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

[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, Zurich, 8057, Switzerland.

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

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

[11] Institute of Environmental Sciences, Jagiellonian University, 30-387, Poland.

[12] Stazione Zoologica Anton Dohrn, Via Francesco Caracciolo 333, Naples, 80122, Italy.

[13] Helmholtz Centre for Environmental Research - UFZ, Halle, 6114, Germany.

[14] Institute of Biology Bucharest, Romanian Academy, Splaiul Independentei 296, Bucharest, 60031, Romania.

[15] Fundación Auilaff, Puerto Varas, Region de Los Lagos, 5550000, Chile.

[16] Graduate School of Agriculture,Tokyo University of Agriculture and Technology, Fuchu, Tokyo, Japan, 183-8509, Japan.

[17] HCV Network, C/O Critchleys Llp, Beaver House, 23-38 Hythe Bridge Street, Oxford, OX1 2EP, United Kingdom.

[18] Institute of Environmental and Interdisciplinary Science and Department of Biology, Carleton University, Ottawa, Ontario, K1S 5B6, Canada.

[19] Centro de Estudios Ambientales y Biodiversidad, Universidad del Valle de Guatemala, 11 calle 15-79 zona 15 V.H. III, 1015, Guatemala.

[20] Institute of Vertebrate Biology, Czech Academy of Sciences, Květná 8, Brno, 603 65, Czech Republic.

[21] Joint Remote Sensing Research Program, School of Earth and Environmental Sciences, The University of Queensland, Brisbane, Queensland, 4067, Australia.

[22] Chair of Nature Conservation and Landscape Ecology, University of Freiburg, Tennenbacherstr. 4, Freiburg, 79115, Germany.

[23] Departamento de Botânica, Instituto de Biociências, Universidade de São Paulo, Rua do Matão nº 277, Cidade Universitária, São Paulo, SP, CEP 05508-090, Brasil.

[24] Forest Ecology and Conservation Group, Conservation Research Institute and Department of Plant Sciences, University of Cambridge, Cambridge, CB2 3EA, UK.

[25] Scientific and Practical Center of the National Academy of Sciences of Belarus for Bioresources, Akademicheskaya str., 27, Minsk, Republic of Belarus, 220072, Belarus.

[26] School of Business, The University of Queensland, Brisbane, Queensland, 4072, Australia.

[27] Eduardo Mondlane University, Faculty of Agronomy and Forest Engineering, 3453 Julius Nyerere Av., Maputo, 257, Mozambique.

[28] Nitidae, 16 Agostinho Neto, Maputo, Mozambique.

[29] Department of Biology and Biochemistry, University of Houston, 3455 Cullen Blvd #342, 77204, USA.

[30] Grupo de Ecología de Invasiones, INIBIOMA, CONICET, Universidad Nacional del Comahue,

Quintral 1250, San Carlos de Bariloche, 8400, Argentina.

[31] Department of Terrestrial Ecology, Norwegian Institute for Nature Research (NINA), P.P. BOX 5685, Torgarden, Trondheim, 7485, Norway.

[32] Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, The University of New South Wales, Sydney, New South Wales, 2052, Australia.

[33] Institute of Plant Sciences, University of Bern, Altenbergrain 21, 3013, Switzerland.

[34] The Biodiversity Consultancy, 3E King's Parade, Cambridge, CB2 1SJ, UK.

[35] Saint-Joseph University, Lebanon, Rue de Damas, Mar Mikhael, Beyrouth 1104 2020, 17-5208, Lebanon.

[36] Bird Protection and Study Society of Serbia, Partizanskih baza 6/43, Novi Sad, 21000, Serbia.

[37] Department of Agriculture, Forestry, and Bioresources, Seoul National University, 200-7213, Gwanak-ro 1, Gwanak-gu, Seoul, 08826, Republic of Korea.

[38] Center for Environmental Biology and Ecosystem Studies, National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan, 305-8506, Japan.

[39] Lendület Ecosystem Services Research Group, Institute of Ecology and Botany, Centre for Ecological Research, Alkotmány út 2-4, Vácrátót, 2163, Hungary.

[40] Faculty of Energy and Ecotechnology (GreenTech), ITMO University, Kronverkskiy Prospekt 49, St Petersburg, 197101, Russia.

[41] School of Biological Sciences, Monash University, 25 Rainforest Walk, Monash University, Clayton, Victoria, 3800, Australia.

[42] School of Environmental and Geographical Sciences, University of Nottingham Malaysia, Jalan Broga, Semenyih, Selangor Darul Ehsan, 43500, Malaysia.

[43] School of Life & Environmental Sciences, Deakin University, 221 Burwood Hwy, Burwood Victoria, 3125, Australia.

[44] The Commonwealth Scientific and Industrial Research Organisation Oceans and Atmosphere, 306 Carmody Road, St Lucia, Queensland 4067, Australia.

[45] 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 languages 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.

Introduction

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 20 of a new Global Biodiversity Framework proposed by the Convention on Biological Diversity (CBD) aiming to ensure that relevant 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

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

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

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

157 on the state of biodiversity, as they serve as the fundamental basis for evaluating past, and shaping
158 future, conservation actions and policies in each country while also informing regional and global
159 assessments. We identified relevant biodiversity assessment reports in 37 countries/territories where
160 the official language is not English, and investigated the proportion of English- and non-English-
161 language references cited in those reports. We further conducted a questionnaire survey with the
162 authors or editors of those reports (see **Methods** for more detail) to identify the barriers and enablers
163 affecting the use of references in English- and non-English languages.
164
165 **Results**
166 We identified a total of 333 eligible reports on biodiversity conservation in 37 countries/territories
167 where English is not an official language. These countries spanned across all four regions defined by
168 the IPBES (Africa, Americas, Asia-Pacific, and Europe-Central Asia) [11] and represented 22% of the
169 166 countries/territories where English is not an official language[12]. As we found only one eligible
170 report in seven out of the 37 countries, for consistency, we selected the one most relevant report in
171 each country/territory based on pre-defined criteria (i.e., 37 reports in total) for investigating the use of
172 references written in different languages (see **Methods** for more details). Most selected reports were
173 about the status of biodiversity, or the environment (including biodiversity) in general, at the national
174 level, but others included national reports to the CBD, national biodiversity strategies and action plans
175 (Supplementary Data S1). The sensitivity of our conclusions to the choice of reports included in the
176 analysis was minimal (see Supplementary Discussion).
177
178 **Use of scientific references in different languages**
179 For each of the 37 selected report we then recorded the number of references cited for each of the
180 following four categories: (i) English-language traditional academic literature (i.e., peer-reviewed
181 journal papers and books, hereafter "English-language academic literature"), (ii) English-language
182 grey literature (i.e., all other literature types not controlled by commercial publishing, such as
183 governmental reports, websites, databases, theses, etc.), (iii) non-English-language traditional
184 academic literature (hereafter "non-English-language academic literature"), and (iv) non-English-
185 language grey literature.

