The role of non-English-language science in informing national biodiversity assessments

Tatsuya Amano[1, 2], Violeta Berdejo-Espinola[1, 2], Munemitsu Akasaka[3, 4], Milton A.U. de Andrade Junior[5], Ndayizeye Blaise[6], Julia Checco[7], F. Gözde Çilingir[8, 9], Geoffroy Citegetse[10], Marina Corella Tor[1, 2], Szymon M. Drobniak[11, 12], Sylvaine Giakoumi[1, 13], Marina Golivets[14], Mihaela C. Ion[15], Javiera P. Jara-Díaz[16], Ryosuke Katayose[17], Felicia P.S. Lasmana[18], Hsien-Yung Lin[19], Erick Lopez[20], Peter Mikula[21, 22], Lucia Morales-Barquero[23], Anne-Christine Mupepele[24], Juan P. Narváez-Gómez[25, 26], Thi Hieu Nguyen[27], Sá Nogueira Lisboa[28, 29], Martin A. Nuñez[30, 31], Diego Pavón-Jordán[32], Patrice Pottier[33], Graham W. Prescott[34, 35], Farah Samad[36], Marko Šćiban[37], Hae-Min Seo[38], Yushin Shinoda[3, 39], Flóra Vajna[40], Svetlana Vozykova[41], Jessica C. Walsh[42], Alison K.S. Wee[43], Hui Xiao[44, 45], Veronica Zamora-Gutierrez[46]

[1] School of Biological Sciences, The University of Queensland, Brisbane, Queensland, 4072, Australia.

[2] Centre for Biodiversity and Conservation Science, The University of Queensland, Brisbane, Queensland, 4072, Australia.

[3] Institute of Agriculture, Tokyo University of Agriculture and Technology, Fuchu, Tokyo, Japan, 183-8509, Japan.

[4] Institute of Global Innovation Research, Tokyo University of Agriculture and Technology, Fuchu, Tokyo, Japan, 183-8509, Japan.

[5] State Department of Sustainable Economic Development, Florianópolis, Santa Catarina, 88032-005, Brazil.

[6] Faculty of Environnemental Sciences, Gitega's Polytechnique University (UPG), Kayanza, Kirema, Burundi.

[7] School of Agriculture and Food Sciences, The University of Queensland, Brisbane, Queensland, 4072, Australia.

[8] Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

[9] Eidg. Forschungsanstalt WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland.

[10] BirdLife International, Mermoz Pyrotechnie, Lot 23 Rue MZ 56, Dakar, Senegal.

[11] School of Biological, Environmental and Earth Sciences, University of New South Wales, Sydney, New South Wales, 2052, Australia.

[12] Institute of Environmental Sciences, Jagiellonian University, Kraków, 30-387, Poland.
[13] Stazione Zoologica Anton Dohrn, Sicily Marine Centre, Lungomare Cristoforo Colombo, 90149 Palermo, Italy.
[14] Helmholtz Centre for Environmental Research - UFZ, Halle, 06120, Germany.
[15] Institute of Biology Bucharest, Romanian Academy, Splaiul Independentei 296, Bucharest, 60031, Romania.
[16] Fundación Auilaff, Puerto Varas, Region de Los Lagos, 5550000, Chile.
[17] Graduate School of Agriculture,Tokyo University of Agriculture and Technology, Fuchu, Tokyo, Japan, 183-8509, Japan.
[18] HCV Network, C/O Critchleys Llp, Beaver House, 23-38 Hythe Bridge Street, Oxford, OX1 2EP, United Kingdom.
[19] Institute of Environmental and Interdisciplinary Science and Department of Biology, Carleton University, Ottawa, Ontario, K1S 5B6, Canada.
[20] Centro de Estudios Ambientales y Biodiversidad, Universidad del Valle de Guatemala, 11 calle 15-79 zona 15 V.H. III, 1015, Guatemala.
[21] Institute of Vertebrate Biology, Czech Academy of Sciences, Květná 8, Brno, 603 65, Czech Republic.
[22] Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, 165 00 Prague, Czech Republic.
[23] Joint Remote Sensing Research Program, School of Earth and Environmental Sciences, The University of Queensland, Brisbane, Queensland, 4067, Australia.
[24] Chair of Nature Conservation and Landscape Ecology, University of Freiburg, Tennenbacherstr. 4, Freiburg, 79115, Germany.
[25] Departamento de Botânica, Instituto de Biociências, Universidade de São Paulo, Rua do Matão nº 277, Cidade Universitária, São Paulo, SP, CEP 05508-090, Brasil.
[26] Forest Ecology and Conservation Group, Conservation Research Institute and Department of Plant Sciences, University of Cambridge, Cambridge, CB2 3EA, UK.
[27] School of Business, The University of Queensland, Brisbane, Queensland, 4072, Australia.
[28] Eduardo Mondlane University, Faculty of Agronomy and Forest Engineering, 3453 Julius Nyerere Av., Maputo, 257, Mozambique.
[29] N'Lab, Maputo 77, Mozambique.
[30] Department of Biology and Biochemistry, University of Houston, 3455 Cullen Blvd #342, 77204,

USA.
[31] Grupo de Ecología de Invasiones, INIBIOMA, CONICET, Universidad Nacional del Comahue, Quintral 1250, San Carlos de Bariloche, 8400, Argentina.
[32] Department of Terrestrial Ecology, Norwegian Institute for Nature Research (NINA), P.O. BOX 5685, Torgarden, 7485, Trondheim, Norway.
[33] Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, The University of New South Wales, Sydney, New South Wales, 2052, Australia.
[34] Institute of Plant Sciences, University of Bern, Altenbergrain 21, 3013, Switzerland.
[35] The Biodiversity Consultancy Ltd, 3E King's Parade, Cambridge, CB2 1SJ, UK.
[36] Saint-Joseph University, Lebanon, Rue de Damas, Mar Mikhael, Beyrouth 1104 2020, 17-5208, Lebanon.
[37] Bird Protection and Study Society of Serbia, Partizanskih baza 6/43, Novi Sad, 21000, Serbia.
[38] Department of Agriculture, Forestry, and Bioresources, College of Agriculture and Life Sciences, Seoul National University, 200-7213, Gwanak-ro 1, Gwanak-gu, Seoul, 08826, Republic of Korea.
[39] Center for Environmental Biology and Ecosystem Studies, National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan, 305-8506, Japan.
[40] Lendület Ecosystem Services Research Group, Institute of Ecology and Botany, Centre for Ecological Research, Alkotmány út 2-4, Vácrátót, 2163, Hungary.
[41] Faculty of Energy and Ecotechnology (GreenTech), ITMO University, Kronverkskiy Prospekt 49, St Petersburg, 197101, Russia.
[42] School of Biological Sciences, Monash University, 25 Rainforest Walk, Monash University, Clayton, Victoria, 3800, Australia.
[43] School of Environmental and Geographical Sciences, University of Nottingham Malaysia, Jalan Broga, Semenyih, Selangor Darul Ehsan, 43500, Malaysia.
[44] School of Life & Environmental Sciences, Deakin University, 221 Burwood Hwy, Burwood Victoria, 3125, Australia.
[45] The Commonwealth Scientific and Industrial Research Organisation Oceans and Atmosphere, 306 Carmody Road, St Lucia, Queensland 4067, Australia.
[46] CONACYT - Centro Interdisciplinario de Investigación para el Desarrollo Integral Regional Unidad Durango (CIIDIR), Instituto Politécnico Nacional, Sigma 119, Colonia 20 de Noviembre II, Durango, Durango, 34220, Mexico.

Abstract

Consulting the best available evidence is key to successful conservation decision-making. While much scientific evidence on conservation continues to be published in non-English languages, a poor understanding of how non-English language science contributes to conservation decision-making is causing global assessments and studies to practically ignore non-English-language literature. By investigating the use of scientific literature in biodiversity assessment reports across 37 countries/territories, we uncover the established role of non-English-language literature as a major information source locally. On average, non-English-language literature constituted 65% of the references cited, and were recognised as relevant knowledge sources by 75% of report authors. This means that by ignoring non-English-language science, international assessments may overlook important information on local/regional biodiversity. A quarter of the authors acknowledged the struggles of understanding English-language literature. This points to the need to aid the use of English-language literature in domestic decision-making, for example, by providing non-English-language abstracts or improving/implementing machine translation.

