

RESEARCH ARTICLE

Diversity and complexity of arthropod references in haiku

Andrew R. Deans ^{*}, Laura Porturas 

Frost Entomological Museum, The Pennsylvania State University, University Park, Pennsylvania, United States of America

 These authors contributed equally to this work.^{*} adeans@psu.edu OPEN ACCESS

Citation: Deans AR, Porturas L (2024) Diversity and complexity of arthropod references in haiku. PLoS ONE 19(4): e0298865. <https://doi.org/10.1371/journal.pone.0298865>

Editor: Christian Wegener, Biocenter, Universität Würzburg, GERMANY

Received: October 26, 2023

Accepted: February 1, 2024

Published: April 3, 2024

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0298865>

Copyright: © 2024 Deans, Porturas. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: We provide here those poems that are in the public domain and those we have permission to share. Uniform Resource Identifiers (URIs) and International Standard Book Numbers (ISBNs) are provided to the source publications, which allow anyone with

Abstract

Haiku are short poems, each composed of about 10 words, that typically describe moments in nature. People have written haiku since at least the 17th century, and the medium continues to be popular with poets, amateurs, educators, and students. Collectively, these poems represent an opportunity to understand which aspects of nature—e.g., which taxa and biological traits—resonate with humans and whether there are temporal trends in their representation or the emotions associated with these moments. We tested this potential using a mix of linguistic and biological methods, in analyses of nearly 4,000 haiku that reference arthropods. We documented the taxa and the life history traits represented in these poems and how they changed over time. We also analyzed the poems for emotion and tone. Our results reveal a mix of predictable trends and compelling surprises, each of which stand to potentially inform engagement strategies. At least 99 families of arthropods, in 28 orders, are represented in these haiku. The eight most commonly referenced taxa, from highest to lowest number of references, include: Lepidoptera, Hymenoptera, Diptera, Coleoptera, Araneae, Orthoptera, Hemiptera, and Odonata. Several common, conspicuous orders were never referenced, including Trichoptera, Plecoptera, and Megaloptera. The most commonly referenced traits relate to ecology (especially habitat, phenology, time of day), behavior (especially sound production), phenotype (especially color), and locomotion (especially flight). The least common traits in haiku relate to arthropod reproduction and physiology. Our analyses revealed few obvious temporal trends in the representations of taxa, biological traits, or emotion and tone. The broader implications of these results and possible future directions are discussed.

Introduction

Life abounds with ephemeral moments that evoke emotion, expose the profound, or otherwise effect inspiration. Imagine a dragonfly gliding effortlessly across a pond, a millipede slowly coiling in self defense, or a caterpillar smoothly descending from the canopy on a line of silk. Each of these seemingly simple events has the potential to awaken or reinforce one's connection to the natural world. People often capture these experiences in short poems, called *haiku*.

access to the Internet to reconstruct the full corpus. All of the digital sources for the poems analyzed in this study are open access. The data sets provided here do have all the data used in our statistical analyses and those data that were used to generate figures.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

A haiku conveys such a moment intentionally and objectively, usually in about ten words, leaving the reader to infer emotional and sensory elements of that moment [1].

The literary origins of haiku date to the 17th century in Japan [2], or even earlier [3], but the popularity of haiku has soared globally over the last few decades [4]. There are now dozens of societies, journals, and competitions dedicated to the craft. The medium is also employed in countless classrooms worldwide, in exercises that teach mindfulness, grammar, and effective communication, among other educational and therapeutic goals [5]. Traditionally, these poems focused on natural events, and each haiku included a word or phrase, referred to as *kigo* [6], that alluded to the season [2]. Consider this haiku, by Mastuo Bashō (1644–1694), who is broadly considered to be the original haiku master [7]:

Spring rain falling

The roof leaks,

Trickling down the wasps' nest.

“Spring rain” serves as the *kigo*. The reader also witnesses the wasps' nest dripping with rain and can imagine the unstated elements Bashō experiences in that moment.

Thousands of haiku are published annually, and certainly tens of thousands or more are written each year. Collectively, they offer opportunities to understand how and which organisms inspire people, which biological traits resonate with the human experience, and where openings may exist for more effective outreach regarding biodiversity and conservation [8, 9]. Given the long history of haiku, one could also examine how biological references in haiku have changed temporally. Poets may incorporate increasingly sophisticated references over time, for example, as shared knowledge of organismal biology increases. Alternatively, the diversity of biology represented in haiku may decrease over time, as people spend more time indoors or otherwise have fewer encounters with wildlife [10–12].

We aim here to explore this potential, by analyzing a set of haiku, broadly defined, using linguistic and biological approaches. Our focus is on haiku that reference insects and their relatives, including arachnids and myriapods (Arthropoda). Here, we address the following questions: (1) What arthropods and which aspects of their biology are represented in haiku? (2) Have arthropod references in haiku changed over time? (3) What can we glean from the language used in haiku, regarding public sentiment towards different arthropod taxa and their associated biological traits?

Materials and methods

Corpus assembly

We assembled a primary corpus of 3,894 relevant haiku, internationally sourced and written by 1,248 individual poets between and 1549 and 2022. These poems were acquired from more than 65 unique sources (>400 volumes), including haiku-focused websites, journal volumes, books, and poetry competitions. About one third of the poems ($n = 1129$) were sourced from a poetry contest specifically focused on insect-related haiku, the Hexapod Haiku Challenge (HHC; [13]). This source includes 236 haiku written by poets under the age of 13, which provides a small window into potential age-related differences in arthropod references. Only haiku that were written in English or had been translated into English were included in our analysis. A complete list of sources is available in [S1 Appendix](#). The haiku were extracted manually from each source and organized in a spreadsheet with relevant metadata, using Dublin

Core [14] terms: “contributor” (who added the poem to the data set), “creator” (poet’s name), “date” (date of publication or poet’s death year, if publication year unknown), “dateSubmitted” (timestamp for when the poem was added to the data set), and “source” (the resource from which the haiku was found). Additional headers for “poem” (the text of the poem itself, pasted without regard to formatting), “identifier” (unique string for reference), and “notes” (any comments made by the contributor) were also used. The primary corpus is available as [S2 File](#), which includes the poems we have permission to share and URIs to those haiku still under copyright.

Scoring biological complexity

A subset of haiku was compiled as a secondary corpus and scored for 69 variables related to biological complexity and the lowest-level taxon represented. Given the labor intensive nature of this process, we limited our scoring to 2,500 haiku [S1 File](#), representing each creator in the primary corpus and including 1,003 HHC poems. Detailed descriptions of our scoring method and variables are provided in supplementary document [S2 Appendix](#). A list of the variable categories and two scoring examples are provided below. To calculate a haiku’s complexity, we scored variables as present (+1) or absent (0) for each major category below and for each of its lower-level variables:

- **locomotion:** ambulatory, cursory, saltatory, natatory, skating, flight, in place
- **reproduction:** courtship, mating, oviposition, hatching, brood care
- **anatomy/phenotype:** legs, antenna, mouthparts, head, wings, genitalia, defensive structure, setae, color/pattern, extended phenotype
- **physiology:** digestion, molting, metamorphosis, vision, chemosensory, hearing, neural, thermoregulation, bioluminescence, phototaxis, biological sex
- **life stage:** egg, immature, pupa, imago (adult)
- **behavior (not reproduction):** mimicry, predation, prey, aposematism, parasitism, foraging, pollination, feeding, defense, sound production, sociality, migration, dispersal, grooming
- **ecology:** domestic, peridomestic, not domestic (wild), association, paleontological, phenological (seasonal), temporal (circadian)

As an example, consider this haiku from 1812, penned by Kobayashi Issa and translated by David G. Lanoue [15]:

the hairy bug
becomes a butterfly . . .
summer moon

We scored it as positive (+1) for setae (“hairy”), which is a variable of anatomy/phenotype (+1); for metamorphosis (“becomes”; +1), which is a variable of physiology (+1); for adult (+1) and larva (+1); for the explicit phenology (“summer”; +1) and allusive temporal (“moon”; +1) references, which are variables of ecology (+1). This matrix yielded a complexity score of 13.04 ($= 9/69 \times 100$) for this poem. Based on the word “butterfly”, the lowest-level taxon was recorded as Papilionoidea. Another poem by Issa, from 1806 [15], scored considerably lower in biological complexity:

while swatting a fly
today again . . .
the mountain temple bell

This poem vaguely references locomotion (people swat at flies that are either flying or in place; +1) and ecology (“mountain temple” hints at the fly’s habitat; +1), and it references life stage (imago; +1). This matrix yields a complexity score of 4.35 (= $3/69 \times 100$). Based on the word “fly”, we recorded the lowest-level taxon as Diptera.