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

188 countries/territories covered in this study (Fig. S1). On average, 65% of the references cited were
189 written in a non-English language (red solid vertical line in Fig. 1a). Non-English-language literature
190 represented over half of the references cited in reports for 28 (76%) countries/territories and over 75%
191 in 15 (41%) countries/territories (Fig. 1a). These were in stark contrast to non-English-language
192 literature representing only 3.4% of the references cited in the IPBES assessment (red broken vertical
193 line in Fig. 1a, based on[10]). The proportion of non-English-language references cited in the reports was
194 significantly higher in countries with a lower English Proficiency Index (a measure of the average
195 English proficiency in each country[13], see **Methods** for more details) (Fig. 1b and Table S1) and in
196 countries with a lower gross domestic product (GDP) per capita (as a measure of economic level in
197 each country, Fig. 1c and Table S1).

198 A considerable proportion of the non-English-language literature cited was grey literature and
199 when focusing only on academic literature, 44% of the academic literature cited in those reports were,
200 on average, written in non-English languages. The proportion of non-English-language academic
201 literature cited in the reports was again significantly higher in countries with a lower GDP per capita
202 (Fig. S2 and Table S1). Some of the countries with a high English Proficiency Index and GDP per
203 capita, such as those in Central and Western Europe, cited a very low proportion (i.e., less than 10%)
204 of non-English-language academic literature (Figs. S1 and S2).
205
206 **Reasons for citing English and non-English-language references**
207 Next, we investigated the barriers and enablers affecting the use of references written in different
208 languages, by contacting at least one author or editor (hereafter "report author") of each report who
209 played a leading role in compiling their reports (as the corresponding author or chief editor in most
210 cases; see **Methods** for the sampling strategy). Their answers to questions in the survey (apart from
211 the questions asking information on authors themselves, such as their first language(s)) are thus
212 expected to represent the experience of the entire author teams.

213 In total we collected answers from 51 authors in 35 of the 37 countries/territories (we could not
214 collect answers from any report authors in Burundi and Serbia). Academics (35%) and national
215 government employees (31%) represented the majority of the survey participants, followed by those at
216 government research institutions (20%), not-for-profit organisations (10%), private sectors (10%), and
217 others (6%: the sum of the percentages exceeds 100, as some participants selected multiple options).
218 All participants had high levels of experience working in conservation, with a median 20 years of

219 experience (Fig. S3).

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

226 English-language academic literature was cited because report authors thought it was relevant
227 ("Relevant", 65%), credible ("High quality", 55%), accessible ("Easy to find", 49%), and widely
228 recognised ("Widely recognised", 51%) (Fig. 2a). Few report authors selected understandability (14%
229 for both "Easy to understand" and "Easy for readers") as a reason for citing English-language
230 academic literature (Fig. 2a). For English-language grey literature, the relevance of references was the
231 only reason that was selected by over half of the report authors (57%, Fig. 2b).
232
233 **Barriers to the use of English-language literature**
234 Although most of the report authors self-reported relatively high English proficiency (Fig. S4, 72%
235 answered that it is easy or very easy to understand an English-language paper), 8% and 24% of them
236 experienced difficulties in searching and understanding English-language literature for their reports,
237 respectively (Fig. 3). The report authors with lower English proficiency were more likely to have
238 experienced such difficulties in searching (generalised linear mixed model: coefficient = 16.42, SE =
239 8.14, $z = 2.02$, $p = 0.044$) and understanding English-language literature (coefficient = 0.85, SE = 0.40,
240 $z = 2.14$, $p = 0.032$; Fig. S5). Further, 8% of the report authors answered that they could not cite
241 relevant English-language literature due to difficulties in understanding it (Fig. 3). About 27% of the
242 report authors indicated that their reports could have improved if they had used more English-language
243 literature; however, a slightly larger proportion of the report authors also indicated that their reports
244 could have improved if they had used more non-English-language literature (Fig. 3).
245
246 **Solutions to aiding the use of English-language literature**
247 We also asked how report authors perceived the two potential solutions to aiding the use of English-
248 language literature (providing non-English-language title, abstract or main text, and using machine
249 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 (Fig. S6).

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 that is relevant to national biodiversity assessments is still being published in non-English-language literature even in this era of supposed English dominance in scientific publishing, which is recognised,

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

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

311 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 (Fig. S6). 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 (Fig. S6). 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

374 importance of grey literature in informing environmental evidence synthesis[34], and our results
375 corroborate that the argument also applies to non-English-language grey literature. Non-English-
376 language grey literature may be especially important as a source of scientific information in countries
377 with low English proficiency, as English proficiency was negatively associated with the proportion of
378 non-English-language references (i.e., academic and grey literature combined) cited but not with the
379 proportion of non-English-language academic literature.

380 This study is likely to have underestimated the overall level of non-English-language literature
381 used in national biodiversity assessments, as we could not sufficiently cover countries in, for example,
382 Western Asia and North Africa, where non-English-language literature is also expected to be
383 frequently used due to lower national levels of English proficiency[13] and limited accessibility to
384 English-language literature. The level of English language barriers for non-academic communities
385 including environmental decision-makers could also be more severe than the level we found in this
386 study, as among our survey respondents, decision-makers (i.e., non-academics in Fig. S4) tended to
387 have lower self-reported English proficiency and were more likely to experience language barriers
388 when citing English-language references (Fig. S5).