Main text

Our ability to tackle global challenges effectively relies on a solid scientific evidence base[1]. Poor uptake of scientific evidence could cause biased and inefficient decisions, potentially leading to ineffective, and even negative, outcomes[2]. Conservation communities, for example, now explicitly recognise the importance of evidence-based decision-making, with Target 21 of the Kunming-Montreal Global Biodiversity Framework proposed by the Convention on Biological Diversity (CBD) aiming to ensure that the best available knowledge guides decision-making for the effective management of biodiversity[3]. We thus urgently need to understand what hinders and facilitates the uptake of scientific evidence in decision-making, in order to better inform practices and policies for addressing global challenges including the ongoing biodiversity crisis.

A number of barriers and enablers have been identified to affect the extent to which scientific evidence is used in environmental decision-making[4], yet there is an important driver that has almost completely been overlooked to date—language barriers. Today non-native English speakers, as well as native English speakers, routinely publish their scientific findings in English. This tendency often hinders access to the latest and relevant scientific evidence for decision-makers whose first language is not English. For example, 54% of protected area directors in Spain identified language (i.e., relevant

128 scientific knowledge being written in English) as a barrier to the use of scientific knowledge in their
129 management[5] while 12% of Swiss conservation professionals also reported language as a reason for
130 not reading academic journals[6]. In contrast, scientific knowledge that is available in a local, non-
131 English language is not only more readily accessible to decision-makers with lower English
132 proficiency, but could also provide locally-relevant evidence, such as knowledge on the ecology and
133 conservation of species and ecosystems in countries where English is not widely spoken[7,8]. Such non-
134 English-language scientific knowledge could be essential for informing environmental decision-
135 making, as biodiversity hotspots, where rich biodiversity is severely threatened, are largely found in
136 regions where English is not widely spoken[9]. In such regions, important scientific knowledge on
137 conservation is also produced by practitioners, who often find it difficult to publish their work in
138 English if their first language is not English and thus may decide to publish it in a non-English
139 language[5].

140 Earlier studies have rarely examined how scientific knowledge that is available in different
141 languages is being used in environmental decision-making, and what drives decision-makers to use or
142 not to use scientific knowledge in English and non-English languages. One exception is a recent study
143 showing that 96.6% of the references cited in global and regional biodiversity assessments by the
144 Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) were in
145 English[10]. This indicates that scientific literature published in non-English languages, which
146 constitutes up to one-third of the existing scientific literature on conservation[5], is hugely underused at
147 the international level. Contrary to this, given that language barriers can impede the use of English-
148 language literature, and much important knowledge is made available in non-English languages,
149 English-language literature would not dominate information sources for national biodiversity
150 assessments in countries where English is not widely spoken. Instead, we expect that scientific
151 literature made available in non-English languages is dominant and well recognised as locally-relevant
152 and readily-accessible information sources in such national assessments. The reliance on non-English-
153 language literature might be especially high in countries with lower English proficiency, where
154 science is more often communicated in a non-English language, and in countries with lower economic
155 development, where both producers and users of scientific information may be unable to afford access
156 to sufficient English-language literature and education.

157 This study investigates the contribution of scientific literature that is available in different
158 languages in informing national biodiversity assessments. We focused on national-level policy reports

on the state of biodiversity, as they serve as the fundamental basis for evaluating past, and shaping future, conservation actions and policies in each country while also informing regional and global assessments. We identified relevant biodiversity assessment reports in 37 countries/territories where the official language is not English, and investigated the proportion of English- and non-English-language references cited in those reports. We further conducted a questionnaire survey with the authors or editors of those reports (see **Methods** for more detail) to identify the barriers and enablers affecting the use of references in English- and non-English languages.

Results

We identified a total of 333 eligible reports on biodiversity conservation in 37 countries/territories where English is not an official language. These countries spanned across all four regions defined by the IPBES (Africa, Americas, Asia-Pacific, and Europe-Central Asia) [11] and represented 22% of the 166 countries/territories where English is not an official language[12]. As we found only one eligible report in seven out of the 37 countries, for consistency, we selected the one most relevant report in each country/territory based on pre-defined criteria (i.e., 37 reports in total) for investigating the use of references written in different languages (see **Methods** for more details). Most selected reports were about the status of biodiversity, or the environment (including biodiversity) in general, at the national level, but others included national reports to the CBD, national biodiversity strategies and action plans (Supplementary Data 1). The sensitivity of our conclusions to the choice of reports included in the analysis was minimal (see Supplementary Discussion and Extended Data Fig. 7).

Use of scientific references in different languages

For each of the 37 selected report we then recorded the number of references cited for each of the following four categories: (i) English-language traditional academic literature (i.e., peer-reviewed journal papers and books, hereafter "English-language academic literature"), (ii) English-language grey literature (i.e., all other literature types not controlled by commercial publishing, such as governmental reports, websites, databases, theses, etc.), (iii) non-English-language traditional academic literature (hereafter "non-English-language academic literature"), and (iv) non-English-language grey literature.

Non-English-language literature (academic and grey literature combined) represented a major source of scientific information in national biodiversity assessments in most of the 37

countries/territories covered in this study (Extended Data Fig. 1). On average, 65% of the references cited were written in a non-English language (red solid vertical line in Fig. 1a). Non-English-language literature represented over half of the references cited in reports for 28 (76%) countries/territories and over 75% in 15 (41%) countries/territories (Fig. 1a). These were in stark contrast to non-English-language literature representing only 3.4% of the references cited in the IPBES assessment (red broken vertical line in Fig. 1a, based on[10]). The proportion of non-English-language references cited in the reports was significantly higher in countries with a lower English Proficiency Index (a measure of the average English proficiency in each country[13], see **Methods** for more details) (Fig. 1b and Extended Data Table 1) and in countries with a lower gross domestic product (GDP) per capita (as a measure of economic level in each country, Fig. 1c and Extended Data Table 1).

A considerable proportion of the non-English-language literature cited was grey literature and when focusing only on academic literature, 44% of the academic literature cited in those reports were, on average, written in non-English languages. The proportion of non-English-language academic literature cited in the reports was again significantly higher in countries with a lower GDP per capita (Extended Data Fig. 2 and Extended Data Table 1). Some of the countries with a high English Proficiency Index and GDP per capita, such as those in Central and Western Europe, cited a very low proportion (i.e., less than 10%) of non-English-language academic literature (Extended Data Figs. 1 and 2).

Reasons for citing English/non-English-language references

Next, we investigated the barriers and enablers affecting the use of references written in different languages, by contacting at least one author or editor (hereafter "report author") of each report who played a leading role in compiling their reports (as the corresponding author or chief editor in most cases; see **Methods** for the sampling strategy). Their answers to questions in the survey (apart from the questions asking information on authors themselves, such as their first language(s)) are thus expected to represent the experience of the entire author teams.

In total we collected answers from 51 authors in 35 of the 37 countries/territories (we could not collect answers from any report authors in Burundi and Serbia). Academics (35%) and national government employees (31%) represented the majority of the survey participants, followed by those at government research institutions (20%), not-for-profit organisations (10%), private sectors (10%), and others (6%: the sum of the percentages exceeds 100, as some participants selected multiple options).

All participants had high levels of experience working in conservation, with a median 20 years of experience (Extended Data Fig. 3).

Relevance of the references was the major reason that report authors cited non-English-language academic literature (75% of report authors selected "Relevant" in Fig. 2a). In contrast, a much smaller proportion of report authors selected accessibility (39% for "Easy to find" and 20% for "Easy to access") and understandability (26% for "Easy to understand" and 18% for "Easy for readers") as a reason for citing non-English-language academic literature (Fig. 2a). The pattern was quite similar to the reasons for citing non-English-language grey literature (Fig. 2b).