Linguistic analyses

We analyzed the secondary corpus for word use, tone, and emotion using the Linguistic Inquiry and Word Count (LIWC) application (LIWC-22; [16]). LIWC-22 compares contents of the corpus with built-in word banks associated with each linguistic category. Output scores for the trait categories from this analysis were appended to our dataset. We focused on the following linguistic variables: positive tone (tone_pos), negative tone (tone_neg), emotion (emotion), positive emotion (emo_pos), negative emotion (emo_neg), and anxiety (emo_anx). These scores are the proportion of words in the haiku that match to the LIWC software word-banks for each of the different linguistic categories/variables. If a haiku has 9 words total and 3 of them match words in the tone_pos word bank, the haiku would have a score of (3/9 =) 0.333 for that variable. If it also has 1 word that matches words in the tone_neg word bank, it receives a (1/9 =) 0.111 for that variable.

Word clusters were identified using AntConc [17, 18] on the primary corpus, focusing especially on the most common arthropod words (with wildcards): butterfly*; spider*|web; ant|ants; firefl*; dragonfl*; crick*; cicad*; mosquito*; moth*; etc. More details regarding AntConc settings, variables, and early results can be found in supplementary document S3 Appendix. The AntConc analyses were used to highlight linguistic trends, references to each taxon over time, and reveal potentially overlooked connections between taxa and biological traits.

Taxon hierarchies and diversity

To more fully understand the taxonomic richness of our dataset and to explore broader taxonomic trends, we placed each reference of an arthropod (i.e., each lowest-level taxon) within a taxonomic hierarchy. Some haiku referenced more than one taxon, and each reference was separated as an independent record. Higher-level taxonomic names were obtained for referenced taxa using the “tax_name” function in the package “taxize” [19], in R Statistical Software version 4.2.1 [20]. We chose to query only the National Center for Biotechnology Information (NCBI) [21] database for this step. Deviations from this query include the following: (a) we use Anthophila to represent bees, rather than the broader taxon Apoidea, and (b) we treat Psocodea as an order, in line with our current understanding of insect classification.

To get a better sense of whether each taxon was over- or underrepresented in haiku, relative to its diversity and likelihood of human encounter, we compared the number of references to their alpha diversity (number of species; see references in Discussion) and the proportion of arthropod occurrences on iNaturalist [22]. iNaturalist is a Web-based application, where people can upload records, including photos and sound recordings, of organisms they encounter. Unlike other, similar resources, for example the photo sharing application Flickr (<https://flickr.com/>), iNaturalist uses a taxonomic hierarchy to organize their data, which allowed us to more readily estimate rates of encounter.

Recording haiku source

Haiku that were sourced from the Hexapod Haiku Challenge were categorized as “HH”, and the other haiku collected websites, journal volumes, books, and poetry competitions were categorized as “other”.

Statistics

A generalized linear model (family = quasipoisson) was used to examine whether time, haiku source, or # of authors was a predictor for taxon diversity (estimated by # of unique taxonomic orders/year) referenced in haiku. The analysis was restricted to years that had 10 or more haiku, and performed using the `glm()` function in base R (R version 4.2.1) [20].

A linear mixed effects model was used to examine whether time, haiku source, or # of authors was a predictor for biological complexity referenced in haiku. Time, haiku source, and # of authors were included as fixed effects, and taxonomic order was included as a random effect. The analysis was restricted to years that had 10 or more haiku, and performed using the function `lmer()` from the package “lme4: Linear Mixed-Effects Models using ‘Eigen’ and S4” [23].

Nonparametric multivariate analyses were used to determine whether any of the LIWC variables related to the tone & emotion of haiku in the dataset (`tone_pos`, `tone_neg`, `emotion`, `emo_pos`, `emo_neg`, `emo_anx`) were dependent on categorical groups of interest (e.g., haiku written about adult insects vs. immature insects, or domestic vs. outdoor arthropods). We used the function `nonpartest()` from the package “nrmv: Nonparametric Comparison of Multivariate Samples” [24]. When all four test statistics can be calculated (ANOVA-Type, Lawley-Hotelling Type, Bartlett-Nanda-Pillai-Type, Wilks’ Λ -Type), it is recommended the Wilks’ Λ -Type statistic be used [25]. For each comparison, permutation replications were set to 1000, and prior to analysis, haiku that referenced more than one of the categorical variables were removed (e.g., a haiku that referenced both adult insects and immature insects would be excluded) to help eliminate the possibility of response values being incorrectly applied to multiple categories.

To summarize the strength and direction of the relationships between LIWC variables and time (date), we calculated Spearman’s Rank correlation coefficient using the `cor.test()` function in base R (R version 4.2.1) [20].

The R packages “dplyr: A Grammar of Data Manipulation” [26] and “tidyr: Tidy Messy Data” [27] were both used to manipulate and organize data. the R package “ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics” [28] were used to generate some figures.

Results

Taxonomic trends

The 2,500 haiku in our secondary corpus included 2,611 references to arthropods, with reference to 24 of the 33 major taxonomic groups in Fig 1 (far right column). The referenced arthropods come from 28 different taxonomic orders and represent at least 99 taxonomic families. Some taxa are predictably well-represented (Fig 2). Lepidoptera, for example, which includes many large, colorful, and otherwise conspicuous species of butterfly and moth, had the most references ($n = 510$). Hymenoptera (ants, wasps, bees; $n = 347$), Diptera (mosquitoes, flies; $n = 314$), Coleoptera (fireflies, beetles; $n = 310$), and Araneae (spiders; $n = 220$) round out the top five most referenced taxa. Each of these taxa includes well over 50,000 named species worldwide, many of which are conspicuous and familiar.

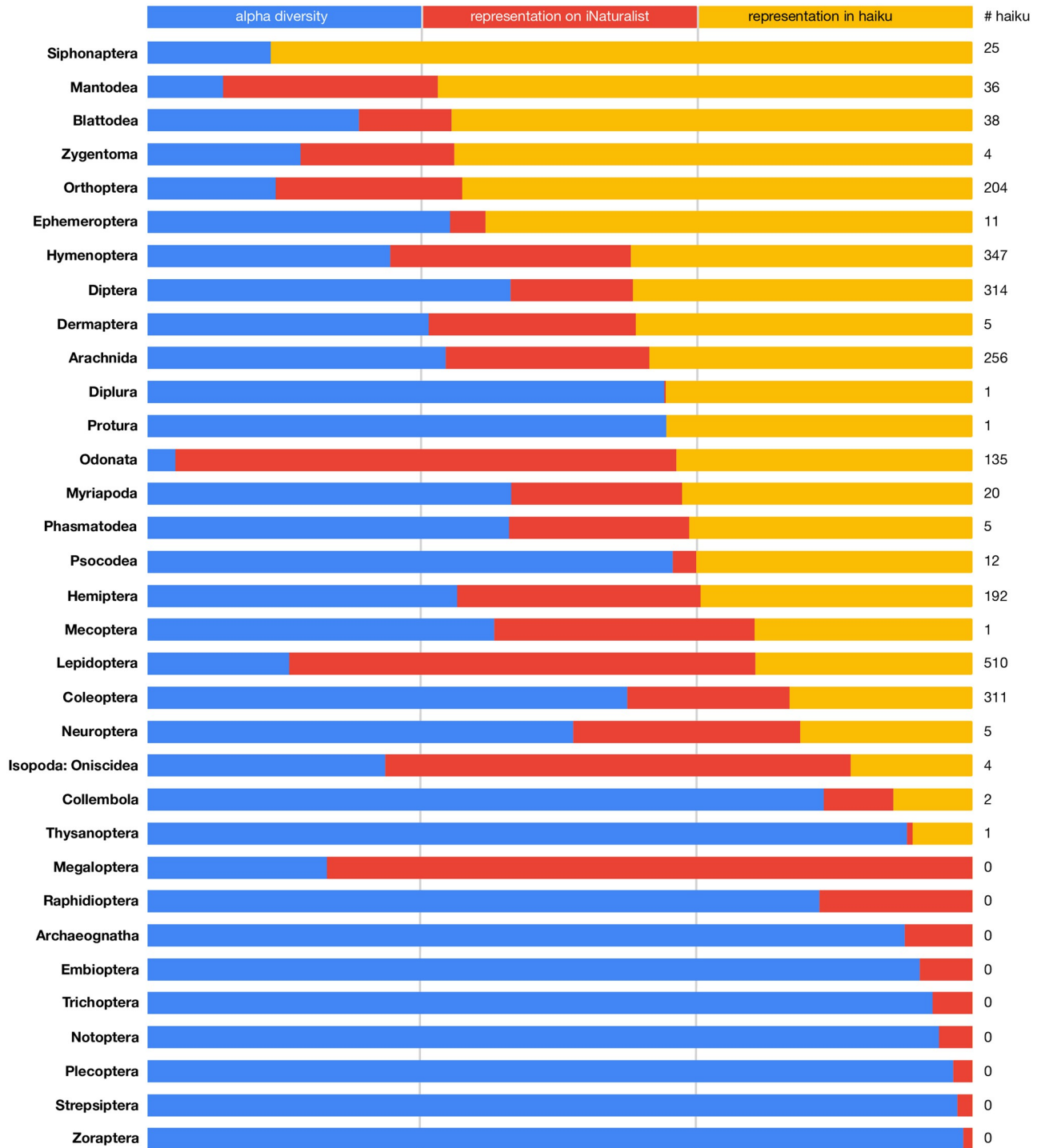


Fig 1. Relative arthropod representation. Bar size represents proportion in each category, relative to other categories. Blue = alpha diversity (number of species); red = representation on iNaturalist; yellow = representation in haiku. Three bars of equal size for a taxon would indicate that its representation is proportional across all categories (i.e., that the taxon comprises the same percentage in each category). See Discussion for references regarding the diversity of each taxon.