389 The national-level usage of scientific literature in different languages uncovered in this study
390 mirrors two major consequences of language barriers in achieving global biodiversity targets for the
391 next decade. A new Global Biodiversity Framework proposed by the CBD aims to "*Ensure that*
392 *relevant knowledge, including the traditional knowledge, innovations and practices of indigenous*
393 *peoples and local communities with their free, prior, and informed consent, guides decision-making*
394 *for the effective management of biodiversity, enabling monitoring, and by promoting awareness,*
395 *education and research*" (Target 20)[3]. On the one hand, we uncovered that non-English-language
396 literature is routinely used as a unique source of relevant scientific information at the national level but
397 almost entirely ignored at the international level. Future assessments and decision-making on
398 biodiversity conservation at the international level must not dismiss relevant knowledge simply due to
399 the language of its publication. This also applies to national-level assessments and decision-making.
400 For example, the distribution of many species spans multiple countries where different non-English
401 languages are spoken[12]. In such a case, transferring relevant knowledge between non-English
402 languages could be key to the conservation of those species. On the other hand, we also revealed that
403 decision-makers face difficulties in identifying and utilising scientific knowledge if relevant
404 knowledge is provided only in English. We must ensure that English-language scientific knowledge is

easily accessible, i.e., available also in a relevant language for its users. This will facilitate the use of the best scientific evidence in environmental decisions across all countries, including those where English is not widely spoken and, quite often, biodiversity is threatened the most[9]. Language barriers in biodiversity conservation, and more generally in other applications of science, have just recently started attracting attention[14]. Some of the solutions provided here are relatively easy to implement (e.g., encouraging the use of non-English-language literature in international assessments, or providing non-English-language abstracts of papers) while others await further developments (e.g., implementing reliable machine translation into literature search systems). We urge scientific communities to turn their eyes to this overlooked issue, and make a concerted effort to understand its consequences and devise and implement solutions.

Methods

Target countries/territories

Our previous work[12] that compiled information on official languages in each country/territory from the World Factbook 2021[35] identified 166 countries/territories where English was not an official language. In this study we aimed to include as many of the 166 countries/territories as possible. We first used a range of approaches (e.g., known networks, social media, e-mail lists, and the website of the translatE project: https://translatesciences.com/) to recruit coordinators for any countries/territories (hereafter referred to as country coordinators) where English is not an official language. The country coordinators were required to have at least a bachelor's degree, but often had higher research degrees, in a relevant discipline, such as ecology or conservation science. We aimed to include as many countries as possible from each of the four different regions of the world defined by the IPBES (Africa, Americas, Asia-Pacific, and Europe-Central Asia)[11]. However, some regions were inevitably under-represented (Supplementary Data S1) 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 references were deliberately avoided, as citing non-English-language references is often discouraged or avoided especially in English-language literature[21]. For each report identified as potentially relevant, we recorded the following information:

- The country/territory of report publication,
- Title of the report in the non-English language and in English (translated if an English title does not exist),
- Publication language(s),
- Organisation(s) that edited/published the report,
- Name and contact of the report editor(s)/author(s),

468 · Publication year,
469 · Broad description of the report topic, and
470 · URL.

471 We then selected the report from each country/territory that best suited the eligible criteria (see
472 Supplementary Data S1). For example, we chose a report on the conservation status of biodiversity
473 over a report describing species found in the country (Criterion 1), a report covering multiple species
474 groups (e.g., plants and animals) over a report focusing only on a single species group (Criterion 2), a
475 newer edition if multiple editions existed for different years (Criterion 5), and a governmental report
476 over a non-governmental report (Criterion 6).
477

478 **Recording the number of references cited**

479 For the selected reports in each country/territory, we counted and recorded the number of references
480 cited, for each of the following four categories: (i) English-language traditional academic literature
481 (i.e., peer-reviewed journal papers and books), (ii) English-language grey literature (i.e., all other
482 literature types not controlled by commercial publishing, such as governmental reports, websites,
483 databases, theses, etc), (iii) non-English-language traditional academic literature, and (iv) non-English-
484 language grey literature. The report selected for Romania included nine other sub-reports, and we thus
485 used the total number of references cited in the report itself and the nine sub-reports.
486

487 **Questionnaire survey with editors/authors**

488 To understand the barriers and enablers affecting the use of references in English- and non-English
489 languages, we conducted a questionnaire survey (Supplementary Text S1) with at least one author or
490 editor of each report. Our aim here was to secure one participant from each country who played as
491 major a role as possible, assuming that their responses would represent the experience of the whole
492 author/editor team (if multiple authors/editors were involved in the report). To achieve this we adopted
493 the following sampling strategy:

494 1. Each country coordinator identified one author/editor who played the most important role (e.g.,
495 corresponding author or chief editor) and invited the author/editor to complete the survey. If more
496 than one author/editor played a similarly important role (e.g., leading authors of multiple relevant
497 chapters), the coordinator contacted more than one author/editor simultaneously (this applied to ten
498 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 S2 for the number of authors/reports whom we contacted and those who completed the survey.

The survey consisted of three sections (see Supplementary Text S1 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.

The survey was conducted in accordance with the University of Queensland's Institutional Human Research Ethics Approval (approval number 2020001838). All participants were at least 18 years old and provided written consent indicating their agreement to participate in the survey. The Participant Information Sheet clarified the voluntary nature of participation, the aims of the research, how the data would be used and that all data would be confidential.

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 S1), 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.

560 The English proficiency of individual report authors was measured by asking how easily each
561 participant could read and understand the full text of an English-language peer-reviewed paper on
562 biodiversity conservation (on a five-point scale: very easy, easy, neutral, difficult, or very difficult),
563 shown in Fig. S4. To test the relationship between the self-reported English proficiency of report
564 authors (the explanatory variable) and their experience of encountering difficulties in searching and
565 understanding English-language literature (Yes or No, the response variable), we applied generalised
566 linear mixed models with a binomial distribution, using country/territory as a random factor to account
567 for multiple participants in ten countries.

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

570 **Data Availability**

571 Data on 333 biodiversity assessment reports identified in 37 countries/territories, on 37 reports used
572 for the analysis, and on 130 reports in 11 countries used for the sensitivity analysis are available as
573 Supplementary Data S1, S2, and S3, respectively. We are unable to make data on the report authors'
574 responses to the survey questions publicly available, as per our agreement with the University of
575 Queensland Ethics office and due to the confidentiality of the data.