English-language academic literature was cited because report authors thought it was relevant ("Relevant", 65%), credible ("High quality", 55%), accessible ("Easy to find", 49%), and widely recognised ("Widely recognised", 51%) (Fig. 2a). Few report authors selected understandability (14% for both "Easy to understand" and "Easy for readers") as a reason for citing English-language academic literature (Fig. 2a). For English-language grey literature, the relevance of references was the only reason that was selected by over half of the report authors (57%, Fig. 2b).

Barriers to the use of English-language literature

Although most of the report authors self-reported relatively high English proficiency (Extended Data Fig. 4, 72% answered that it is easy or very easy to understand an English-language paper), 8% and 24% of them experienced difficulties in searching and understanding English-language literature for their reports, respectively (Fig. 3). The report authors with lower English proficiency were more likely to have experienced such difficulties in searching (generalised linear mixed model: coefficient = 16.42, SE = 8.14, $z = 2.02$, $p = 0.044$) and understanding English-language literature (coefficient = 0.85, SE = 0.40, $z = 2.14$, $p = 0.032$; Extended Data Fig. 5). Further, 8% of the report authors answered that they could not cite relevant English-language literature due to difficulties in understanding it (Fig. 3). About 27% of the report authors indicated that their reports could have improved if they had used more English-language literature; however, a slightly larger proportion of the report authors also indicated that their reports could have improved if they had used more non-English-language literature (Fig. 3).

Solutions to aiding the use of English-language literature

We also asked how report authors perceived the two potential solutions to aiding the use of English-

language literature (providing non-English-language title, abstract or main text, and using machine translation), proposed by earlier studies[5,7,14]. About half the report authors indicated that non-English-language titles and abstracts would help them search for (51%) and understand (56%) English-language literature (Fig. 4a). The availability of non-English-language main text, in addition to title and abstract, for English-language literature did not affect the proportion greatly (47% and 59% indicated that it helps for searching and understanding English-language literature, respectively: Fig. 4a), indicating that the availability of non-English-language title and abstract is a key first step. Although most report authors did not frequently use machine translation (Fig. 4b), approximately a quarter and half of them reported that machine translation helped them search for and understand English-language literature, respectively (Fig. 4c). For those who did not find machine translation helpful, the main reason was inadequate quality (Extended Data Fig. 6).

Discussion

Our results uncover the widespread use of non-English-language literature as a source of information in national biodiversity assessments. There was a considerable inter-country variation in the proportion of non-English language references cited, with countries with lower English proficiency and lower economic development citing more non-English-language references. This result implies the following two, not mutually exclusive, possibilities. First, knowledge producers (i.e., those producing scientific literature, such as scientists and practitioners) in countries/territories with lower English proficiency and lower economic development may be more likely to publish their work in a non-English language (i.e., the official language of the country, or any other dominant language). This is either due to their own low English proficiency, or in consideration of the low English proficiency and financial difficulty in accessing English-language literature among the anticipated users of the scientific information they are publishing. This could be leading to a higher availability of important scientific knowledge in non-English-language literature. Second, report authors in those countries/territories may struggle more with searching, understanding, and accessing English-language literature due to the lack of English proficiency or necessary funds, resulting in a heavier reliance on non-English-language literature.

The survey results seem to support the first possibility; most report authors indicated that they cited those non-English-language references because they were truly relevant to the report, and not necessarily because they were more easily accessible or understandable. Clearly, scientific knowledge

283 that is relevant to national biodiversity assessments is still being published in non-English-language
284 literature even in this era of supposed English dominance in scientific publishing, which is recognised,
285 and actively used, as an important information source across countries/territories where English is not
286 widely spoken. While the quality of non-English-language science may tend to be lower than that of
287 English-language science[7], studies published in non-English languages are known to provide unique
288 scientific information, such as information on local species in countries/territories where relevant
289 English-language studies are not available[7,8]. Examples of such cases found in this study include a
290 Japanese-language review on historical changes in grassland area in Japan[15], cited in the Japan
291 Biodiversity Outlook 2[16], a simplified Chinese-language study on the relative value of total ecosystem
292 services to the regional GDP in the Xishuangbanna region[17], cited in China's fifth national report on
293 the implementation of the CBD[18], and a Spanish-language study reporting the impact of deforestation
294 on the erosion in the Magdalena River drainage basin[19], cited in a national report on the status and
295 trends of Colombia's biodiversity[20].

296 Such scientific knowledge available in non-English languages is, however, far less frequently used
297 in international biodiversity assessments compared to assessments of any countries/territories covered
298 in this study[10]. English-language literature cited in international assessments is unlikely to cover
299 scientific knowledge published in non-English languages, as citing non-English-language literature is
300 often discouraged in English-language publications, [21] and non-English-language studies are
301 commonly excluded from English-language meta-analysis and systematic reviews[22]. This means that
302 international assessments may overlook important, locally and regionally-relevant scientific
303 information on biodiversity conservation. IPBES biodiversity assessments, for example, involve
304 experts with diverse linguistic backgrounds[10], who are likely to be aware of the importance of non-
305 English-language literature and also have relevant language skills for searching and understanding it.
306 Yet, the assessments are essentially based on English-language literature. This suggests that the non-
307 use of non-English-language literature in IPBES assessments could be the result of its importance not
308 properly emphasised[23] and hence its citation being discouraged or refrained. Indeed, the IPBES guide
309 on the production of assessments states, "Contributions [from contributing authors] should be
310 supported, as far as possible, with references from peer-reviewed and *internationally available*
311 literature" [24], which could implicitly discourage contributing authors to cite non-English-language
312 literature. This disregard for relevant non-English-language literature in international assessments
313 could be a serious issue, given that these reports are meant to be a global synthesis of national-level

information.

The survey results also highlight the consequences of language barriers to the use of English-language literature in national biodiversity assessments. Although language barriers did not seem to actually prevent report authors from citing English-language literature, a non-negligible proportion (a quarter) of report authors, especially those with lower self-reported English proficiency, struggled with understanding English-language literature when compiling their reports. The higher reliance on non-English-language academic literature in countries with lower economic development signals the significance of financial inaccessibility as another barrier to the use of English-language academic literature. Most report authors recognise English-language academic literature as a relevant, high-quality, and widely-recognised source of scientific knowledge, but they require extra effort and funds to search for, access, and understand them. Environmental decision-makers are known to face over 200 barriers to the use of science in their decisions[4]; the additional effort required to understand English-language literature could present yet another substantial burden for them, potentially leading to a poorer uptake of relevant scientific evidence.

Providing a non-English-language title and abstract of English-language literature is supported by almost half the report authors as a promising solution to overcoming the language barrier to the use of English-language literature. Although an increasing number of English-language journals allow authors to provide non-English-language abstracts, and sometimes main texts, of their papers, no studies to date have assessed the actual effectiveness of this practice. Our results provide concrete evidence that supplying non-English-language abstracts could help lower language barriers to the use of English-language scientific knowledge. This approach, however, is still far from being a common practice across disciplines. We need a concerted effort from scientific communities to make this solution more pervasive; authors should make sure to provide at least the title and abstract, and the main text if possible, of their English-language papers in other relevant language(s) in an easily understandable way for non-experts, while more journals, especially those targeted at international readers, should allow and actively encourage authors to do so. The visibility of non-English-language abstracts matters too, as many journals that do provide non-English-language abstracts still publish them only as a part of supplementary information, which is very hard for readers to find. Non-English-language abstracts should be presented together with English-language abstracts, as is the case in, for example, British Ecological Society journals.