<https://doi.org/10.1371/journal.pone.0298865.g001>

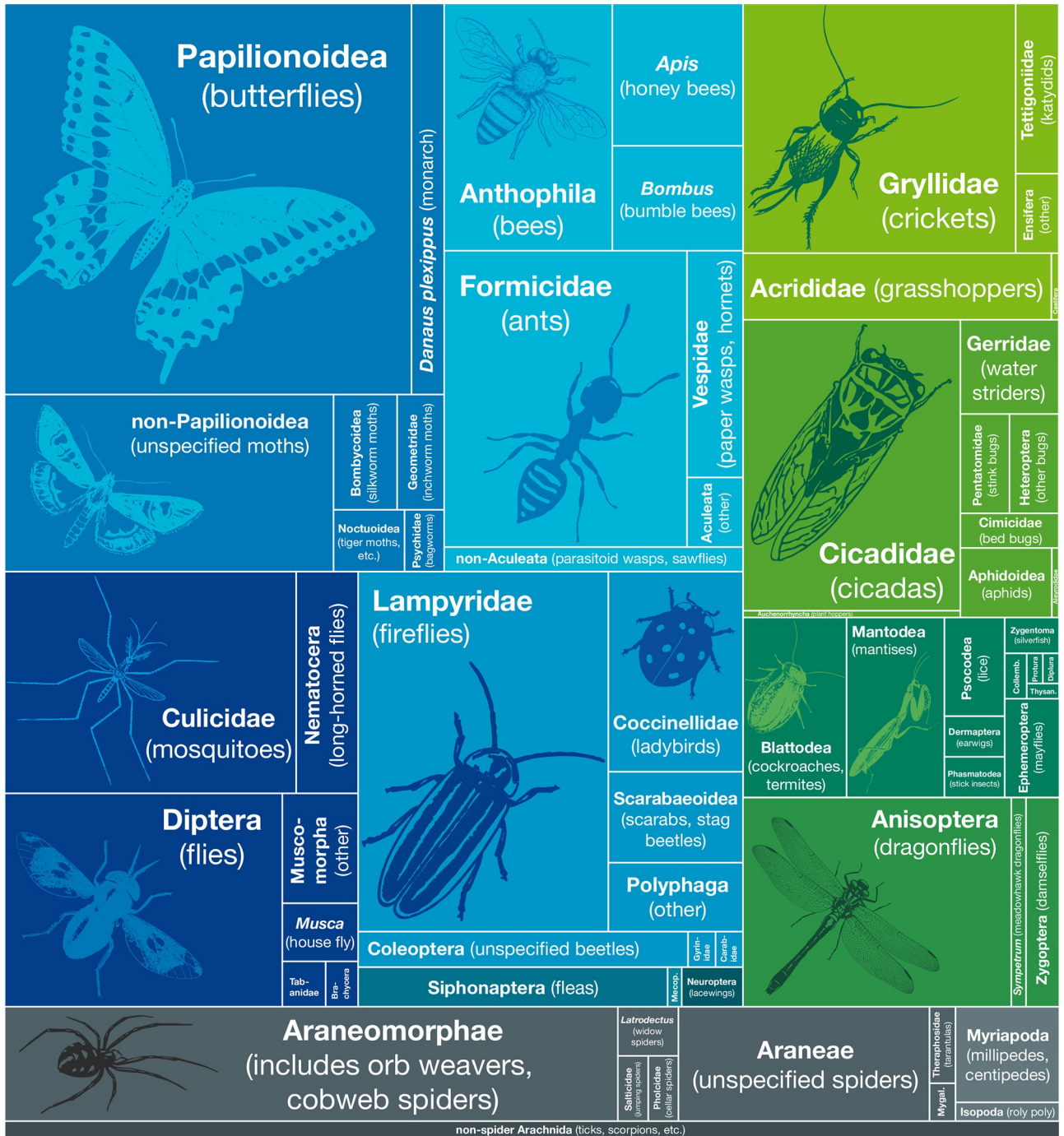


Fig 2. Taxon representation in haiku. Rectangle sizes are proportional and represent major taxonomic names referenced in haiku. The arrangement is roughly phylogenetic, with blue rectangles representing Holometabola, green rectangles non-holometabolous Hexapoda, and gray representing non-hexapod Arthropoda.

<https://doi.org/10.1371/journal.pone.0298865.g002>

Nine higher-level taxa were not referenced in any haiku (Fig 1, bottom rows), and four were only referenced once. Predictably, some of these taxa are depauperate and/or tend to have species of very small size. Zoraptera (zero references), for example, are small, cryptic insects with stringent habitat requirements [29] and <50 species. Protura (one reference) are even smaller

in size, generally requiring a microscope to observe. Supplementary [S1 File](#) contains the raw numbers for representation by taxon, which are summarized by [Fig 2](#) and [S1 Table](#).

As described above, we compared the taxonomic representation against alpha diversity and proportion of arthropod occurrences on iNaturalist, to get a better sense of whether each taxon was over- or underrepresented in haiku, relative to its diversity and likelihood of human encounter. Lepidoptera is the most frequently referenced taxon in our corpus, for example, accounting for 20.9% of arthropod references (yellow bar in [Fig 1](#)). Lepidoptera is overrepresented in haiku, relative to the number of species in the taxon (13.6% of all arthropod species; blue bar in [Fig 1](#)) but underrepresented relative to their frequency of encounter on iNaturalist (44.95% of the almost 14,000,000 arthropod observations recorded on that platform; red bar in [Fig 1](#)). The representation of Hemiptera in haiku (7.87% of references) closely resembles their species diversity (8.95% of all arthropod species) and likelihood of encounter (7.06% of arthropod occurrences on iNaturalist). Siphonaptera (fleas), on the other hand, are vastly overrepresented in haiku (1.02% of arthropod references), relative to their species-level diversity (2,075 species; 0.18% of arthropods) and frequency of encounter on iNaturalist ($n = 73$; 0.0005% of occurrences). These trends are discussed further in the taxon-specific summaries below, which also includes the results related to subordinal taxa.

Examining the relative proportions of taxa, binned by decade ([S1 Fig](#)) also reveals several notable trends. References to fleas (Siphonaptera), for example, are much more abundant in haiku prior to 1900, possibly because of improvements in pest control [30], hygiene related to livestock/pets, and increased urbanization [31]. References to bed bugs (Hemiptera: Cimicidae), however, are all contemporary (2010–2022), likely due to their recent resurgence globally [32, 33]. References to a third group of common, flightless parasites, sucking lice (Psocodea: Anoplura), are relatively evenly distributed across time. The relative consistency of taxonomic representation since 1950 ([S1 Fig](#)) is also notable. However, we cannot discount the fact that our uneven temporal sampling of haiku, or even mistranslation into English, may explain or at least influence some of these trends. Subordinal trends are presented and discussed in the taxon summaries below.

We determined that there were significant positive effects of both time and number of authors on the taxonomic diversity referenced in haiku ($t(33) = 4.09$, $p < .001$ and $t(33) = 3.63$, $p < .001$, respectively), and a significant negative effect of haiku source “other” ($t(33) = -2.23$, $p = 0.026$).