576 **Code Availability**

577 All codes used in the analysis are available at: http://doi.org/10.17605/OSF.IO/Y94ZT.
578

References

1 Sutherland, W. J., Pullin, A. S., Dolman, P. M. & Knight, T. M. The need for evidence-based conservation. *Trends in Ecology and Evolution* **19**, 305-308 (2004).

2 Sutherland, W. J. & Wordley, C. F. R. Evidence complacency hampers conservation. *Nature Ecology & Evolution* **1**, 1215-1216, doi:10.1038/s41559-017-0244-1 (2017).

3 Convention on Biological Diversity. First draft of the post-2020 Global Biodiversity Framework. (2021).

4 Walsh, J. C., Dicks, L. V., Raymond, C. M. & Sutherland, W. J. A typology of barriers and enablers of scientific evidence use in conservation practice. *Journal of Environmental Management* **250**, 109481, doi:10.1016/j.jenvman.2019.109481 (2019).

5 Amano, T., González-Varo, J. P. & Sutherland, W. J. Languages are still a major barrier to global science. *PLOS Biology* **14**, e2000933 (2016).

6 Fabian, Y. *et al.* How to close the science-practice gap in nature conservation? Information sources used by practitioners. *Biological Conservation* **235**, 93-101, doi:10.1016/j.biocon.2019.04.011 (2019).

7 Amano, T. *et al.* Tapping into non-English-language science for the conservation of global biodiversity. *PLOS Biology* **19**, e3001296, doi:10.1371/journal.pbio.3001296 (2021).

8 Konno, K. *et al.* Ignoring non-English-language studies may bias ecological meta-analyses. *Ecology and Evolution* **10**, 6373-6384, doi:10.1002/ece3.6368 (2020).

9 Amano, T. & Sutherland, W. J. Four barriers to the global understanding of biodiversity conservation: wealth, language, geographical location and security. *Proceedings of the Royal Society B: Biological Sciences* **280**, 20122649, doi:10.1098/rspb.2012.2649 (2013).

10 Lynch, A. J. *et al.* Culturally diverse expert teams have yet to bring comprehensive linguistic diversity to intergovernmental ecosystem assessments. *One Earth* **4**, 269-278, doi:https://doi.org/10.1016/j.oneear.2021.01.002 (2021).

11 Brooks, T. M. *et al.* Analysing biodiversity and conservation knowledge products to support regional environmental assessments. *Scientific Data* **3**, 160007, doi:10.1038/sdata.2016.7 (2016).

12 Negret, P. J. *et al.* Language barriers in global bird conservation. *PLOS ONE* **17**, e0267151, doi:10.1371/journal.pone.0267151 (2022).

13 EF Education First. EF English Proficiency Index. (2020).

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,

643 2018).
644 28 Grames, E. M., Stillman, A. N., Tingley, M. W. & Elphick, C. S. An automated approach to
645 identifying search terms for systematic reviews using keyword co‐occurrence networks.
646 *Methods in Ecology and Evolution* **10**, 1645-1654, doi:10.1111/2041-210x.13268 (2019).
647 29 Joshi, P., Santy, S., Budhiraja, A., Bali, K. & Choudhury, M. in *Proceedings of the 58th Annual*
648 *Meeting of the Association for Computational Linguistics* 6282–6293 (Association for
649 Computational Linguistics, 2020).
650 30 Amano, T. *et al.* Global distribution and drivers of language extinction risk. *Proceedings of the*
651 *Royal Society B: Biological Sciences* **281**, 20141574, doi:10.1098/rspb.2014.1574 (2014).
652 31 Cámara-Leret, R. & Bascompte, J. Language extinction triggers the loss of unique medicinal
653 knowledge. *Proceedings of the National Academy of Sciences* **118**, e2103683118,
654 doi:10.1073/pnas.2103683118 (2021).
655 32 Pinto, S. & Araújo e Sá, M. H. Scientific research and languages in Portuguese Higher
656 Education Institutions. *Language Problems and Language Planning* **44**, 20-44,
657 doi:https://doi.org/10.1075/lplp.00054.pin (2020).
658 33 Hunter, N. B., North, M. A. & Slotow, R. The marginalisation of voice in the fight against
659 climate change: The case of Lusophone Africa. *Environmental Science & Policy* **120**, 213-221,
660 doi:10.1016/j.envsci.2021.03.012 (2021).
661 34 Haddaway, N. R. & Bayliss, H. R. Shades of grey: Two forms of grey literature important for
662 reviews in conservation. *Biological Conservation* **191**, 827–829,
663 doi:http://dx.doi.org/10.1016/j.biocon.2015.08.018 (2015).
664 35 Central Intelligence Agency. *The World Factbook 2021* (Central Intelligence Agency:
665 https://www.cia.gov/the-world-factbook/, 2021).
666 36 R Core Team. *R: A language and environment for statistical computing* (R Foundation for
667 Statistical Computing, 2021).
668 37 The World Bank. GDP per capita, PPP (current international $) (The World Bank, 2022).
669 38 Fox, J. & Weisberg, S. *An R companion to applied regression, third edition* (Sage, Thousand
670 Oaks CA, 2019).
671 39 Auguie, B. *gridExtra: miscellaneous functions for "grid" graphics. R package version 2.3.*
672 (2017).
673 40 Becker, R. A., Wilks, A. R., Brownrigg, R., Minka, T. P. & Deckmyn, A. *maps: draw*
674 *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).

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, 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

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, V.V.N., 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.