Machine translation also seems to be recognised by report authors as a potential solution to aiding the understanding of English-language literature. The quality of machine translation has improved drastically over the years[25], and machine translation is increasingly being used in science communication, for example, to assist communication with patients in health settings[26]. However, understandably, concerns over the accuracy of machine translation, especially when applied to scientific terms[27], still limit its broader implementation in science communication[26]. The inadequate quality of machine translation was also recognised by some of the report authors who participated in the survey (Extended Data Fig. 6). This is also likely why most academic journals have not integrated machine translation on their websites. Similarly, many major literature search systems (e.g., Web of Science and Scopus) display their platforms in some non-English languages, but do not fully integrate machine translation into their systems; this was another reason why report authors did not think that machine translation could help with English-language literature searches (Extended Data Fig. 6). Attempts to multi-lingualise literature searches using machine translation are emerging (e.g., litsearchr package in R translates search strings into multiple languages[28]), although the effectiveness of these attempts should be further explored. Another issue with regards to the use of machine translation in science communication is that the small number of languages with a dominant online presence, such as English, Spanish, German, Japanese, and French, are over-represented in the recent evolution of technologies and applications associated with machine translation[29]. Most of the world's languages still face a serious lack of digital language resources needed for developing and improving machine translation for that language. Those languages with fewer speakers are often spoken in biodiversity hotspots, and thus are key to communicating science[30] as well as accessing traditional knowledge relating to those hotspots[31]. There is thus a risk that relying on machine translation alone could further exacerbate the existing disparity among speakers of different languages. The true effectiveness and applicability of machine translation to scientific communication is a complex issue warranting a separate discussion, and is beyond the scope of this paper. However, while its limitations should be kept in mind, machine translation does offer the potential to aid the transfer of scientific knowledge across languages, especially with its quality improving over time, and in particular when those languages with sufficient online presence are concerned.

Our results also highlight the importance of non-English-language grey literature in informing national biodiversity assessments. Across 37 countries/territories, 65% of the references cited were, on average, non-English-language grey literature. In many countries, for example, masters and PhD

theses are often written in a non-English language[32] and not necessarily published later in more internationally-visible, peer-reviewed journals[33]. Similarly, most governmental reports are usually only available in a local, non-English language. There is now an increasing recognition of the importance of grey literature in informing environmental evidence synthesis[34], and our results corroborate that the argument also applies to non-English-language grey literature. Non-English-language grey literature may be especially important as a source of scientific information in countries with low English proficiency, as English proficiency was negatively associated with the proportion of non-English-language references (i.e., academic and grey literature combined) cited but not with the proportion of non-English-language academic literature.

This study is likely to have underestimated the overall level of non-English-language literature used in national biodiversity assessments, as we could not sufficiently cover countries in, for example, Western Asia and North Africa, where non-English-language literature is also expected to be frequently used due to lower national levels of English proficiency[13] and limited accessibility to English-language literature. The level of English language barriers for non-academic communities including environmental decision-makers could also be more severe than the level we found in this study, as among our survey respondents, decision-makers (i.e., non-academics in Extended Data Fig. 4) tended to have lower self-reported English proficiency and were more likely to experience language barriers when citing English-language references (Extended Data Fig. 5).

The national-level usage of scientific literature in different languages uncovered in this study mirrors two major consequences of language barriers in achieving global biodiversity targets for the next decade. The Kunming-Montreal Global Biodiversity Framework proposed by the CBD aims to "*Ensure that the best available data, information and knowledge, are accessible to decision makers, practitioners and the public to guide effective and equitable governance, integrated and participatory management of biodiversity*" (Target 21) [3]. On the one hand, we uncovered that non-English-language literature is routinely used as a unique source of relevant scientific information at the national level but almost entirely ignored at the international level. Future assessments and decision-making on biodiversity conservation at the international level must not dismiss relevant knowledge simply due to the language of its publication. This also applies to national-level assessments and decision-making. For example, the distribution of many species spans multiple countries where different non-English languages are spoken[12]. In such a case, transferring relevant knowledge between non-English languages could be key to the conservation of those species. On the other hand, we also revealed that

406 decision-makers face difficulties in identifying and utilising scientific knowledge if relevant
407 knowledge is provided only in English. We must ensure that English-language scientific knowledge is
408 easily accessible, i.e., available also in a relevant language for its users. This will facilitate the use of
409 the best scientific evidence in environmental decisions across all countries, including those where
410 English is not widely spoken and, quite often, biodiversity is threatened the most[9]. Language barriers
411 in biodiversity conservation, and more generally in other applications of science, have just recently
412 started attracting attention[14]. Some of the solutions provided here are relatively easy to implement
413 (e.g., encouraging the use of non-English-language literature in international assessments, or providing
414 non-English-language abstracts of papers) while others await further developments (e.g.,
415 implementing reliable machine translation into literature search systems). We urge scientific
416 communities to turn their eyes to this overlooked issue, and make a concerted effort to understand its
417 consequences and devise and implement solutions.
418
419 **Methods**
420 **Ethics declaration**
421 The survey in this study was conducted in accordance with the University of Queensland's
422 Institutional Human Research Ethics Approval (approval number 2020001838). All participants were
423 at least 18 years old and provided written consent indicating their agreement to participate in the
424 survey. The Participant Information Sheet clarified the voluntary nature of participation, the aims of
425 the research, how the data would be used and that all data would be confidential.
426
427 **Target countries/territories**
428 Our previous work[12] that compiled information on official languages in each country/territory from the
429 World Factbook 2021[35] identified 166 countries/territories where English was not an official language.
430 In this study we aimed to include as many of the 166 countries/territories as possible. We first used a
431 range of approaches (e.g., known networks, social media, e-mail lists, and the website of the translatE
432 project: https://translatesciences.com/) to recruit coordinators for any countries/territories (hereafter
433 referred to as country coordinators) where English is not an official language. The country
434 coordinators were required to have at least a bachelor's degree, but often had higher research degrees,
435 in a relevant discipline, such as ecology or conservation science. We aimed to include as many
436 countries as possible from each of the four different regions of the world defined by the IPBES
437 (Africa, Americas, Asia-Pacific, and Europe-Central Asia) [11]. However, some regions were inevitably

under-represented (Supplementary Data 1) because (i) we were unable to find country coordinators who were willing or able to collaborate, despite considerable efforts made and (ii) in some countries all reports identified did not meet our selection criteria (see **Identifying national reports on biodiversity assessments**). For example, the country coordinators from nine countries (Albania, Bolivia, Cambodia, Côte d'Ivoire, Estonia, Lithuania, Macedonia, Mongolia, and Montenegro) were unable to complete the required tasks. Although we also found willing country coordinators in Bangladesh, Maldives, Myanmar, Nepal, and Sri Lanka, all reports identified from Bangladesh, Maldives, Nepal, and Sri Lanka were published in English while the country coordinator in Myanmar could not keep contributing due to the military coup. See **Discussion** for the potential consequences of geographical bias in the sampled countries/territories. All country coordinators who completed the required tasks were involved in this study as coauthors.

Identifying national reports on biodiversity assessments

We first identified relevant national reports on biodiversity assessments in each country/territory. Each country coordinator used a range of approaches (e.g., personal knowledge, opinions of colleagues, online searches, etc) to identify as many relevant reports as possible in the country/territory, using all of the following eligibility criteria:

1. The report must be about biodiversity and/or its conservation (but reports on the conservation status of biodiversity are preferred) across the entire country/territory (i.e., cannot be about a specific region within a country/territory).
2. The report must cover at least an entire group of species, such as bird species or pollinators (but reports covering broader species groups are preferred).
3. The report must be written in a non-English language, or have a non-English-language version, in addition to an English version.
4. The report must have at least 15 references including at least one non-English-language reference cited, with the list of references cited made available.
5. The report must have been published during the past 15 years (i.e., in 2005 or later, but newer reports are preferred).
6. The report must be published by either the government or other organisations, such as universities or conservation NGOs (but governmental reports are preferred).

We used eligibility criteria 3 and 4 above to exclude reports where citations to non-English-language

469 references were deliberately avoided, as citing non-English-language references is often discouraged
470 or avoided especially in English-language literature[21]. For each report identified as potentially
471 relevant, we recorded the following information:

472 · The country/territory of report publication,
473 · Title of the report in the non-English language and in English (translated if an English title does
474 not exist),
475 · Publication language(s),
476 · Organisation(s) that edited/published the report,
477 · Name and contact of the report editor(s)/author(s),
478 · Publication year,
479 · Broad description of the report topic, and
480 · URL.