Biological trends

The insect biology referenced in these 2,500 haiku was similarly sophisticated yet uneven ([Fig 3](#)). The average haiku, composed of merely nine or ten words, had a complexity score of 7.52, meaning that just over five out of the 69 traits we scored (three major trait categories plus two more specific traits) were referenced. The poem with the highest complexity score (20.29), written by Gina Burns in 2010, referenced 14 different biological traits (9 specific traits in 5 major categories):

Tiny eggs on glass

Like jeweled beads for our eyes

Hatch, fly, eat, mate, lay

Seventeen haiku scored as zero for biological complexity, for example this haiku by John Soules [34].

theme: habitat (not domestic (wild), $n = 506$; domestic, $n = 369$; peridomestic, $n = 261$), time of day (temporal, $n = 416$), and time of year (phenology; $n = 354$). Other variables in the top 10 relate to life stage (adult, $n = 1,529$), locomotion (flying, $n = 398$; in place, $n = 227$), behavior (sound production, $n = 308$), and phenotype (color/pattern, $n = 230$; extended phenotype, $n = 194$). By contrast, only 112 haiku (<5%) refer to aspects of reproduction, the least popular category of biological references.

To examine temporal trends in the representation of different biological traits in haiku, we plotted the relative proportions of the top 20 biological traits (S3 Fig) and the major trait categories, binned by decade (S2 Fig). Unsurprisingly, the biological traits and the larger biological categories show similar trends through time. Of note, the relative proportions of different biological traits referenced remain fairly stable through time, with the exception of those decades with small sample sizes.

We found that neither time, number of authors, nor haiku source influenced the complexity of biological traits references ($t(332) = 1.90$, $p = 0.058$, $t(332) = 0.23$, $p = 0.819$, and $t(332) = 0.91$, $p = 0.36$, respectively).

Tone and emotional language trends

The following comparisons test whether the defined categories influence the language related tone and emotion in haiku.

Adult vs. immature arthropods. For this comparison, only the ANOVA-Type statistic could be computed, and we found no evidence that the language related to tone and emotion differs between the haiku referencing adult ($n = 1500$) vs. immature ($n = 111$) arthropods (ANOVA-Type statistic_{3,056,1252,642} = 1.043, $p = 0.373$, permutation $p = 0.369$).

Domestic vs. outdoor arthropods. Despite the fact that increasing urbanization and encounters with insects indoors are correlated with more negative opinion of insects [35, 36], we found no evidence that language differs between the haiku referencing domestic ($n = 367$), peridomestic ($n = 259$), or not-domestic ($n = 506$) (Wilks' Λ -Type statistic_{12,2248} = 1.119, $p = 0.340$; permutation $p = 0.082$).

Between taxonomic orders. Only orders with 10 or more references were included. Our results reveal there are differences in the language used between taxonomic orders (Araneae $n = 227$, Blattodea = 38, Coleoptera = 311, Diptera = 314, Ephemeroptera = 11, Hemiptera = 192, Hymenoptera = 347, Lepidoptera = 510, Mantodea = 35, Odonata = 135, Orthoptera = 204, Psocodea = 12, Scorpiones = 15, Siphonaptera = 25) (Wilks' Λ -Type statistic_{78,000,13001,710} = 2.901, $p < 0.001$; permutation $p < 0.001$). The relative effects (probability that a random sample from one group is higher than a random sample from another group) of taxon order on the LIWC variables indicating attitude were strongest in haiku referencing Scorpiones (scorpions); relative effects for negative tone = 0.80627, negative emotion = 0.73256, anxiety = 0.62604, and emotion = 0.70240. Haiku referencing scorpions also had a lower relative effect regarding positive tone = 0.43248. Haiku referencing Psocodea (lice) had more subtly elevated relative effects for negative tone = 0.58240, negative emotion = 0.55433, and emotion = 0.57172, but surprisingly also had an elevated effect for positive tone = 0.58218. Haiku referencing Ephemeroptera (mayflies) had a slightly elevated relative effect for positive tone = 0.55293, and a low relative effect regarding negative tone = 0.42361. Surprisingly, haiku referencing Mantodea (mantids) had an elevated negative tone = 0.56744. The remaining relative effects of taxon order on the LIWC variables ranged between 0.45 and 0.55, which are very close to 0.50, indicating little to no probability of differences in language that reflect attitude.

Flying vs. non-flying other types of locomotion. For this comparison, only the ANOVA-Type statistic could be computed, and we found no evidence that language use differs

between the haiku referencing flying insects ($n = 376$) vs. arthropods using other types of locomotion ($n = 404$) (ANOVA-Type statistic_{3,072,2386.679} = 0.958, $p = 0.413$; permutation $p = 0.41$).

Singing vs. non-singing. For this comparison, only the ANOVA-Type statistic could be computed, and we found evidence for differences in language used between haiku referencing singing insects ($n = 308$) vs. haiku that did not reference singing ($n = 2192$) (ANOVA-Type statistic_{2,92,3145.172} = 8.49, $p < 0.001$, permutation $p < 0.001$). However, the effects were very modest, and the haiku written about singing insects were slightly more likely to have both positive and negative language in the haiku (relative effects: positive tone = 0.53683, negative tone = 0.53377, negative emotion = 0.53457, and emotion = 0.53869).

Color vs. other. For this comparison, only the ANOVA-Type statistic could be computed, and we did find that there were significant differences between haiku referencing insect coloration ($n = 230$) vs. those that did not ($n = 2270$) (ANOVA-Type statistic_{2,938,2444.1} = 2.818, $p = 0.039$, permutation $p = 0.035$). The effects were very modest, with the haiku referencing color in insects containing only slightly more positive language in haiku than those that did not reference insect coloration (relative effect: positive tone = 0.52491). Research elsewhere suggests that colorful insects are less likely to elicit disgust in an observer, than their less colorful counterparts [37].

Emotion over time. To test whether there was detectable change in LIWC variables over time, we calculated Spearman's correlation coefficient. Our results reveal no correlations between these variables and time: positive tone ($\rho = 0.02086308$, $p = 0.2971$); negative tone ($\rho = 0.000325706$, $p = 0.987$); emotion ($\rho = -0.01276468$, $p = 0.5235$); positive emotion ($\rho = 0.005387698$, $p = 0.7877$); negative emotion ($\rho = -0.01772976$, $p = 0.3756$); anxiety ($\rho = 0.02020234$, $p = 0.3126$).

Discussion

Most commonly referenced taxa

Araneae (spiders). Spiders are frequently referenced in haiku (263 poems in the primary corpus, 220 in secondary corpus), but this is mostly a recent phenomenon. More than 98% of the references were published from 1960 onward. Taxonomically, there is a strong bias towards araneomorph spiders (72% of references in the secondary corpus), especially including those species that make orb and cob webs, when compared to the other major spider lineages: Mygalomorphae (3% of spider references) and Mesothelae (0%). About 25% of spider references could not be determined below Araneae. Active hunters, like the charismatic jumping spiders (Araneomorphae: Salticidae; 3% of spider references) and wolf spiders (Araneomorphae: Lycosidae; 1%) are hardly ever referenced. References of the unmistakable, diverse (1,047 spp.; [38]), and compelling tarantulas (Mygalomorphae: Theraphosidae), were scant in haiku (2% of spider references). The proportion of tarantula records in iNaturalist (<2% of spider records; [22]), however, is similarly sparse.

The most common biological attributes referenced in haiku about spiders relate to their extended phenotype, i.e., their webs ($n = 131$), and locomotion ($n = 75$). This result mirrors what we see in the AntConc analyses on the primary corpus. The top nine associated words from the cluster analysis of "spider", for example, are: web, webs, weaves, crawls, silk, jumping, moves, descends, and thread. Spiders were also more likely to be referenced as inside or around a home or building (domestic, $n = 70$; peridomestic, $n = 23$) than in nature (not domestic, $n = 33$), and they had the highest proportion of domestic references relative to the other arthropod orders analyzed. Interestingly, spiders comprise only 16% of all arthropod

taxa found indoors in a recent survey [39], and both abundance and diversity of spiders outdoors far exceeds those found indoors.

Odonata (dragonflies, damselflies). Odonata were referenced in 182 haiku (135 in the secondary corpus), and, unlike for spiders, the references reveal no obvious historical bias. Culturally, they have been popular insects since the 1600s. References to Anisoptera (dragonflies) abound, encompassing 89% of all Odonata references. Only ~5% of these references, however, included enough information to determine a taxon below suborder. Zygoptera (damselflies) makes up slightly more than half of all odonate species (3,218 of 6,330 described spp.; 50.8% [40]) and ~32% of odonate iNaturalist records [22] and yet comprise a mere 10% of Odonata references in the secondary corpus. Anisoptera on average are larger, faster, and otherwise more conspicuous, which may account for the bias in their representation.

Flight ($n = 54$) was the most common biological trait referenced in haiku about Odonata, followed by references to color ($n = 35$). The cluster analysis of the primary corpus likewise yields terms related to flying (hover, hovers, flying, lands) and color (red, blue). Odonata also had the highest proportion of wild habitat references (not domestic, $n = 48$ or 75% of the habitat references for this taxon). Given the conspicuous nature of adult dragonflies and damselflies, with their bright colors and adept flight, life stage references were dominated by adults (122 references to imago vs. one for immature and three for egg). Immature odonates (called “naiads” [41]) live underwater and are camouflaged with their surroundings. They are not as readily observed as adults.