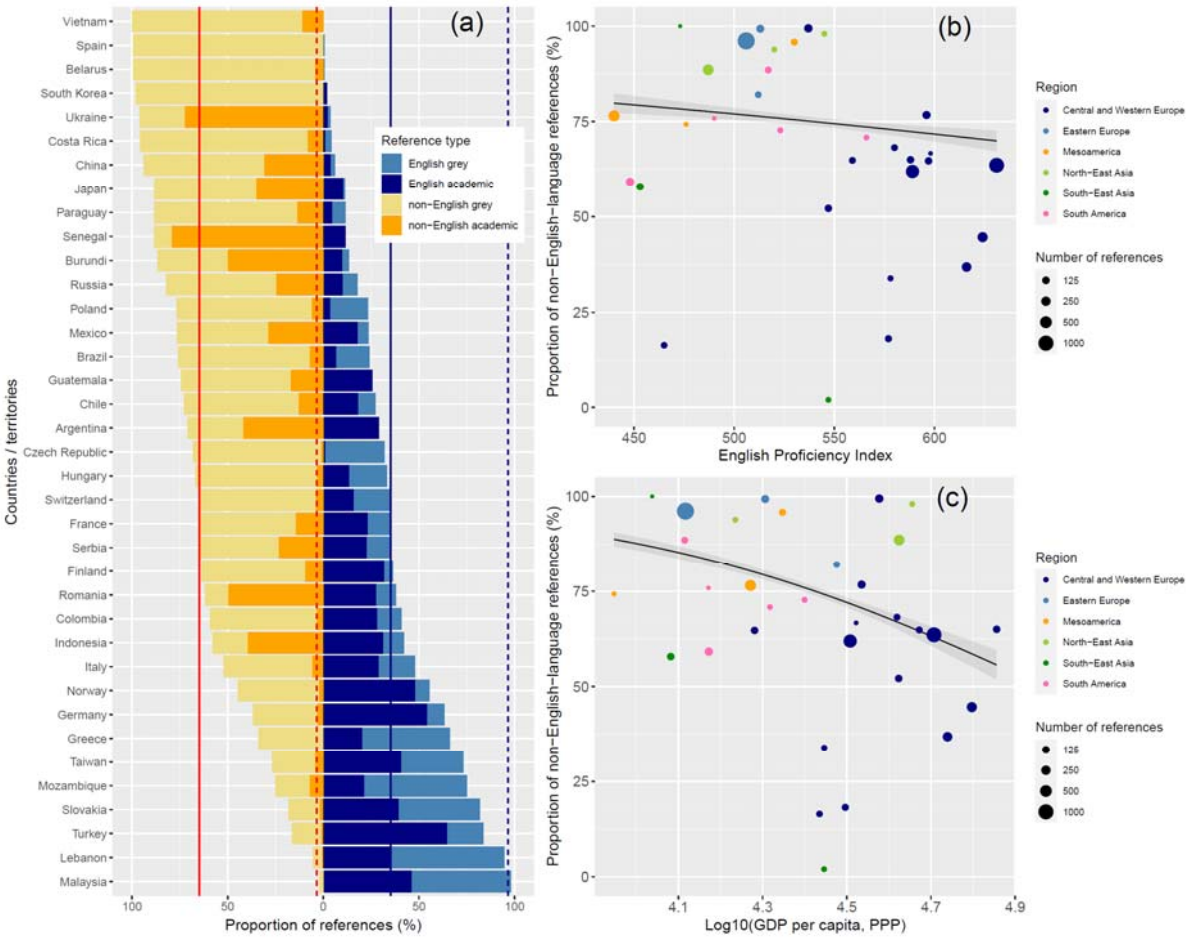
Visualisation: T.A.

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, V.V.N., 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.

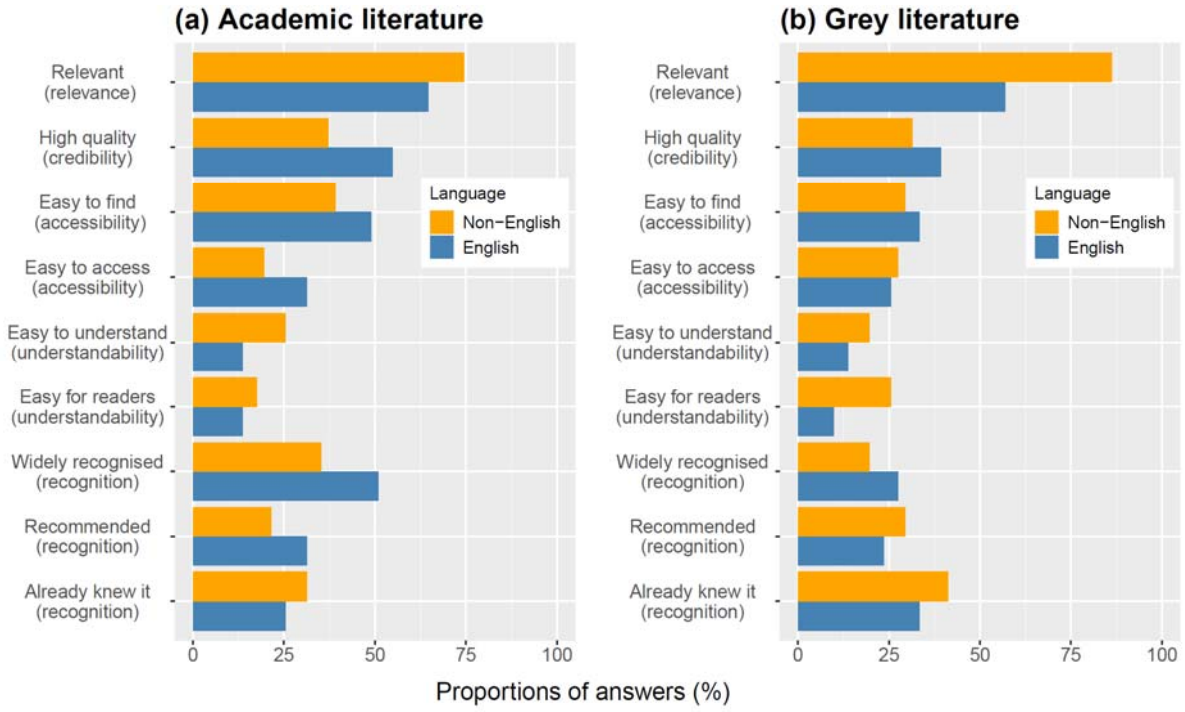
Competing interests

The authors declare no competing interests.