481 We then selected the report from each country/territory that best suited the eligible criteria (see
482 Supplementary Data 1). For example, we chose a report on the conservation status of biodiversity over
483 a report describing species found in the country (Criterion 1), a report covering multiple species
484 groups (e.g., plants and animals) over a report focusing only on a single species group (Criterion 2), a
485 newer edition if multiple editions existed for different years (Criterion 5), and a governmental report
486 over a non-governmental report (Criterion 6).
487

488 **Recording the number of references cited**

489 For the selected reports in each country/territory, we counted and recorded the number of references
490 cited, for each of the following four categories: (i) English-language traditional academic literature
491 (i.e., peer-reviewed journal papers and books), (ii) English-language grey literature (i.e., all other
492 literature types not controlled by commercial publishing, such as governmental reports, websites,
493 databases, theses, etc), (iii) non-English-language traditional academic literature, and (iv) non-English-
494 language grey literature. The report selected for Romania included nine other sub-reports, and we thus
495 used the total number of references cited in the report itself and the nine sub-reports.
496

497 **Questionnaire survey with editors/authors**

498 To understand the barriers and enablers affecting the use of references in English- and non-English
499 languages, we conducted a questionnaire survey (Supplementary Text 1) with at least one author or

editor of each report. Our aim here was to secure one participant from each country who played as major a role as possible, assuming that their responses would represent the experience of the whole author/editor team (if multiple authors/editors were involved in the report). To achieve this we adopted the following sampling strategy:

1. Each country coordinator identified one author/editor who played the most important role (e.g., corresponding author or chief editor) and invited the author/editor to complete the survey. If more than one author/editor played a similarly important role (e.g., leading authors of multiple relevant chapters), the coordinator contacted more than one author/editor simultaneously (this applied to ten countries: Argentina, Chile, China, Costa Rica, Hungary, Indonesia, Japan, Malaysia, Russia, and Slovakia). If the author(s)/editor(s) did not respond, the country coordinator sent at least two reminders.
2. Where at least one author/editor from Step 1 completed the survey, the country coordinator stopped the sampling process, and we used the data submitted as a representative sample of the country/territory. If we had more than one participant from a country/territory, we used data from all participants (this was accounted for in the analysis; see **Analysis**).
3. If no author/editor participated in Step 1, the country coordinator identified and contacted another author/editor who played the second most important role (e.g., second author, or another senior editor). In some countries, the author/editor whom the country coordinator contacted first referred us to another author/editor, in which case the country coordinator contacted that author/editor. Again if the author(s)/editor(s) did not respond, the country coordinator sent at least two reminders.
4. Each country coordinator repeated Steps 2 and 3 until at least one author/editor had participated from each country/territory.

All correspondence was conducted via email and the survey was sent as an attached Microsoft Word file between September 2020 and July 2021 (depending on countries/territories). The completed survey was submitted electronically in a Microsoft Word file to the relevant country coordinator, who anonymised the response before sending it to the data analyst. None of the country coordinators participated in the survey themselves. In two countries (Burundi and Serbia) we were not able to collect data from any author/editor although the respective country coordinator contacted all relevant authors/editors and sent at least two reminders. Those two countries were therefore excluded from the relevant part of the analysis. See Supplementary Data 2 for the number of authors/reports whom we contacted and those who completed the survey.

The survey consisted of three sections (see Supplementary Text 1 for more detail). The first section (Q1-5) comprised questions on demographic information, such as the first language and self-reported English proficiency of report authors. The second section (Q6-16) included questions on reasons for citing different types of references and the level of English-language barriers perceived by report authors. The third section (Q17-26) includes questions on potential solutions to facilitating the use of English-language literature in national reports on biodiversity conservation. Here we focused on two potential solutions (providing non-English-language title, abstract or main text, and using machine translation) proposed by earlier studies[5,7,14]. To maximise the response rate, the survey was translated by relevant country coordinators into French, Italian, Japanese, Korean, simplified Chinese, Romanian, Russian, Spanish, Turkish, Ukrainian, and Vietnamese, before being shared with report authors in countries where those languages are an official language.

Analysis

Some survey participants did not answer some questions, in which case we recorded these answers as missing values (i.e., NA) and excluded them from the analysis. One participant selected both Yes and Unsure, or both Yes and No, in three questions asking if participants experienced English-language barriers (Questions 11, 12, and 13 in Supplement Text 1), for which we recorded Yes as the answer, assuming that the participant experienced those English-language barriers at least to some degree.

We applied generalised linear models with a binomial distribution, implemented in R 4.1.2[36], to test the association between (i) the proportion of non-English-language references (i.e., academic and grey literature combined) or (ii) the proportion of non-English-language academic literature in each report as the response variables, and the English Proficiency Index[13] and $\log_{10}$-transformed GDP per capita (based on purchasing power parity, current international $) in 2020[37] of each country as the explanatory variables. The English Proficiency Index measures the average English proficiency in each country, based on an 800 point scale, with scores less than 450 representing the Very Low Proficiency, 450-499 the Low Proficiency, 500-549 the Moderate Proficiency, 550-599 the High Proficiency, and 600-800 the Very High Proficiency bands, respectively[13]. GDP per capita measures the level of economic development in each country. The English Proficiency Index was not available in Burundi, Lebanon, Mozambique, Senegal, and Taiwan and GDP per capita was also unavailable in Taiwan. Those five countries/territories were therefore excluded from this analysis. Our hypothesis was that the use of non-English-language literature was more prevalent in countries/territories with

lower English proficiency and lower economic development. The variance inflation factor for the two explanatory variables (calculated using the R package "car"[38]) was 1.94, indicating a low level of multicollinearity.

The English proficiency of individual report authors was measured by asking how easily each participant could read and understand the full text of an English-language peer-reviewed paper on biodiversity conservation (on a five-point scale: very easy, easy, neutral, difficult, or very difficult), shown in Extended Data Fig. 4. To test the relationship between the self-reported English proficiency of report authors (the explanatory variable) and their experience of encountering difficulties in searching and understanding English-language literature (Yes or No, the response variable), we applied generalised linear mixed models with a binomial distribution, using country/territory as a random factor to account for multiple participants in ten countries.

We also used the following R packages: gridExtra[39], maps[40], patchwork[41], RColorBrewer[42], and tidyverse[43].

Data Availability

Data on 333 biodiversity assessment reports identified in 37 countries/territories, on 37 reports used for the analysis, and on 130 reports in 11 countries used for the sensitivity analysis are available as Supplementary Data 1, 2, and 3, respectively. We are unable to make data on the report authors' responses to the survey questions publicly available, as per our agreement with the University of Queensland Ethics office and due to the confidentiality of the data.

Code Availability

All codes used in the analysis are available at: http://doi.org/10.17605/OSF.IO/Y94ZT.

Corresponding author

Correspondence to T.A. (t.amano@uq.edu.au)

Acknowledgements

Thanks to all authors and editors of the reports who participated in our survey, V. V. Natykanets for his help in data collection, and M. Amano for all the support. This work was funded by the Australian Research Council Future Fellowship FT180100354, the University of Queensland strategic funding (T.A.), the Australian Research Council DECRA Fellowship DE180100202 (S.M.D.), Coordenação de

Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) – Finance Code 001 (J.P.N.G), and the Swiss National Science Foundation (G.W.P.).

Author Contributions Statement

Conceptualization: T.A.

Formal analysis: T.A., V.B.-E.

Funding acquisition: T.A., S.M.D., J.P.N.G., G.W.P.

Investigation: T.A., V.B.-E., M.A., M.A.U.d.A.J., N.B., J.C., F.G.Ç., G.C, M.C.T., S.M.D., S.G., M.G., M.C.I., J.P. J.-D., R.K., F.P.S.L, H.-Y.L., E.L., P.M., L.M.B, A.-C.M., J.P.N.G, T.H.N., S.N.L, M.A.N., D.P.-J., P.P., G.W.P. , F.A.S, M.S., H.-M.S., Y.S., V.F., S.V., A.K.S.W., H.X., V.Z.G.

Methodology: T.A., V.B.-E., J.C.W.

Project administration: T.A., V.B.-E.

Validation: T.A., V.B.-E.