Orthoptera (crickets, katydids, grasshoppers). “Cricket”, “katydid”, and “grasshopper” appear to be used interchangeably by translators of Japanese haiku [15, 42], which complicates our taxonomic assessment. Nevertheless, references to Orthoptera (308 haiku in the primary corpus; 204 in the secondary corpus) appear to be dominated by Ensifera, especially crickets (Gryllidae; 68% of orthopteran references). Ensifera accounts for ~58% of all described orthopteran species [43] and ~39% of orthopteran records on iNaturalist [22] but 79% of all Orthoptera references in the secondary corpus. Caelifera (grasshoppers and related orthopterans; ~42% of Orthoptera spp., ~60% of orthopteran records on iNaturalist) was represented by only 21% of the orthopteran references, most which were grasshoppers (Acrididea, probably mainly Acrididae). A recent survey of domestic environments found that the vast majority of home invaders are Ensifera (e.g., 100% of orthopteran home invaders in [39]), which may explain this bias in haiku. Ensifera also has more species that sing, compared to Caelifera, and these species tend to have louder calls than caeliferans [44]. Similar to Odonata, Orthoptera appear in haiku fairly consistently across time.

As expected, haiku that mention Orthoptera were dominated by references to sound production ($n = 125$, out of 204 haiku). Six of the top eight words revealed in the AntConc clusters include: chirp, chirps, song, cricky, sings, singing. These insects are well-known songsters and have been kept as companions by many cultures around the world [45, 46]. Other commonly referenced traits related to seasonality ($n = 42$) and time of day ($n = 65$). A defining feature of Orthoptera is their elongate, saltatorial hind legs, and yet we could only find eight references to jumping. Zero haiku referenced these insects’ fighting behaviors, which otherwise is well represented culturally [45].

Hemiptera (aphids, cicadas, true bugs). References to Hemiptera are fairly common and consistent across time in the primary corpus, with 274 references in haiku (192 in the secondary corpus). There is a bias towards Auchenorrhyncha (~68% of Hemiptera references), driven by references to Cicadidae (>98% of Auchenorrhyncha references; 67% of all Hemiptera references). Cicadas, which tend to be large, conspicuous, and gregarious insects that sing loudly, are common cultural icons [47]. The next two most commonly referenced insects are water striders (Heteroptera: Gerridae), which skate along the surface of bodies of water,

and stink bugs (Heteroptera: Pentatomidae), also referred to as “fart bugs” by translators. Despite the extraordinary diversity of Sternorrhyncha (scale insects, whiteflies, aphids, etc.), which comprises >18,000 spp. [48] and is represented by >53,000 records in iNaturalist [22], and which includes species that are common pests but also species that produce important products (dyes, lac, manna) [49], these insects made up a paltry 7% of Hemiptera references. Almost all Sternorrhyncha references were about aphids (Aphidoidea).

As expected, given the popularity of Cicadidae, biological references were dominated by sound production (n = 89), with AntConc clusters including terms like chirr, chirrs, sing, sings, chorus, cries, cry, singing, and song. Other commonly referenced traits include their natural habitat (not domestic, n = 50) and allusions to their seasonality (n = 39) and time of day (n = 39).

Hymenoptera (wasps, bees, ants). References to Hymenoptera were found in 366 haiku in the primary corpus (347 in the secondary corpus) and were dominated by those species capable of stinging (Aculeata). Aculeates accounted for 96% of the references. For comparison, Aculeata includes 65,258 described species (42.7% of all Hymenoptera; [50]) and an astounding 1,241,890 records on iNaturalist (88.7% of all Hymenoptera records on that site; [22]). Bees (Anthophila; 43% of hymenopteran references) and ants (Formicidae; 42% of hymenopteran references), which together comprise more than 29,000 species, or about 20% of described Hymenoptera [50], are the most commonly referenced aculeates. As with spiders, haiku about Hymenoptera are mostly a recent phenomenon; only ~6% of the references were published prior to 1959 (18/347 in secondary corpus). Given their ecosystem services (e.g., pollination), natural products (wax, honey), the long, global history of apiculture [51], the social nature of many species, and the prevalence of these insects in the environment (e.g., more than half the total biomass of terrestrial arthropods globally is ants [52, 53]) and in our homes [39], we expect to see many references to bees and ants in haiku. The paucity of references before 1959 is surprising.

Biological trait references for Hymenoptera mostly allude to foraging (n = 78) and related traits: pollination (n = 21), flying (n = 31), walking (n = 52). Top clusters from AntConc include words like scurry, walks, crawling, march, circles. Habitat was another commonly referenced trait: domestic (n = 31, primarily Formicidae), peridomestic (n = 47), not domestic (n = 71). Despite Aculeata accounting for 96% of the hymenopterans in haiku, poets made relatively few references to defense (n = 44) or to the sting itself (defensive structure, n = 5). Many aculeate species are social, and there were 74 references to sociality in the secondary corpus.

Coleoptera (beetles, fireflies). References to beetles (347 haiku in the primary corpus; 311 in secondary corpus) largely focused on species in Polyphaga (91%), especially fireflies (Lampyridae; 59% of Coleoptera references) and ladybird beetles (Coccinellidae; 18% of Coleoptera references). Given the cultural importance of these taxa [47, 54, 55], it is not surprising to see them represented frequently in haiku. However, Scarabaeoidea (e.g., scarabs, stag beetles) also includes many large, conspicuous, and culturally relevant beetles [56, 57]; yet they make up only about 9% of the beetle references in haiku. Another surprisingly underrepresented taxon is Cerambycidae, the longhorn beetles, which was referenced in only one haiku. These beetles are extraordinarily diverse (>30,000 species, ~8% of all beetle species; [58]) and are represented by 210,413 records on iNaturalist (13.1% of beetle records and the second most reported beetle family; [22]). Carabidae likewise was represented in only two haiku, despite having about 40,000 species (>10% of all beetle species; [58]) and 126,483 records in iNaturalist (~8% of all beetle records; [22]). References to Coleoptera across time are relatively consistent.

Bioluminescence (n = 111) was the most frequently referenced biological trait, which is unsurprising, given the popularity of fireflies in haiku. Other common traits relate to the behavior of flashing of fireflies, especially night references (temporal, n = 86), phenology

($n = 38$), and flight ($n = 55$). The top 20 AntConc clusters include terms related to locomotion (flits, flitting, come, flies), phenology (first), and bioluminescence (calling, lights).

Diptera (mosquitoes, midges, flies, maggots). We found 197 references in the primary corpus but 314 in the secondary corpus, where determinations were more confidently assigned. The word “fly” returned numerous irrelevant results in AntConc (flying, butterfly, firefly, etc.) Diptera references were fairly evenly distributed over time. Taxonomically, we see an high representation of mosquitoes (Culicidae; 42% of Diptera references), relative to other families. For context, mosquitoes comprise 3,725 species ($\sim 2.4\%$ of all Diptera; [59]) and 18,205 records in iNaturalist ($\sim 2.8\%$ of all Diptera records; [22]). Mosquitoes are abundant and often pestiferous in the adult stage, with many species capable of vectoring important human and veterinary diseases. Other biting flies, including horse flies (Tabanidae), black flies (Simuliidae), and tsetse flies (Glossinidae), together received as many references as the common house fly (Muscidae: *Musca*). At least 39% of Diptera references were not determinable below the order.

The polysemic nature of the word “fly” complicated our AntConc cluster analysis, but we did find many words associated with mosquito behavior (swarms, swarming, whine, whines, pesky) and cultural control (net, nets, and “smudge”, which refers to burning substances that produce a repellent smoke). Of the biological traits scored in the secondary corpus, ecological variables were by far the richest, especially habitat ($n = 167$), phenology ($n = 55$), and temporal ($n = 46$) references. Behavior was the second richest category, especially traits related to feeding (foraging, $n = 18$; parasitism, $n = 38$) and sound production ($n = 25$).

Lepidoptera (butterflies, moths). Lepidoptera are the most commonly referenced insects in haiku, with 498 references in the primary corpus and 510 references in the secondary corpus. About 68% of lepidopteran references are about butterflies (Papilionoidea). Monarchs, *Danaus plexippus* (Linnaeus, 1758), and cabbage whites, *Pieris rapae* (Linnaeus, 1758), were the most identifiable species in the secondary corpus. These species are abundant, well-known (e.g., see [60]), and widespread, with 156,878 and 61,613 records, respectively, on iNaturalist [22]. The relative commonness of bagworm references (Psychidae; 1,350 spp., $< 1\%$ of all Lepidoptera spp.; [61], compared to other small families, was surprising to us but mirrors references to this taxon in other cultural contexts [62–64]. Lepidoptera references were distributed fairly evenly across time.