715

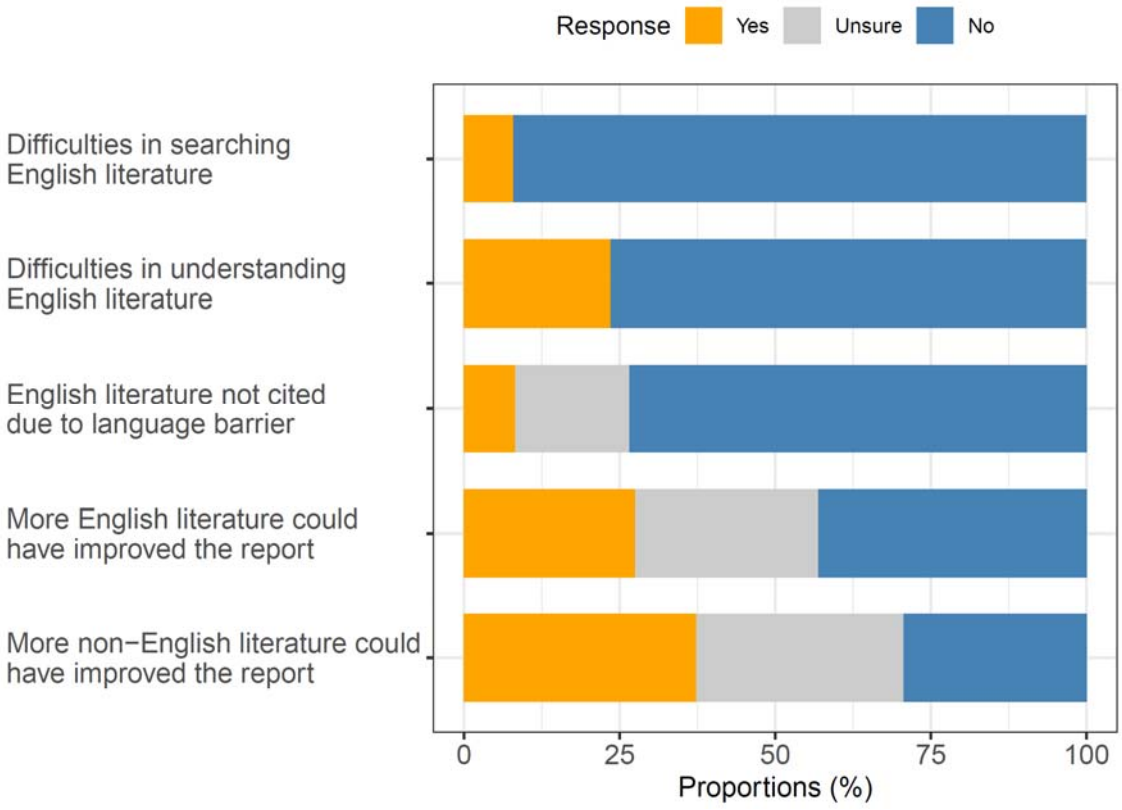
716 **Fig. 1. The proportion of references cited in national biodiversity assessments by language and**
717 **literature type.** (a) The proportion of English-language academic (dark blue) and grey (pale blue)
718 literature, and non-English-language academic (orange) and grey (yellow) literature. The red and blue
719 solid lines indicate the mean proportion of non-English- and English-language references cited in
720 national biodiversity assessments across 37 countries/territories, respectively, while the red and blue
721 broken lines represent the mean proportion of non-English- and English-language references in the
722 eight biodiversity assessment reports by the Intergovernmental Science-Policy Platform on
723 Biodiversity and Ecosystem Services (IPBES) [10], respectively. The relationship between the proportion
724 of non-English-language references cited (academic and grey literature combined) and (b) the English
725 Proficiency Index (see **Methods** for more details) and (c) gross domestic product (GDP) per capita
726 (based on purchasing power parity (PPP), current international $) of each country. The size of each dot
727 indicates the total number of references cited in the report. The colours indicate regions (subregions
728 defined by the IPBES[11]). The regression curves (and 95% confidence intervals as shaded areas) are
729 based on the fitted generalised linear models with a binomial distribution (see Table S1).



730

731 **Fig. 2. Reasons for citing English- and non-English-language (a) academic and (b) grey**
732 **literature in national biodiversity assessments.** The authors of national biodiversity assessments
733 were allowed to select multiple reasons. The x-axis shows the proportion of the report authors who
734 selected each reason. See Questions 6-9 in Supplementary Text S1 for the full description of each
735 reason. Answers were collected from 51 authors in 35 countries/territories (we could not collect
736 answers from the report authors in Burundi and Serbia).

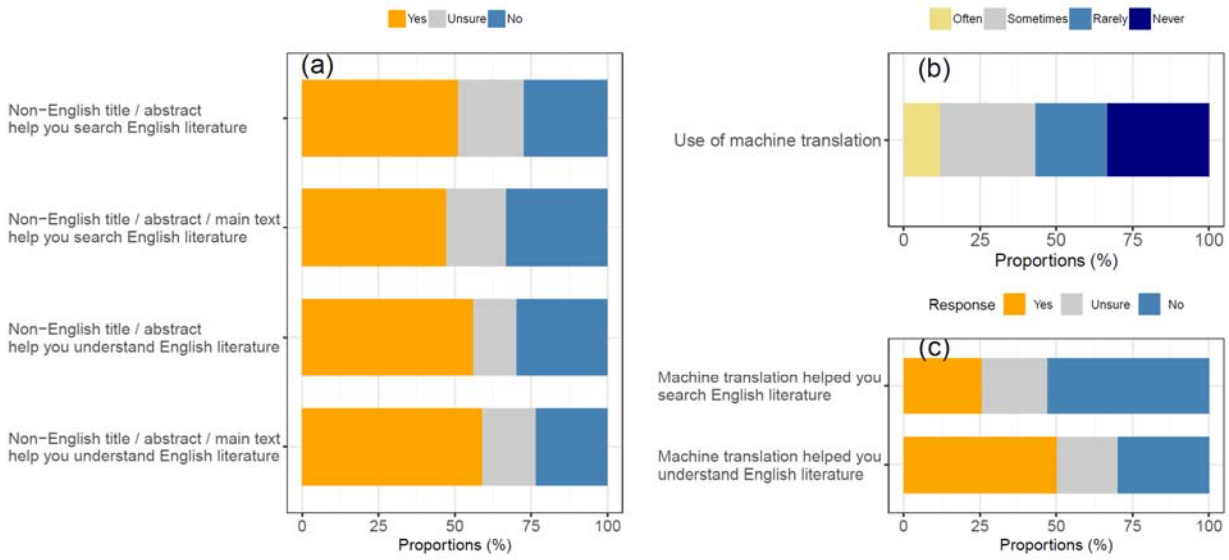
737



738

739 **Fig. 3. Proportions of authors of national biodiversity assessment reports who have experienced**
740 **English language barriers.** Those who have experienced difficulties in searching (n = 50),
741 understanding (n = 51) English-language literature, those who could not cite English-language
742 literature due to difficulties in understanding (n = 49), and those who recognised that citing more
743 English-language or non-English-language literature could have improved their reports (n = 51).