Visualisation: T.A.

Writing - original draft: T.A.

Writing - review & editing: T.A., V.B.-E., M.A., M.A.U.d.A.J., N.B., J.C., F.G.Ç., G.C, M.C.T., S.M.D., S.G., M.G., M.C.I., J.P. J.-D., R.K., F.P.S.L, H.-Y.L., E.L., P.M., L.M.B, A.-C.M., J.P.N.G, T,H.N., S.N.L, M.A.N., D.P.-J., P.P., G.W.P. , F.A.S, M.S., H.-M.S., Y.S., V.F., S.V., J.C.W., A.K.S.W., H.X., V.Z.G.

Competing Interests Statement

The authors declare no competing interests.

612 **Figure Legends**

613 **Fig. 1. The proportion of references cited in national biodiversity assessments by language and**
614 **literature type.** (a) The proportion of English-language academic (dark blue) and grey (pale blue)
615 literature, and non-English-language academic (orange) and grey (yellow) literature. The red and blue
616 solid lines indicate the mean proportion of non-English- and English-language references cited in
617 national biodiversity assessments across 37 countries/territories, respectively, while the red and blue
618 broken lines represent the mean proportion of non-English- and English-language references in the
619 eight biodiversity assessment reports by the Intergovernmental Science-Policy Platform on
620 Biodiversity and Ecosystem Services (IPBES) [10], respectively. The relationship between the proportion
621 of non-English-language references cited (academic and grey literature combined) and (b) the English
622 Proficiency Index (see **Methods** for more details) and (c) gross domestic product (GDP) per capita
623 (based on purchasing power parity (PPP), current international $) of each country. The size of each dot
624 indicates the total number of references cited in the report. The colours indicate regions (subregions
625 defined by the IPBES[11]). The regression curves (shown as black solid lines, and 95% confidence
626 intervals as shaded areas) are based on the fitted generalised linear models with a binomial distribution
627 (see Extended Data Table 1).
628
629 **Fig. 2. Reasons for citing English- and non-English-language (a) academic and (b) grey**
630 **literature in national biodiversity assessments.** The authors of national biodiversity assessments
631 were allowed to select multiple reasons. The x-axis shows the proportion of the report authors who
632 selected each reason. See Questions 6-9 in Supplementary Text 1 for the full description of each
633 reason. Answers were collected from 51 authors in 35 countries/territories (we could not collect
634 answers from the report authors in Burundi and Serbia).
635
636 **Fig. 3. Proportions of authors of national biodiversity assessment reports who have experienced**
637 **English language barriers.** Those who have experienced difficulties in searching (n = 50),
638 understanding (n = 51) English-language literature, those who could not cite English-language
639 literature due to difficulties in understanding (n = 49), and those who recognised that citing more
640 English-language or non-English-language literature could have improved their reports (n = 51).
641
642 **Fig. 4. Potential solutions to facilitating the use of English-language literature.** (a) The proportion

of report authors who indicated that a non-English-language title, abstract, and main text of English-language literature would help them search and understand English-language literature. (b) The frequency of use of machine translation when searching and/or reading English-language literature for the reports. Note that no report authors selected “Always” and so this option is now shown. (c) The proportion of report authors who indicated that machine translation helped them search and understand English-language literature. Answers were collected from 51 authors in 35 countries/territories (we could not collect answers from the report authors in Burundi and Serbia), apart from two questions (“Non-English title/abstract help you understand English literature” in (a) and “Machine translation helped you understand English literature” in (c)) where answers were available only from 50 authors.

653 **References**

654 1 Sutherland, W. J., Pullin, A. S., Dolman, P. M. & Knight, T. M. The need for evidence-based
655 conservation. *Trends in Ecology and Evolution* **19**, 305-308 (2004).
656 2 Sutherland, W. J. & Wordley, C. F. R. Evidence complacency hampers conservation. *Nature*
657 *Ecology & Evolution* **1**, 1215-1216, doi:10.1038/s41559-017-0244-1 (2017).
658 3 Convention on Biological Diversity. Kunming-Montreal Global biodiversity framework, Draft
659 decision submitted by the President. (2022).
660 4 Walsh, J. C., Dicks, L. V., Raymond, C. M. & Sutherland, W. J. A typology of barriers and
661 enablers of scientific evidence use in conservation practice. *Journal of Environmental*
662 *Management* **250**, 109481, doi:10.1016/j.jenvman.2019.109481 (2019).
663 5 Amano, T., González-Varo, J. P. & Sutherland, W. J. Languages are still a major barrier to
664 global science. *PLOS Biology* **14**, e2000933 (2016).
665 6 Fabian, Y. *et al.* How to close the science-practice gap in nature conservation? Information
666 sources used by practitioners. *Biological Conservation* **235**, 93-101,
667 doi:10.1016/j.biocon.2019.04.011 (2019).
668 7 Amano, T. *et al.* Tapping into non-English-language science for the conservation of global
669 biodiversity. *PLOS Biology* **19**, e3001296, doi:10.1371/journal.pbio.3001296 (2021).
670 8 Konno, K. *et al.* Ignoring non-English-language studies may bias ecological meta-analyses.
671 *Ecology and Evolution* **10**, 6373-6384, doi:10.1002/ece3.6368 (2020).
672 9 Amano, T. & Sutherland, W. J. Four barriers to the global understanding of biodiversity
673 conservation: wealth, language, geographical location and security. *Proceedings of the Royal*
674 *Society B: Biological Sciences* **280**, 20122649, doi:10.1098/rspb.2012.2649 (2013).
675 10 Lynch, A. J. *et al.* Culturally diverse expert teams have yet to bring comprehensive linguistic
676 diversity to intergovernmental ecosystem assessments. *One Earth* **4**, 269-278,
677 doi:https://doi.org/10.1016/j.oneear.2021.01.002 (2021).
678 11 Brooks, T. M. *et al.* Analysing biodiversity and conservation knowledge products to support
679 regional environmental assessments. *Scientific Data* **3**, 160007, doi:10.1038/sdata.2016.7
680 (2016).
681 12 Negret, P. J. *et al.* Language barriers in global bird conservation. *PLOS ONE* **17**, e0267151,
682 doi:10.1371/journal.pone.0267151 (2022).
683 13 EF Education First. EF English Proficiency Index. (2020).
684 14 Amano, T., Rios Rojas, C., Boum II, Y., Calvo, M. & Misra, B. B. Ten tips for overcoming

language barriers in science. *Nature Human Behaviour* **5**, 1119–1122, doi:10.1038/s41562-021-01137-1 (2021).

15 Ogura, J. Changes in grassland area in Japan. *Journal of Kyoto Seika University*, 159-172 (2006).

16 Ministry of the Environment. Japan Biodiversity Outlook 2. (2016).

17 Jing, Z.-P. & Ma, Y.-X. Dynamic evaluation on ecosystem service values of Xishuangbanna, Yunnan, China. *Journal of Central South University of Forestry & Technology* **32**, 87-93 (in Chinese with English summary) (2012).

18 Ministry of Environmental Protection of China. China's fifth national report on the implementation of the Convention on Biological Diversity. (2014).

19 Restrepo, J. D. The impact of deforestation on the erosion in the Magdalena River drainage basin (1980-2010). *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales* **39**, 250-267 (in Spanish with English summary) (2015).

20 Moreno, L. A. & Andrade, G. I. Biodiversidad 2019. Estado y tendencias de la biodiversidad continental de Colombia. (2019).

21 Lazarev, V. S. & Nazarovets, S. A. Don't dismiss citations to journals not published in English. *Nature* **556**, 174, doi:10.1038/d41586-018-04169-2 (2018).

22 Neimann Rasmussen, L. & Montgomery, P. The prevalence of and factors associated with inclusion of non-English language studies in Campbell systematic reviews: a survey and meta-epidemiological study. *Systematic Reviews* **7**, doi:10.1186/s13643-018-0786-6 (2018).

23 Nuñez, M. A. & Amano, T. Monolingual searches can limit and bias results in global literature reviews. *Nature Ecology & Evolution* **5**, 264, doi:10.1038/s41559-020-01369-w (2021).

