. (doi:10.1111/j.1365-2435.2012.01961.x)
  39. Crothers L *et al.* 2016 Warning signal properties covary with toxicity but not testosterone or aggregate carotenoids in a poison frog. *Evol. Ecol.* **30**, 601–621. (doi:10.1007/s10682-016-9830-y)
  40. Medina I, Wallenius T, Head M. 2020 No honesty in warning signals across life stages in an aposematic bug. *Evol. Ecol.* **34**, 59–72. (doi:10.1007/s10682-019-10025-0)
  41. Kemp DJ *et al.* 2015 An integrative framework for the appraisal of coloration in nature. *Am. Nat.* **185**, 705–724. (doi:10.1086/681021)
  42. Maia R, White TE. 2018 Comparing colors using visual models. *Behav. Ecol.* **29**, 649–659. (doi:10.1093/beheco/ary017)
  43. Viechtbauer W. 2010 Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36**, 1–48. (doi:10.18637/jss.v036.i03)
  44. R Core Team. 2018 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
  45. Egger M, Smith GD, Schneider M, Minder C. 1997 Bias in meta-analysis detected by a simple, graphical test. *BMJ* **315**, 629–634. (doi:10.1136/bmj.315.7109.629)
  46. Bonansea MI, Vaira M. 2012 Geographic and intrapopulational variation in colour and patterns of an aposematic toad, *Melanophryniscus rubriventris* (Amphibia, Anura, Bufonidae). *Amphib.-Reptil.* **33**, 11–24. (doi:10.1163/156853811X619754)
  47. Winters AE, Stevens M, Mitchell C, Blomberg SP, Blount JD. 2014 Maternal effects and warning signal honesty in eggs and offspring of an aposematic ladybird beetle. *Funct. Ecol.* **28**, 1187–1196. (doi:10.1111/1365-2435.12266)
  48. Wheeler CA, Millar JG, Cardé RT. 2015 Multimodal signal interactions in the ladybeetle, *Hippodamia convergens*, aposematic system. *Chemoecology* **25**, 123–133. (doi:10.1007/s00049-014-0181-2)
  49. Amézquita A, Ramos Ó, González MC, Rodríguez C, Medina I, Simões PI, Lima AP. 2017 Conspicuousness, color resemblance, and toxicity in geographically diverging mimicry: the pan-Amazonian frog *Allobates femoralis*. *Evolution* **71**, 1039–1050. (doi:10.1111/evo.13170)
  50. Briolat ES, Zagrobelyny M, Olsen CE, Blount JD, Stevens M. 2018 Sex differences but no evidence of quantitative honesty in the warning signals of six-spot burnet moths (*Zygaena Filipendulae* L.). *Evolution* **72**, 1460–1474. (doi:10.1111/evo.13505)
  51. Paul SC, Stevens M, Pell JK, Birkett MA, Blount JD. 2018 Parental phenotype not predator cues influence egg warning coloration and defence levels. *Anim. Behav.* **140**, 177–186. (doi:10.1016/j.anbehav.2018.04.014)
  52. Stuckert AMM, Saporito RA, Summers K. 2018 An empirical test indicates only qualitatively honest aposematic signaling within a population of vertebrates. *J. Herpetol.* **52**, 201–208. (doi:10.1670/17-047)
  53. Winters AE *et al.* 2018 Toxicity and taste: unequal chemical defences in a mimicry ring. *Proc. R. Soc. B* **285**, 20180457. (doi:10.1098/rspb.2018.0457)
  54. Preibler K, Gippner S, Lüddecke T, Krause ET, Schulz S, Vences M, Steinfartz S. 2019 More yellow more toxic? Sex rather than alkaloid content is correlated with yellow coloration in the fire salamander. *J. Zool.* **308**, 293–300. (doi:10.1111/jzo.12676)
  55. Reudler JH, Lindstedt C, Pakkanen H, Lehtinen I, Mappes J. 2015 Costs and benefits of plant allelochemicals in herbivore diet in a multi enemy world. *Oecologia* **179**, 1147–1158. (doi:10.1007/s00442-015-3425-0)
  56. Chouteau M, Dezeure J, Sherratt TN, Llaurens V, Joron M. 2019 Similar predator aversion for natural prey with diverse toxicity levels. *Anim. Behav.* **153**, 49–59. (doi:10.1016/j.anbehav.2019.04.017)