744



745

746

747 **Fig. 4. Potential solutions to facilitating the use of English-language literature.** (a) The proportion
748 of report authors who indicated that a non-English-language title, abstract, and main text of English-
749 language literature would help them search and understand English-language literature. (b) The
750 frequency of use of machine translation when searching and/or reading English-language literature for
751 the reports. Note that no report authors selected "Always" and so this option is now shown. (c) The
752 proportion of report authors who indicated that machine translation helped them search and understand
753 English-language literature. Answers were collected from 51 authors in 35 countries/territories (we
754 could not collect answers from the report authors in Burundi and Serbia), apart from two questions
755 ("Non-English title/abstract help you understand English literature" in (a) and "Machine translation
756 helped you understand English literature" in (c)) where answers were available only from 50 authors.

757

Supplementary Information

Table S1. Results of generalised linear models (with binomial distribution) of factors explaining variations in the proportion of non-English-language references and academic literature. Significant results are shown in bold. EPI: English Proficiency Index. GDP per capita: gross domestic product per capita (based on purchasing power parity, current international $).

Response variable	Explanatory variable	Coefficient	Standard error	z	P
Proportion of non-English-language references	Intercept	11.53	0.59		
	EPI	**-0.0028**	**0.00070**	**-3.98**	**6.83×10^{-5}**
	Log_{10} (GDP per capita)	**-2.02**	**0.18**	**-11.20**	**$< 2.0 \times 10^{-16}$**
Proportion of non-English-language academic literature	Intercept	23.44	0.88		
	EPI	0.0017	0.00099	1.74	0.082
	Log_{10} (GDP per capita)	**-5.42**	**0.27**	**-20.33**	**$< 2.0 \times 10^{-16}$**

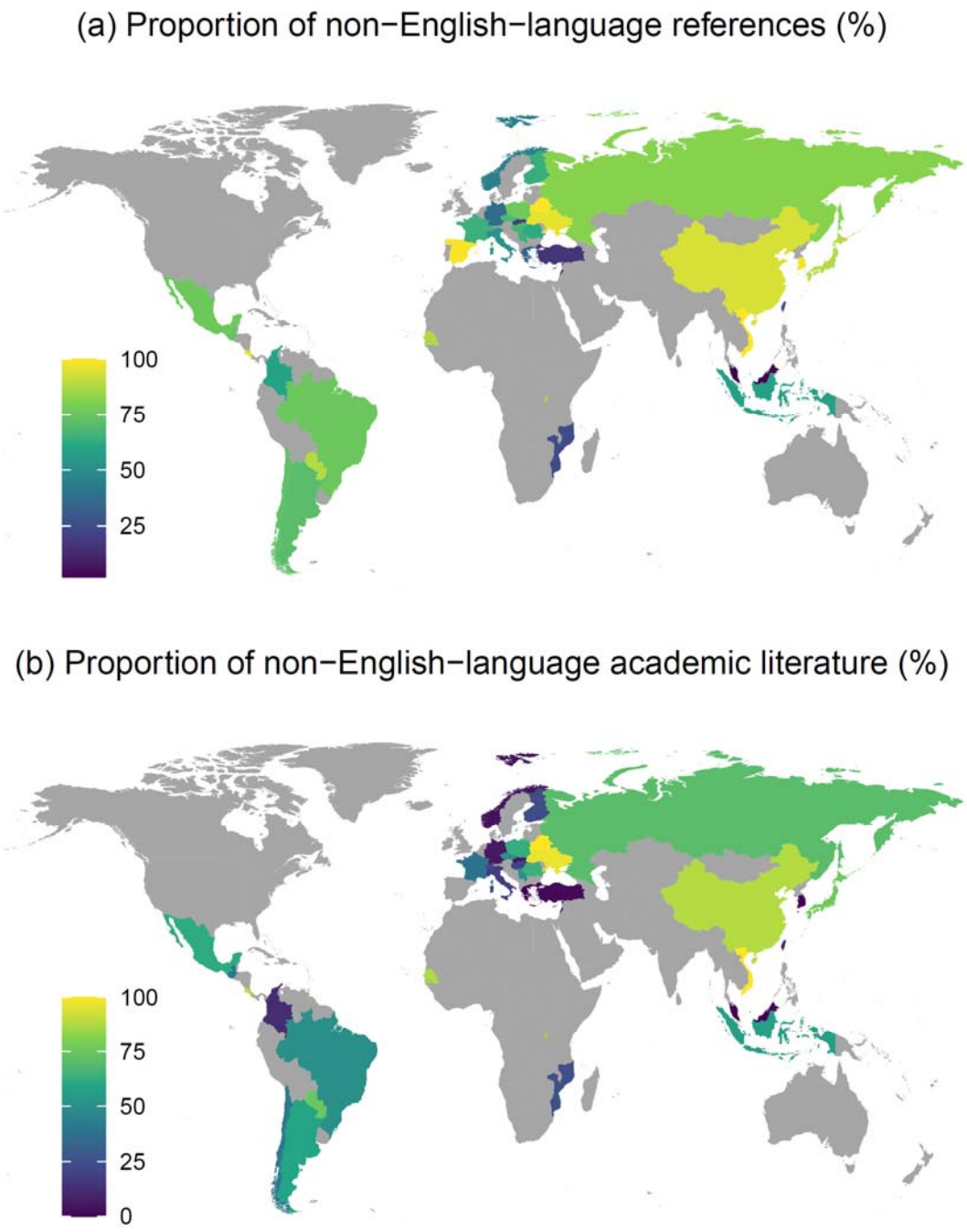
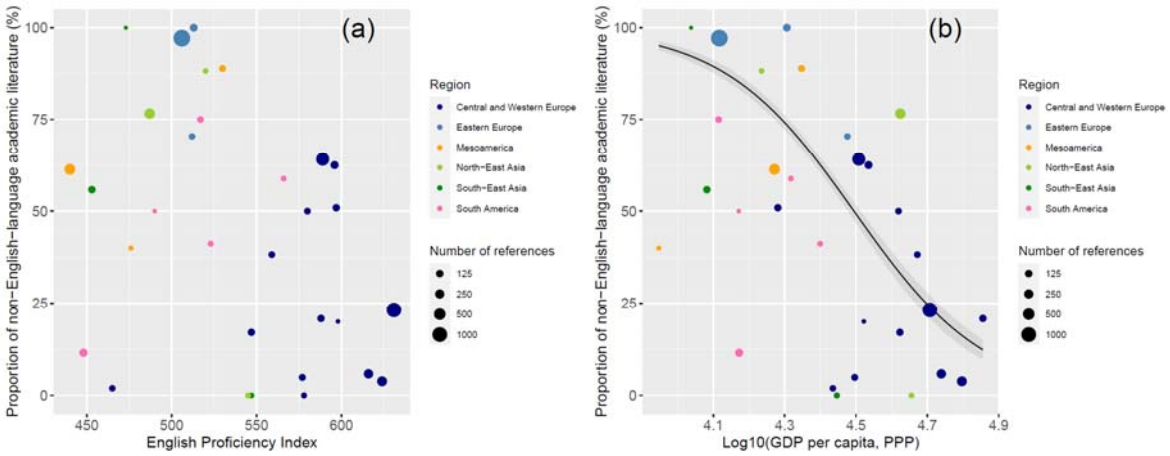