24 IPBES. IPBES Guide on the production of assessments. (2018).

25 Graham, Y., Baldwin, T., Moffat, A. & Zobel, J. Is machine translation getting better over time? *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, 443–451, doi:10.3115/v1/e14 (2014).

26 Dew, K. N., Turner, A. M., Choi, Y. K., Bosold, A. & Kirchhoff, K. Development of machine translation technology for assisting health communication: A systematic review. *Journal of Biomedical Informatics* **85**, 56-67, doi:https://doi.org/10.1016/j.jbi.2018.07.018 (2018).

27 Tehseen, I., Tahir, G. R., Shakeel, K. & Ali, M. in *Artificial Intelligence Applications and Innovations. AIAI 2018. IFIP Advances in Information and Communication Technology, vol 519* (eds Iliadis L., Maglogiannis I., & Plagianakos V.) 196-206 (Springer International Publishing,

2018).

28 Grames, E. M., Stillman, A. N., Tingley, M. W. & Elphick, C. S. An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. *Methods in Ecology and Evolution* **10**, 1645-1654, doi:10.1111/2041-210x.13268 (2019).

29 Joshi, P., Santy, S., Budhiraja, A., Bali, K. & Choudhury, M. in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* 6282-6293 (Association for Computational Linguistics, 2020).

30 Amano, T. *et al.* Global distribution and drivers of language extinction risk. *Proceedings of the Royal Society B: Biological Sciences* **281**, 20141574, doi:10.1098/rspb.2014.1574 (2014).

31 Cámara-Leret, R. & Bascompte, J. Language extinction triggers the loss of unique medicinal knowledge. *Proceedings of the National Academy of Sciences* **118**, e2103683118, doi:10.1073/pnas.2103683118 (2021).

32 Pinto, S. & Araújo e Sá, M. H. Scientific research and languages in Portuguese Higher Education Institutions. *Language Problems and Language Planning* **44**, 20-44, doi:https://doi.org/10.1075/lplp.00054.pin (2020).

33 Hunter, N. B., North, M. A. & Slotow, R. The marginalisation of voice in the fight against climate change: The case of Lusophone Africa. *Environmental Science & Policy* **120**, 213-221, doi:10.1016/j.envsci.2021.03.012 (2021).

34 Haddaway, N. R. & Bayliss, H. R. Shades of grey: Two forms of grey literature important for reviews in conservation. *Biological Conservation* **191**, 827–829, doi:http://dx.doi.org/10.1016/j.biocon.2015.08.018 (2015).

35 Central Intelligence Agency. *The World Factbook 2021* (Central Intelligence Agency: https://www.cia.gov/the-world-factbook/, 2021).

36 R Core Team. *R: A language and environment for statistical computing* (R Foundation for Statistical Computing, 2021).

37 The World Bank. GDP per capita, PPP (current international $) (The World Bank, 2022).

38 Fox, J. & Weisberg, S. *An R companion to applied regression, third edition* (Sage, Thousand Oaks CA, 2019).

39 Auguie, B. *gridExtra: miscellaneous functions for "grid" graphics. R package version 2.3.* (2017).

40 Becker, R. A., Wilks, A. R., Brownrigg, R., Minka, T. P. & Deckmyn, A. *maps: draw geographical maps. R package version 3.4.0* (Available at: https://CRAN.R-

project.org/package=maps, 2021).

41 Pedersen, T. L. *patchwork: the composer of plots. R package version 1.1.2* (Available at: https://CRAN.R-project.org/package=patchwork, 2022).

42 Neuwirth, E. *RColorBrewer: ColorBrewer Palettes. R package version 1.1-2* (Available at: https://cran.r-project.org/package=RColorBrewer, 2014).

43 Wickham, H. *et al.* Welcome to the tidyverse. *Journal of Open Source Software* **4**, 1686, doi:https://doi.org/10.21105/joss.01686 (2019).

Fig. 1: References cited in national biodiversity assessments by language and literature type.