  57. Poulton EB. 1890 *The colours of animals*. London, UK: Paul, Trench, Trübner & Co.
  58. Cott HB. 1940 *Adaptive coloration in animals*. London, UK: Methuen & Co. Ltd.
  59. Krinsky NI. 1989 Antioxidant functions of carotenoids. *Free Radic. Biol. Med.* **7**, 617–635. (doi:10.1016/0891-5849(89)90143-3)
  60. Chew BP, Park JS. 2004 Carotenoid action on the immune response. *J. Nutr.* **134**, 257S–261S. (doi:10.1093/jn/134.1.257S)
  61. Gamberale-Stille G, Guilford T. 2004 Automimicry destabilizes aposematism: predator sample-and-reject behaviour may provide a solution. *Proc. R. Soc. B* **271**, 2621–2625. (doi:10.1098/rspb.2004.2893)
  62. Skelhorn J, Rowe C. 2006 Avian predators taste–reject aposematic prey on the basis of their chemical defence. *Biol. Lett.* **2**, 348–350. (doi:10.1098/rsbl.2006.0483)
  63. Skelhorn J, Rowe C. 2010 Birds learn to use distastefulness as a signal of toxicity. *Proc. R. Soc. B* **277**, 1729–1734. (doi:10.1098/rspb.2009.2092)
  64. Williams BL, Hanifin CT, Brodie ED, Brodie ED. 2010 Tetrodotoxin affects survival probability of rough-skinned newts (*Taricha granulosa*) faced with TTX-resistant garter snake predators (*Thamnophis sirtalis*). *Chemoecology* **20**, 285–290. (doi:10.1007/s00049-010-0057-z)
  65. Lindquist N, Hay ME. 1996 Palatability and chemical defense of marine invertebrate larvae. *Ecol. Monogr.* **66**, 431–450. (doi:10.2307/2963489)
  66. Winters AE, Green NF, Wilson NG, How MJ, Garson MJ, Marshall NJ, Cheney KL. 2017 Stabilizing selection on individual pattern elements of aposematic signals. *Proc. R. Soc. B* **284**, 20170926. (doi:10.1098/rspb.2017.0926)
  67. Green NF, Urquhart HH, van den Berg CP, Marshall NJ, Cheney KL. 2018 Pattern edges improve predator learning of aposematic signals. *Behav. Ecol.* **29**, 1481–1486.
  68. Rönkä K, De Pasqual C, Mappes J, Gordon S, Rojas B. 2018 Colour alone matters: no predator generalization among morphs of an aposematic moth. *Anim. Behav.* **135**, 153–163. (doi:10.1016/j.anbehav.2017.11.015)
  69. Marples NM, van Veelen W, Brakefield PM. 1994 The relative importance of colour, taste and smell in the protection of an aposematic insect *Coccinella septempunctata*. *Anim. Behav.* **48**, 967–974. (doi:10.1006/anbe.1994.1322)



70. Halpin CG, Rowe C. 2017 The effect of distastefulness and conspicuous coloration on the post-attack rejection behaviour of predators and survival of prey. *Biol. J. Linnean Soc.* **120**, 236–244.
71. Sillén-Tullberg B. 1985 Higher survival of an aposematic than of a cryptic form of a distasteful bug. *Oecologia* **67**, 411–415. (doi:10.1007/BF00384948)
72. Roper TJ, Wistow R. 1986 Aposematic colouration and avoidance learning in chicks. *Q. J. Exp. Psychol. B* **38**, 141–149.
73. Roper TJ, Redston S. 1987 Conspicuousness of distasteful prey affects the strength and durability of one-trial avoidance learning. *Anim. Behav.* **35**, 739–747. (doi:10.1016/S0003-3472(87)80110-0)
74. Osorio D, Jones CD, Vorobyev M. 1999 Accurate memory for colour but not pattern contrast in chicks. *Curr. Biol.* **9**, 199–202. (doi:10.1016/S0960-9822(99)80089-X)
75. White TE, Vogel-Ghibely N, Butterworth NJ. 2020 Flies exploit predictable perspectives and backgrounds to enhance iridescent signal salience and mating success. *Am. Nat.* **195**, 733–742. (doi:10.1086/707584)
76. Prudic KL, Skemp AK, Papaj DR. 2007 Aposematic coloration, luminance contrast, and the benefits of conspicuousness. *Behav. Ecol.* **18**, 41–46. (doi:10.1093/beheco/arl046)
77. White TE, Rojas B, Mappes J, Rautiala P, Kemp DJ. 2017 Colour and luminance contrasts predict the human detection of natural stimuli in complex visual environments. *Biol. Lett.* **13**, 20170375. (doi:10.1098/rsbl.2017.0375)
78. Newport C, Green NF, McClure EC, Osorio DC, Vorobyev M, Marshall NJ, Cheney KL. 2017 Fish use colour to learn compound visual signals. *Anim. Behav.* **125**, 93–100. (doi:10.1016/j.anbehav.2017.01.003)
79. Raška J, Štys P, Exnerová A. 2017 How variation in prey aposematic signals affects avoidance learning, generalization and memory of a salticid spider. *Anim. Behav.* **130**, 107–117. (doi:10.1016/j.anbehav.2017.06.012)
80. Gamberale-Stille G, Guilford T. 2003 Contrast versus colour in aposematic signals. *Anim. Behav.* **65**, 1021–1026. (doi:10.1006/anbe.2003.2098)
81. Bukovac Z, Shrestha M, Garcia JE, Burd M, Dorin A, Dyer AG. 2017 Why background colour matters to bees and flowers. *J. Comp. Physiol. A* **203**, 369–380. (doi:10.1007/s00359-017-1175-7)
82. Johnsen S. 2012 *The optics of life: a biologist's guide to light in nature*. Princeton, NJ: Princeton University Press.