As with most other arthropod references, the adult stage is the primary focus (imago, $n = 388$). Lepidoptera did have the most references to immature stages ($n = 92$), though, given the familiarity of caterpillars. The most frequently referenced traits overall were flight ($n = 161$), color/pattern ($n = 92$), and habitat (not domestic, $n = 121$, peridomestic, $n = 65$; domestic, $n = 40$). The top 27 AntConc clusters include terms related mainly to size (little), locomotion (flits, flitting, lands, dance, fly), color (white, blue, yellow, black), and to the type of lepidopteran (tiger, monarch, luna).

General impressions

Altogether, the references to biological traits mirror what is generally understood about children’s knowledge of insect biology and which aspects of insect biology we suspected would resonate with members of the general public. Children readily recall details about an insect’s ecology [65] and feeding behavior [66], for example, but have a poor understanding about how insects reproduce [67] and about arthropod anatomy beyond the presence of wings, legs, and an exoskeleton [68, 69]. We found no equivalent knowledge surveys for adults, who comprise the vast majority of poets in our data set, but it is possible that these trends hold true in the haiku-writing public, independent of age.

Regardless, our results provide insights into which arthropods and biological traits are notable (or not) to people. This knowledge provides entomologists and educators with target subjects for broader engagement. For example, aquatic arthropods provide an extraordinary array of ecosystem services [70], serve as critical bioindicators [71] and sources for bioinspiration [72], and they are commonly collected as bait. Haiku about these arthropods and their biology, however, are surprisingly rare. For example, not a single poem in our primary corpus mentioned Trichoptera (caddisflies, 14,999 species), Plecoptera (stoneflies, 3,788 species), nor Megaloptera (alderflies, dobsonflies, 354 species). A single poem referred to Chironomidae, which, with >7,300 species, is the most diverse family of aquatic insects [73]. These taxa represent a stunningly diverse array of compelling life history traits, and many of the species are common, conspicuous, and charismatic. The absence of aquatic arthropods in haiku may represent a gap in the public's knowledge of these organisms and hence an opportunity for novel engagement.

Alternatively, entomologists could use the results of these analyses to strengthen their outreach with more references to the known. Reinforcing programs and curricula with well-understood taxa and traits—butterflies and their colors, spiders and their webs, fireflies and bioluminescence, etc.—might allow for more efficient transfer of knowledge about arthropods.

We also recognize some limitations of our study. First, we intended our results to reveal, in part, something about the human emotions related to certain taxa and traits and how they may have changed over time. Our analyses of the secondary corpus, however, revealed very few obvious trends regarding emotion and tone, probably because haiku are short ($\bar{x} = 9.8176$ words) and intentionally subtle. The reader is meant to infer the emotion. Additionally, applying statistics to questions related to language is inherently challenging because language is essentially non-random [74]. However, our analyses were performed on the scores generated by LIWC analyses, rather than on the language itself. The results are accompanied by the relative effects when significance was found, which provide context for the magnitude of effect found from the statistical tests.

Second, our corpus is quite uneven in its representation of haiku across time (S4 Fig), poet type, and geographic region. Amateur writers, especially children, might be more likely to incorporate words that are represented in the LIWC-22 word banks (i.e., to use language that is more overtly emotional). The poems by children, however, are represented in our data set only from 2008–2012. Interestingly, the strongest emotional trend revealed in these poems relate to the negativity surrounding scorpions. Most of the scorpion haiku ($n = 14/15$) are from the Hexapod Haiku Challenge submissions, and they were largely written by children under the age of 13 ($n = 12/15$). Likewise, poets in certain geographic regions may limit their representation of insects to those taxa that are readily observed in those regions and/or those taxa that are traditionally included in the medium (i.e., literary inertia). For example, the poems in our corpus that were written prior to the 1920s are exclusively from Japan, whereas the poems from after that decade are increasingly international.

Conclusion

Computational lexicology [75] and related quantitative analyses of culture [76] and knowledge [77] continue to reveal insights that potentially inform education, outreach, and conservation. Poetry stands as a compelling medium to further prospect for these insights, given the deliberative word choice and frequent inclusion of nature. Haiku, which are poems that intentionally and frequently represent moments in nature, are especially relevant in this context. Our analyses of nearly 4,000 representative poems (more than 37,000 words) reveal many surprising results, articulated above, that have already catalyzed conversations locally about future

engagement activities. Studies of larger, more even bodies of text would undoubtedly allow for broader and more robust conclusions about how different arthropods are perceived by the general public and whether outreach programs can be further adapted for more effective engagement.

Supporting information

S1 Table. Arthropod representation in haiku. Columns represent (L to R): taxon, including common names; number of described species in each taxon [78, 79]; percent of total relevant arthropod species; number of occurrences for each taxon in iNaturalist [22]; percent of relevant arthropod occurrences in iNaturalist; number of references in the 2500 haiku corpus for each taxon; percent of all arthropod references in the corpus. Rows in **bold** had zero references in haiku. Rows highlighted in yellow are taxa with few species and no references in haiku. Rows in orange represent taxa that are surprisingly underrepresented.
(PDF)

S1 File. Secondary corpus with results. This spreadsheet contains all the haiku analyzed as the secondary corpus, with results from the trait scoring and the LIWC-22 analysis.
(XLSX)

S2 File. Primary corpus. This spreadsheet contains all the haiku analyzed as the primary corpus, which includes the haiku in the secondary corpus. This corpus was used with AntConc, to get lists of words associated with arthropods.
(XLSX)

S1 Fig. Taxonomic representation through time. This figure illustrates the proportion of haiku that reference any particular taxon at a particular period of time.
(TIF)

S2 Fig. Trait category representation through time. This figure illustrates the proportion of haiku that reference the major trait categories in each decade.
(TIF)

S3 Fig. Top 20 traits through time. This figure illustrates the proportion of haiku that reference the top 20 traits in each decade.
(TIF)

S4 Fig. Number of haiku in the primary corpus over time. This histogram illustrates the number of poems per year that are represented in the primary corpus.
(TIF)

S1 Appendix. Corpus construction. This document describes the strategies and resources used in constructing the corpora of haiku that reference arthropods. This document is also available through Penn State's institutional repository, ScholarSphere: <https://doi.org/10.26207/gpxg-q347>.
(PDF)

S2 Appendix. Biological variables and scoring methods. This document describes the process used to score haiku for biological complexity and taxon. This document is also available through Penn State's institutional repository, ScholarSphere: <https://doi.org/10.26207/04h1-a695>.
(PDF)

S3 Appendix. AntConc analyses. This document describes the cluster analyses and results using the AntConc application. This document is also available through Penn State's institutional repository, ScholarSphere: <https://doi.org/10.26207/35sv-ep54>. (PDF)

Acknowledgments

We thank Heather Froehlich (University of Arizona), for her invaluable assistance throughout the project; Anne Burgevin, for inspiring us to further study haiku; Hojun Song (Texas A&M University), for his input regarding Orthoptera diversity and biology; and David Lanoue (Xavier University) for allowing us to include his translations of Issa haiku in this research. We also appreciate the insights provided by our anonymous reviewers, comments and questions greatly improve the manuscript.

Author Contributions

Conceptualization: Andrew R. Deans, Laura Porturas.

Data curation: Andrew R. Deans, Laura Porturas.

Formal analysis: Andrew R. Deans, Laura Porturas.

Methodology: Andrew R. Deans.

Project administration: Andrew R. Deans.

Visualization: Andrew R. Deans, Laura Porturas.

Writing – original draft: Andrew R. Deans, Laura Porturas.

Writing – review & editing: Andrew R. Deans, Laura Porturas.