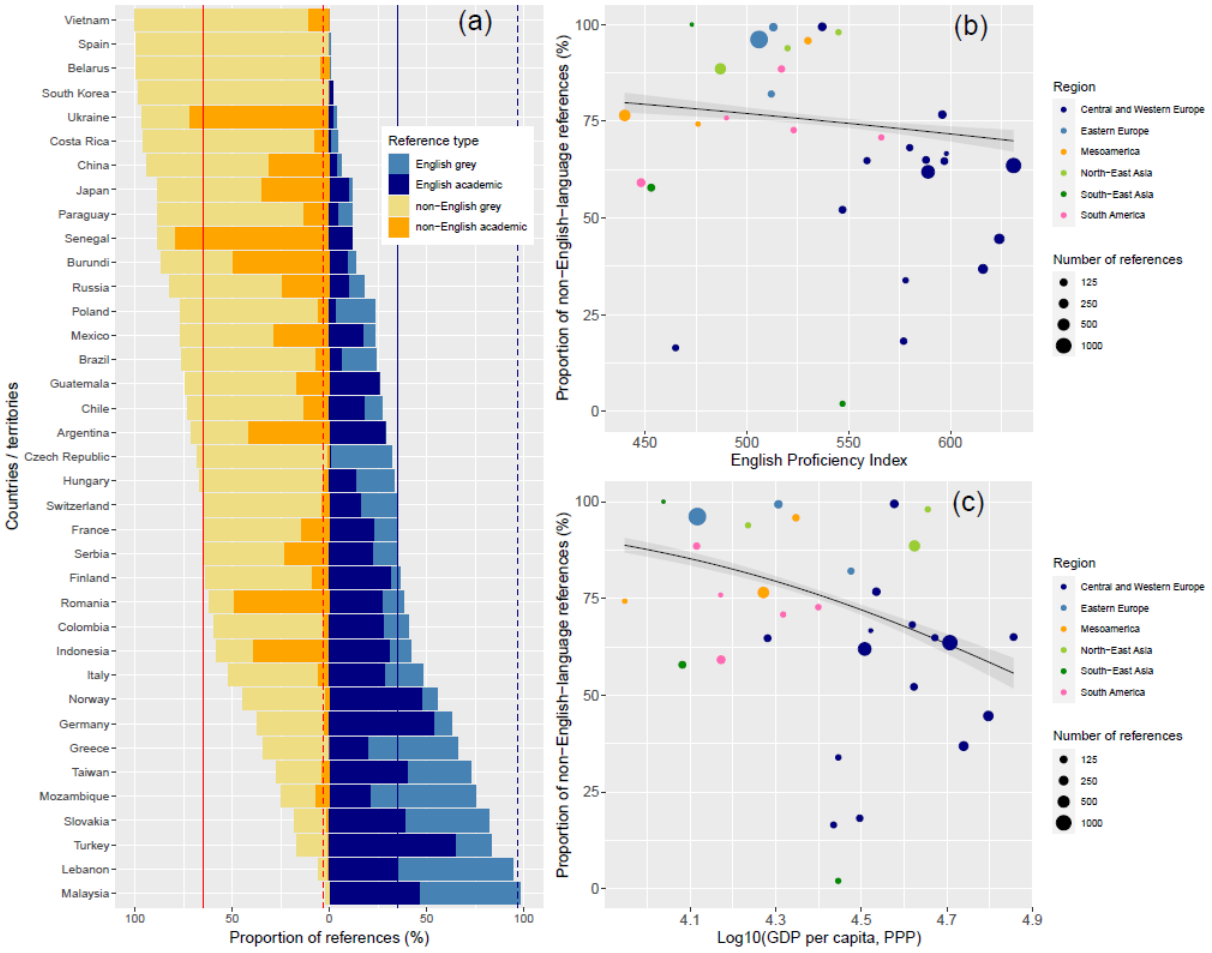
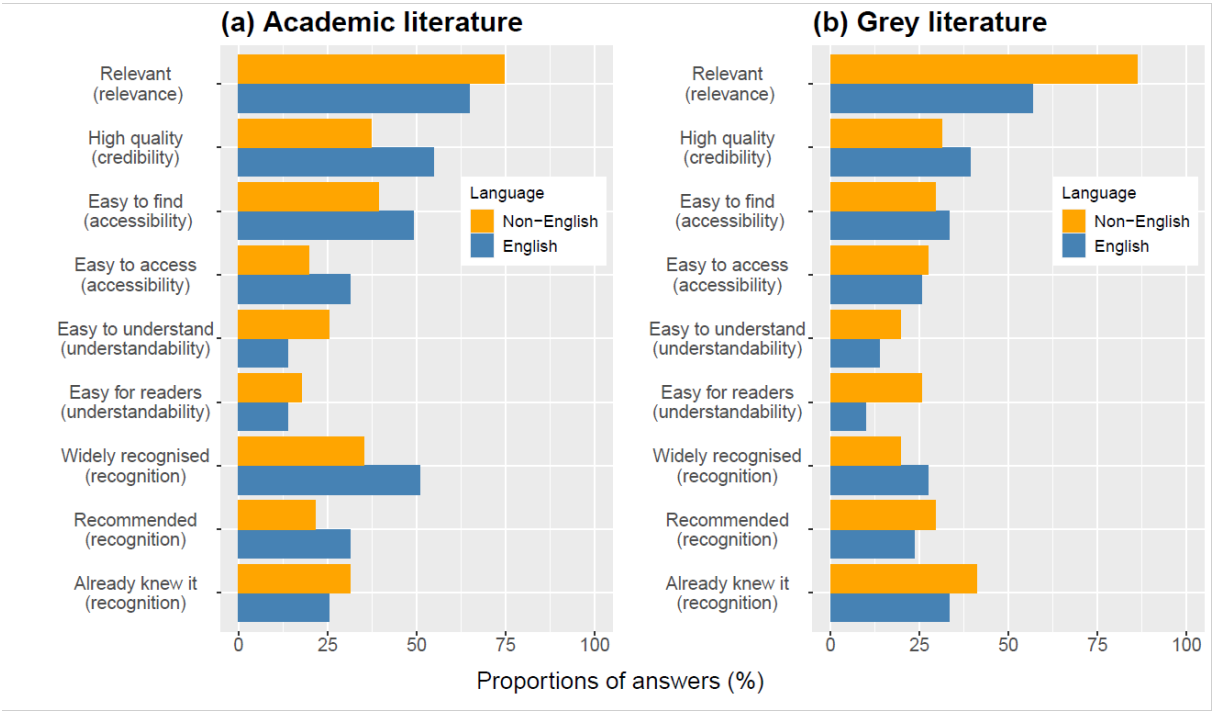
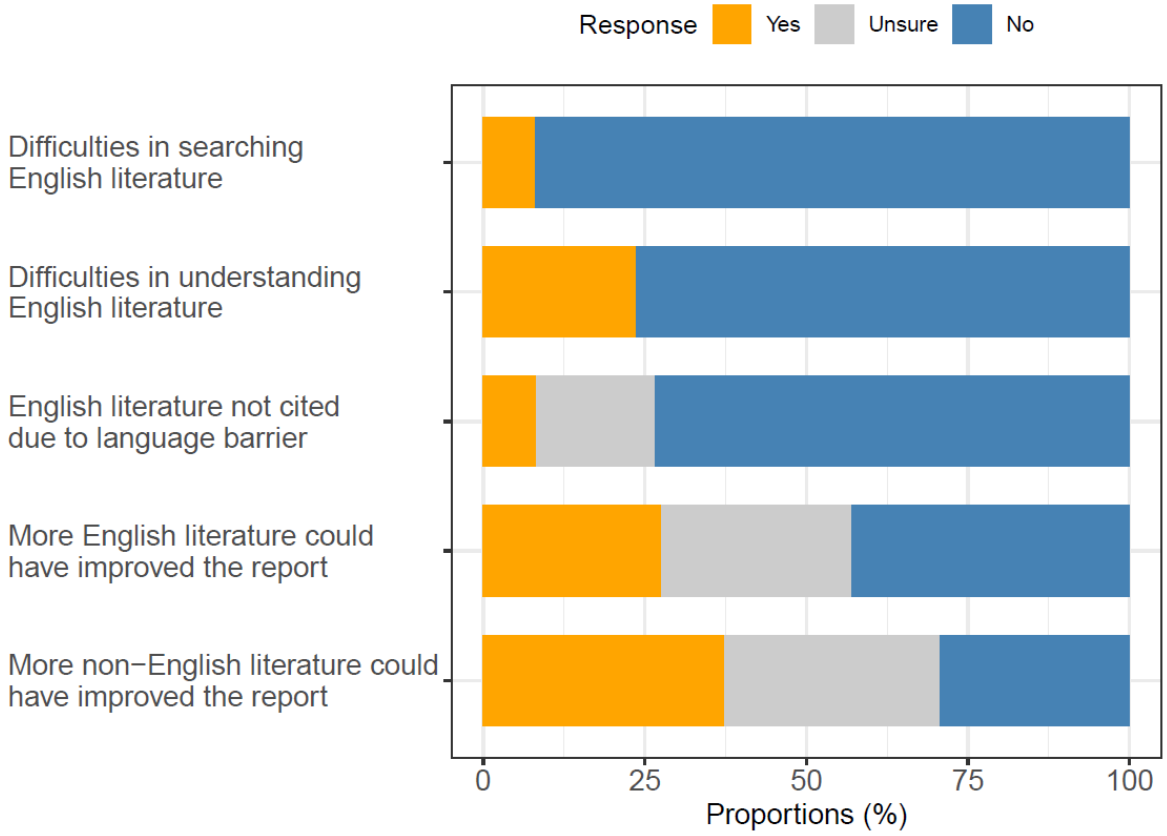


Fig. 2: Reasons for citing English- and non-English-language literature in national biodiversity assessments.

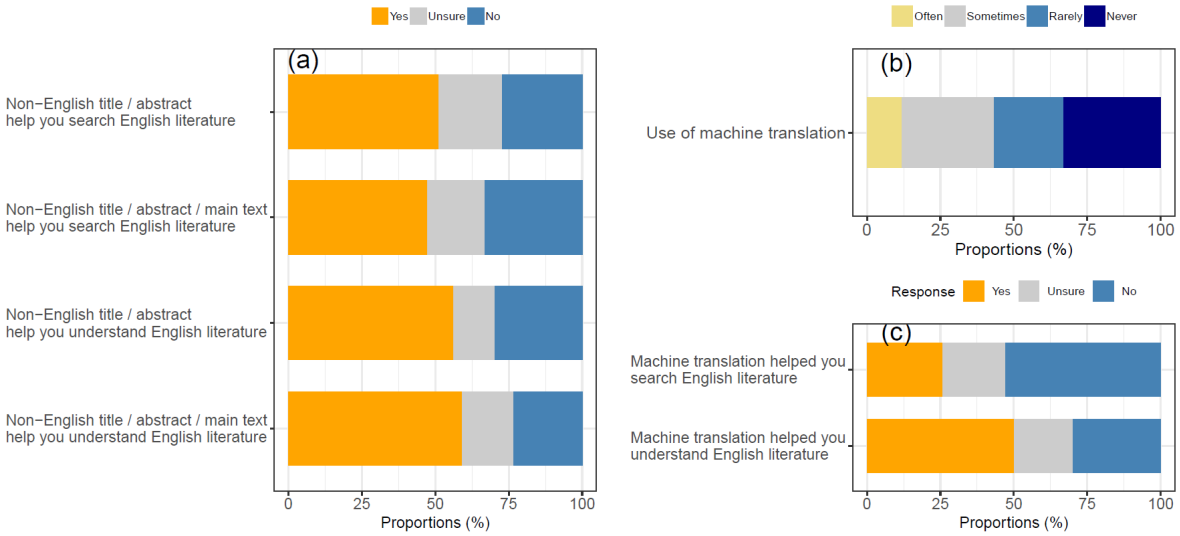


762 **Fig. 3: Authors who have experienced English-language barriers.**



763

764 **Fig. 4: Potential solutions to facilitating the use of English-language literature.**



765