83. Rojas B, Devillechabrolle J, Endler JA. 2014 Paradox lost: variable colour-pattern geometry is associated with differences in movement in aposematic frogs. *Biol. Lett.* **10**, 20140193. (doi:10.1098/rsbl.2014.0193)
84. White TE, Latty T. 2020 Flies improve the salience of iridescent sexual signals by orienting toward the sun. *Behav. Ecol.* **31**, 1401–1409. (doi:10.1093/beheco/araa098)
85. Briolat ES, Burdfield-Steel ER, Paul SC, Rönkä KH, Seymoure BM, Stankowich T, Stuckert AMM. 2019 Diversity in warning coloration: selective paradox or the norm? *Biol. Rev.* **94**, 388–414.
86. Renoult JP, Kelber A, Schaefer HM. 2017 Colour spaces in ecology and evolutionary biology. *Biol. Rev.* **92**, 292–315. (doi:10.1111/brv.12230)
87. Maia R, Gruson H, Endler JA, White TE. 2019. pavo 2: New tools for the spectral and spatial analysis of colour in R. *Methods Ecol. Evol.* **10**, 1097–1107. (doi:10.1111/2041-210X.13174)
88. Rojas B, Burdfield-Steel E, Pakkanen H, Suisto K, Maczka M, Schulz S, Mappes J. 2017 How to fight multiple enemies: target-specific chemical defences in an aposematic moth. *Proc. R. Soc. B* **284**, 20171424. (doi:10.1098/rspb.2017.1424)
89. Burdfield-Steel ER, Schneider JM, Mappes J, Dobler S. 2020 Testing the effectiveness of pyrazine defences against spiders. *Chemoecology* **30**, 139–146.
90. Hamalainen L, Mappes J, Thorogood R, Valkonen JK, Karttunen K, Salmi T, Rowland HM. 2020 Predators' consumption of unpalatable prey does not vary as a function of bitter taste perception. *Behav. Ecol.* **31**, 383–392. (doi:10.1093/beheco/arz199)
91. Rowland HM, Rockwell Parker M, Jiang P, Reed DR, Beauchamp GK. 2015 Comparative taste biology with special focus on birds and reptiles. In *Handbook of olfaction and gustation* (ed. RL Doty), pp. 957–982. Hoboken, NJ: John Wiley & Sons, Inc.
92. Cotton S, Fowler K, Pomiankowski A. 2004 Do sexual ornaments demonstrate heightened condition-dependent expression as predicted by the handicap hypothesis? *Proc. Biol. Sci.* **271**, 771–783. (doi:10.1098/rspb.2004.2688)
93. Emlen DJ, Warren IA, Johns A, Dworkin I, Lavine LC. 2012 A mechanism of extreme growth and reliable signaling in sexually selected ornaments and weapons. *Science* **337**, 860–864. (doi:10.1126/science.1224286)
94. Weaver RJ, Santos ES, Tucker AM, Wilson AE, Hill GE. 2018 Carotenoid metabolism strengthens the link between feather coloration and individual quality. *Nat. Commun.* **9**, 73. (doi:10.1038/s41467-017-02649-z)
95. White TE. 2020 Structural colours reflect individual quality: a meta-analysis. *Biol. Lett.* **16**, 20200001. (doi:10.1098/rsbl.2020.0001)
96. Endler JA, Mappes J. 2017 The current and future state of animal coloration research. *Phil. Trans. R. Soc. B* **372**, 20160352. (doi:10.1098/rstb.2016.0352)
97. White TE, Umbers KDL. 2021 Data for: Meta-analytic evidence for quantitative honesty in aposematic signals Zenodo. (doi:10.5281/zenodo.4671045)