References

1. Gurga L. Toward an aesthetic for English-language haiku; 2000. Available at <http://web.archive.org/web/20160717194526/https://thehaikufoundation.org/omeka/files/original/e1c5a595ac45c6022a1530ef34ef1b49.pdf>. Accessed 8 November 2022.
2. Higginson WJ, Harter P. The haiku handbook: How to write, share, and teach haiku. Kodansha Amer Incorporated; 1989.
3. Kern AL. The rise and fall and spring of haiku. In: Leupp GP, min Tao D, editors. The Tokugawa World. 1st ed. Routledge; 2021. p. 781–798. Available from: <https://www.taylorfrancis.com/chapters/edit/10.4324/9781003198888-52/rise-fall-spring-haiku-adam-kern>.
4. Kern, AL. The Penguin Book of Haiku. Penguin UK; 2018.
5. Nguyen HN, Roth WM. An analysis of haiku teaching discourse: from talking about to doing haiku. Journal of Pedagogical Research. 2019; 3(3):113–125. <https://doi.org/10.33902/jpr.v3i3.93>
6. Yamamoto K, Higginson WJ. The Five Hundred Essential Japanese Season Words; 1992. Available at the Haiku Foundation Digital Library, <https://web.archive.org/web/20221205201607/https://www.thehaikufoundation.org/omeka/items/show/821>. Accessed 2022-12-23.
7. Reichhold J. Basho: the complete haiku. Kodansha America, Incorporated; 2013.
8. Ladle RJ, Correia RA, Do Y, Joo GJ, Malhado AC, Proulx R, et al. Conservation culturomics. Frontiers in Ecology and the Environment. 2016; 14(5):269–275. <https://doi.org/10.1002/fee.1260>
9. Luke SH, Roy HE, Thomas CD, Tilley LAN, Ward S, Watt A, et al. Grand challenges in entomology: Priorities for action in the coming decades. Insect Conservation and Diversity. 2023; 16(2):173–189. <https://doi.org/10.1111/icad.12637>
10. Deans AR. Cultural impacts of the biodiversity crisis and our disconnection from nature. Journal of Brief Ideas. 2022. <https://doi.org/10.5281/zenodo.6288546>
11. Langer L, Burghardt M, Borgards R, Böhning-Gaese K, Seppelt R, Wirth C. The rise and fall of biodiversity in literature: A comprehensive quantification of historical changes in the use of vernacular labels for

- biological taxa in Western creative literature. *People and Nature*. 2021; 3(5):1093–1109. <https://doi.org/10.1002/pan3.10256>
12. Kesebir S, Kesebir P. A growing disconnection from nature is evident in cultural products. *Perspectives on Psychological Science*. 2017; 12(2):258–269. <https://doi.org/10.1177/1745691616662473> PMID: 28346112
 13. Deans AR. Hexapod Haiku Challenge; 2022. Previously hosted by the author at the NC State University Insect Museum. Now hosted through the Frost Entomological Museum at The Pennsylvania State University. Available at <https://web.archive.org/web/20230104183851/https://sites.psu.edu/frost/hexapod-haiku-challenge/> Accessed 2022-12-19.
 14. DCMI Usage Board. DCMI Metadata Terms; 2020. Available at <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/> Accessed 2022-11-10.
 15. Lanoue DG. Haiku of Kobayashi Issa; 2022. Available at <https://web.archive.org/web/20220923170208/http://haikuguy.com/issa/> Accessed 2022-11-10.
 16. Pennebaker JW, Boyd RL, Booth RJ, Ashokkumar A, Francis ME. Linguistic Inquiry and Word Count (LIWC); 2022. Available at <https://www.liwc.app/> Accessed 2022-11-10.
 17. Anthony L. AntConc (Version 4.1.3) [Computer Software]; 2022. Tokyo, Japan: Waseda University. Available at <https://www.laurenceanthony.net/software> Accessed 2022-11-10.
 18. Froehlich H. Corpus analysis with AntConc. 2015. *Programming Historian*. 2015; 4. <https://doi.org/10.46430/phen0043>
 19. Chamberlain S, Szoecs E, Foster Z, Arendsee Z, Boettiger C, Ram K, et al. taxize: Taxonomic information from around the web; 2020. R package version 0.9.98, Accessed 16 August 2023. Available from: <https://github.com/ropensci/taxize>.
 20. R Core Team. R: A Language and Environment for Statistical Computing; 2021. Available from: <https://www.R-project.org/>.
 21. Schoch CL, Ciuffo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, et al. NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database*. 2020; 2020:baaa062. <https://doi.org/10.1093/database/baaa062> PMID: 32761142
 22. iNaturalist contributors, iNaturalist. iNaturalist Research-grade Observations. iNaturalist.org Occurrence dataset; 2022. Available at <https://doi.org/10.15468/ab3s5x> Accessed via GBIF.org on 2023-01-04.
 23. Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H, et al. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-35.1; 2023. Available from: <https://github.com/lme4/lme4/>.
 24. Burchett W, Ellis A. nrmv: nonparametric comparison of multivariate samples. R package version 2.4.0.; 2017. Available at <https://cran.r-project.org/package=nrmv> Accessed 2023-4-19.
 25. Arboretti R, Bathke A, Bonnini S, Bordignon P, Carrozzo E, Corain L, et al. 5. In: *Analyzing Survey Data Using Multivariate Rank-Based Inference*. Springer International Publishing; 2018. p. 71–84. Available from: https://doi.org/10.1007/978-3-319-91740-5_5.
 26. Wickham H, François R, Henry L, Müller K, Vaughn D. dplyr: A grammar of data manipulation; 2023. Available from: <http://dplyr.tidyverse.org/>.
 27. Wickham H, Vaughan D, Girlich M. tidyr: Tidy Messy Data; 2023. Available from <https://tidyr.tidyverse.org>.
 28. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York; 2016. Available from <https://ggplot2.tidyverse.org>.
 29. Wilson EO. Some ecological characteristics of ants in New Guinea rain forests. *Ecology*. 1959; 40(3):437–447. <https://doi.org/10.2307/1929761>
 30. Rust Michael K. Recent advancements in the control of cat fleas. *Insects*. 2020; 11(10):668. <https://doi.org/10.3390/insects11100668> PMID: 33003488
 31. Markkula M, Tiittanen K. Prevalence of bed bugs, cockroaches and human fleas in Finland. *Annales Entomologici Fennici*. 1970; 36(2):99–107.
 32. Romero A, Potter MF, Potter DA, Haynes KF. Insecticide resistance in the bed bug: A factor in the pest's sudden resurgence? *Journal of Medical Entomology*. 2007; 44(2): 175–178. <https://doi.org/10.1093/jmedent/44.2.175> PMID: 17427684
 33. Potter MF. Bed bugs through history In *Advances in the Biology and Management of Modern Bed Bugs* (eds S.L. Doggett, D.M. Miller and C.-Y. Lee); 2011. p. 9–25. Available from: <https://doi.org/10.1002/9781119171539.ch1>.
 34. Soules J. Mayfly Frogpond. *The Journal of the Haiku Society of America*. 2010; 33(3).
 35. Fukano Y, Soga M. Why do so many modern people hate insects? The urbanization–disgust hypothesis. *Science of The Total Environment*. 2021; 777:146229 (1–8). <https://doi.org/10.1016/j.scitotenv.2021.146229>

36. Schoelitsz B, Poortvliet PM, Takken W. Factors driving public tolerance levels and information-seeking behaviour concerning insects in the household environment. *Pest Management Science*. 2018; 74(6):1478–1493. <https://doi.org/10.1002/ps.4839> PMID: 29274106
37. Breuer GB, Schlegel J, Kauf P, Rupf R. The Importance of Being Colorful and Able to Fly: Interpretation and implications of children's statements on selected insects and other invertebrates. *International Journal of Science Education*. 2015; 37(16):2664–2687. <https://doi.org/10.1080/09500693.2015.1099171>
38. World Spider Catalog contributors. World Spider Catalog. Version 23.5. Natural History Museum Bern; 2023. <http://wsc.nmbe.ch> Accessed 2023-01-10.
39. Bertone MA, Leong M, Bayless KM, Malow TL, Dunn RR, Trautwein MD. Arthropods of the great indoors: characterizing diversity inside urban and suburban homes. *PeerJ*. 2016; 4:e1582. <https://doi.org/10.7717/peerj.1582> PMID: 26819844
40. Bybee SM, Kalkman VJ, Erickson RJ, Frandsen PB, Breinholt JW, Suvorov A, et al. Phylogeny and classification of Odonata using targeted genomics. *Molecular Phylogenetics and Evolution*. 2021; 160:107–115. <https://doi.org/10.1016/j.ympev.2021.107115> PMID: 33609713
41. Bybee SM, Hansen Q, Buesse S, Cahill Wightman HM, Branham MA. For consistency's sake: the precise use of larva, nymph and naiad within Insecta. *Systematic Entomology*. 2015; 40(4):667–670. <https://doi.org/10.1111/syen.12136>
42. Blyth R. *Haiku, Volume One*. Tokyo: Hokuseido Press; 1981.
43. Cigliano MM, Braun H, Eades DC, Otte D. Orthoptera Species File. Version 5.0/5.0; 2023. <http://Orthoptera.SpeciesFile.org> Accessed 2023-01-10.
44. Song H, Béthoux O, Shin S, Donath A, Letsch H, Liu S, et al. Author Correction: Phylogenomic analysis sheds light on the evolutionary pathways towards acoustic communication in Orthoptera. *Nature Communications*. 2020; 11(1):5674. <https://doi.org/10.1038/s41467-020-19626-8> PMID: 33144590
45. Neto EMC. Entertainment with insects: singing and fighting insects around the world. A brief review. *Etnobiología*. 2003; 3(1):20–28.
46. Bidau CJ. Patterns in Orthoptera biodiversity. II. The cultural dimension. *Journal of Insect Biodiversity* 2014; 2(21):1–15 Available at: <https://www.mapress.com/jib/article/view/2014.2.21>.
47. Hogue CL. Cultural entomology. *Annual Review of Entomology*. 1987; 32(1):181–199. <https://doi.org/10.1146/annurev.en.32.010187.001145>
48. Hardy NB. The biodiversity of Sternorrhyncha: scale insects, aphids, psyllids, and whiteflies. In: *Insect biodiversity: science and society*. 2nd ed. Wiley Online Library; 2018. p. 591–625. Available from: <https://doi.org/10.1002/9781118945582.ch20>.
49. Kondo T, Gullan PJ. Beneficial scale insects. In: Kondo T, Gullan PJ, editors. *Encyclopedia of Scale Insect Pests*. CAB International, Wallingford, U.K.; 2022. p. 1–7. Available from: <https://doi.org/10.1079/9781800620643.0001>.
50. Aguiar AP, Deans AR, Engel MS, Forshage M, Huber JT, Jennings JT, et al. Order Hymenoptera. In: Zhang, Z.-Q. (Ed.) *Animal biodiversity: an outline of higher-level classification and survey of taxonomic richness* (Addenda 2013). *Zootaxa*. 2013; 3703(1):51–62. <https://doi.org/10.11646/zootaxa.3703.1.12>
51. Morge G. Entomology in the western world in antiquity and in medieval times. In: Smith RF, Mittler TE, Smith CN, editors. *History of entomology*. Annual Reviews, Inc. Palo Alto, Calif.; 1973. p. 40–45.
52. Eggleton P. The State of the world's insects. *Annual Review of Environment and Resources*. 2020; 45(1):61–82. <https://doi.org/10.1146/annurev-environ-012420-050035>
53. Tuma J, Eggleton P, Fayle TM. Ant-termite interactions: an important but under-explored ecological linkage. *Biological Reviews*. 2020; 95(3):555–572. <https://doi.org/10.1111/brv.12577> PMID: 31876057
54. Lewis SM, Thancharoen A, Wong CH, López-Palafox T, Santos PV, Wu C, et al. Firefly tourism: Advancing a global phenomenon toward a brighter future. *Conservation Science and Practice*. 2021; 3(5):e391. <https://doi.org/10.1111/csp2.391>
55. Oba Y, Branham MA, Fukatsu T. The terrestrial bioluminescent animals of Japan. *Zoological Science*. 2011; 28(11):771–789. <https://doi.org/10.2108/zsj.28.771> PMID: 22035300
56. Tournant P, Joseph L, Goka K, Courchamp F. The rarity and overexploitation paradox: stag beetle collections in Japan. *Biodiversity and Conservation*. 2012; 21:1425–1440. <https://doi.org/10.1007/s10531-012-0253-y>
57. Ratcliffe BC. Scarab beetles in human culture. *The Coleopterists Bulletin*. 2006; 60(mo5):85–101. [https://doi.org/10.1649/0010-065X\(2006\)60%5B85:SBIHC%5D2.0.CO;2](https://doi.org/10.1649/0010-065X(2006)60%5B85:SBIHC%5D2.0.CO;2)
58. Bouchard P, Smith ABT, Douglas H, Gimmel ML, Brunke AJ, Kanda K. 11. In: *Biodiversity of Coleoptera*. John Wiley & Sons, Ltd; 2017. p. 337–417. Available from: <https://doi.org/10.1002/9781118945568.ch11>.

59. Courtney GW, Pape T, Skevington JH, Sinclair BJ. 9. In: Biodiversity of Diptera. John Wiley & Sons, Ltd; 2017. p. 229–278. Available from: <https://doi.org/10.1002/9781118945568.ch9>.
60. Vanderstock A, Grandi-Nagashiro C, Kudo G, Latty T, Nakamura S, White TE, et al. For the love of insects: gardening grows positive emotions (biophilia) towards invertebrates. *Journal of Insect Conservation*. 2022; 26(5):751–762. <https://doi.org/10.1007/s10841-022-00419-x>
61. van Nieuwerkerken EJ, Kaila L, Kitching IJ, Kristensen NP, Lees DC, Minet J, et al. Order Lepidoptera Linnaeus, 1758. In: Zhang, Z.-Q.(Ed.) *Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness*. Zootaxa. 2011; 3148(1):212–221. <https://doi.org/10.11646/zootaxa.3148.1.41>
62. Schmidt-Jeffris RA, Nelson J. Gotta catch'em all! Communicating entomology with Pokémon. *American Entomologist*. 2018; 64(3):159–164. <https://doi.org/10.1093/ae/tmy048>
63. Kawahara Akito Y Thirty-foot telescopic nets, bug-collecting video games, and beetle pets: Entomology in modern Japan. *American Entomologist*. 2007; 53(3):160–172. <https://doi.org/10.1093/ae/53.3.160>
64. van Huis, A Cultural significance of Lepidoptera in sub-Saharan Africa. *Journal of Ethnobiology and Ethnomedicine* 2019; 15:26. <https://doi.org/10.1186/s13002-019-0306-3>
65. Shepardson DP. Bugs, butterflies, and spiders: Children's understandings about insects. *International Journal of Science Education*. 2002; 24(6):627–643. <https://doi.org/10.1080/09500690110074765>
66. Shepardson DP. Of butterflies and beetles: First graders' ways of seeing and talking about insect life cycles. *Journal of Research in Science Teaching*. 1997; 34(9):873–889. [https://doi.org/10.1002/\(SICI\)1098-2736\(199711\)34:9%3C873::AID-TEA3%3E3.0.CO;2-R](https://doi.org/10.1002/(SICI)1098-2736(199711)34:9%3C873::AID-TEA3%3E3.0.CO;2-R)
67. Barrow LH. What do elementary students know about insects? *Journal of Elementary Science Education*. 2002; 14(2):53–60. <https://doi.org/10.1007/BF03173848>
68. Kiliç S. Students' knowledge about the internal structure of mice and cockroaches in their environment. *Educational Research and Reviews*. 2013; 8(19):1797. <https://doi.org/10.5897/ERR2013.1572>
69. Sieg AK, Teibtner R, Dreesmann D. Don't know much about bumblebees?—A study about secondary school students' knowledge and attitude shows educational demand. *Insects*. 2018; 9(2). <https://doi.org/10.3390/insects9020040>
70. Suter GW, Cormier SM. Why care about aquatic insects: Uses, benefits, and services. *Integrated Environmental Assessment and Management*. 2015; 11(2):188–194. <https://doi.org/10.1002/ieam.1600> PMID: 25376941
71. Chowdhury S, Dubey VK, Choudhury S, Das A, Jeengar D, Sujatha B, et al. Insects as bioindicator: A hidden gem for environmental monitoring. *Frontiers in Environmental Science*. 2023; 11:273. <https://doi.org/10.3389/fenvs.2023.1146052>
72. Stewart RJ, Ransom TC, Hlady V. Natural underwater adhesives. *Journal of Polymer Science Part B: Polymer Physics*. 2011; 49(11):757–771. <https://doi.org/10.1002/polb.22256> PMID: 21643511
73. Courtney GW, Cranston PS. Chapter 40—Order Diptera. In: Thorp JH, Rogers DC, editors. *Thorp and Covich's Freshwater Invertebrates (Fourth Edition)*. fourth edition ed. Boston: Academic Press; 2015. p. 1043–1058.
74. Kilgarriff A Language is never, ever, ever, random. *Corpus Linguistics and Linguistic Theory*. 2005; 1(2):263–276. <https://doi.org/10.1515/cllt.2005.1.2.263>
75. Michel JB, Shen YK, Aiden AP, Veres A, Gray MK, Team TGB, et al. Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*. 2011; 331(6014):176–182. <https://doi.org/10.1126/science.1199644> PMID: 21163965
76. Kitamura S. Quantifying the Influence of Climate on Human Mind and Culture: Evidence from Visual Art. *arXiv preprint arXiv:230715540*. 2023;.
77. Cruz SM, Grozinger CM. Mapping student understanding of bees: Implications for pollinator conservation. *Conservation Science and Practice*. 2023; 5(3):e12902. <https://doi.org/10.1111/csp2.12902>
78. Zhang ZQ. Phylum Arthropoda von Siebold, 1848 In: Zhang, Z.-Q.(Ed.) *Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness*. Zootaxa. 2011; 3148(1):99–103. <https://doi.org/10.11646/zootaxa.3148.1.14>
79. Sfendourakis S, Taiti S. Patterns of taxonomic diversity among terrestrial isopods. *ZooKeys*. 2015; 515:13–25. <https://doi.org/10.3897/zookeys.515.9332>