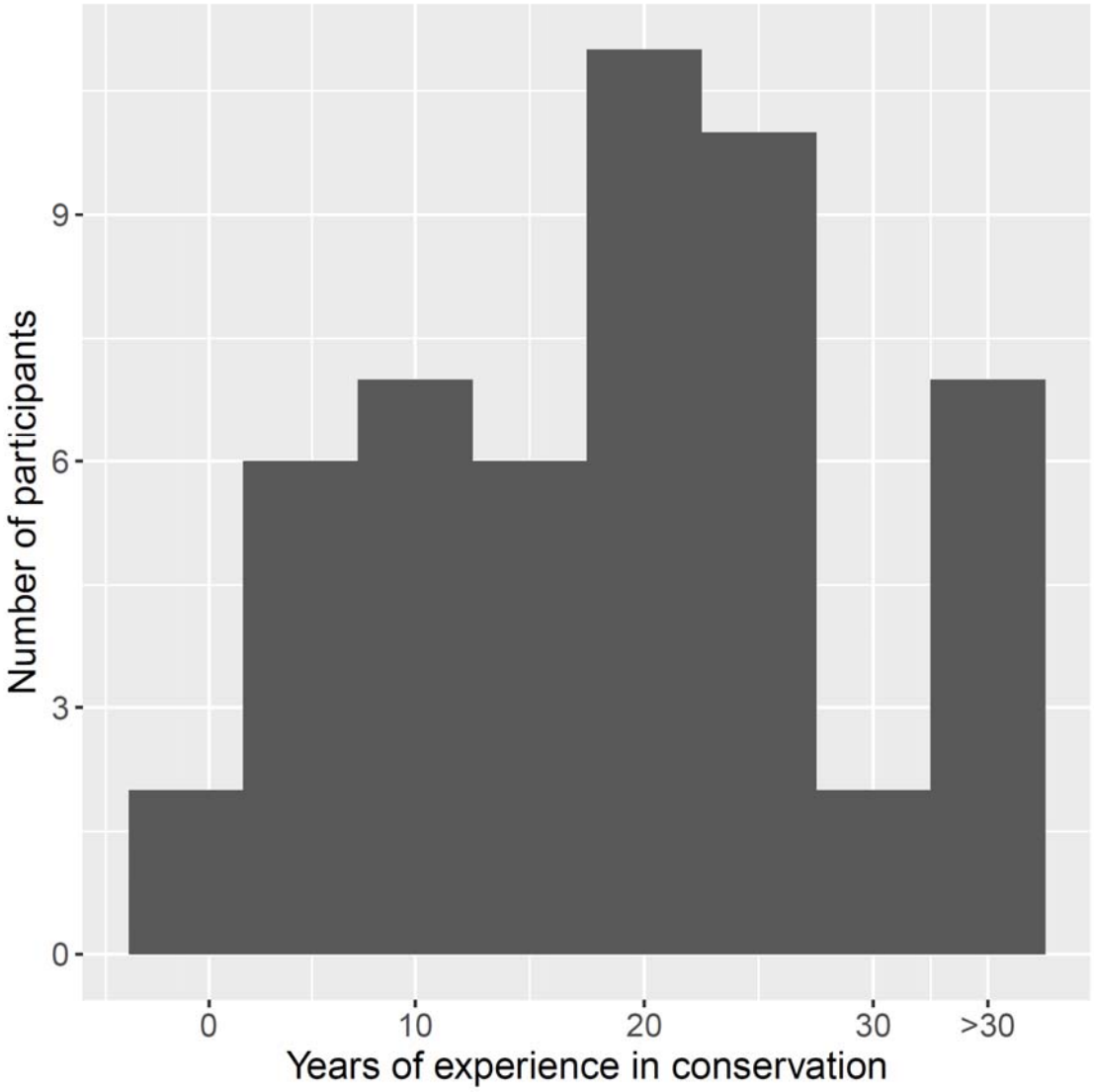
Fig. S1. Geographic variations in the proportion of non-English-language references cited in national biodiversity assessments. (a) The proportion of non-English language references (academic and grey literature combined) and (b) the proportion of non-English-language academic literature in each country. Countries without any records are shown in grey.



773

774 **Fig. S2. Factors associated with the proportion of non-English-language academic literature**
775 **cited in national biodiversity assessments.** The relationship between the proportion of non-English-
776 language academic literature and (a) the English Proficiency Index (see **Methods** for more details) and
777 (b) gross domestic product (GDP) per capita (based on purchasing power parity (PPP), current
778 international $) of each country. The size of each dot indicates the total number of academic literature
779 cited in the report. The colours indicate regions (subregions defined by the IPBES[11]). The regression
780 curve (and 95% confidence interval as a shaded area) in (b) is based on the fitted generalised linear
781 model with a binomial distribution (see Table S1).

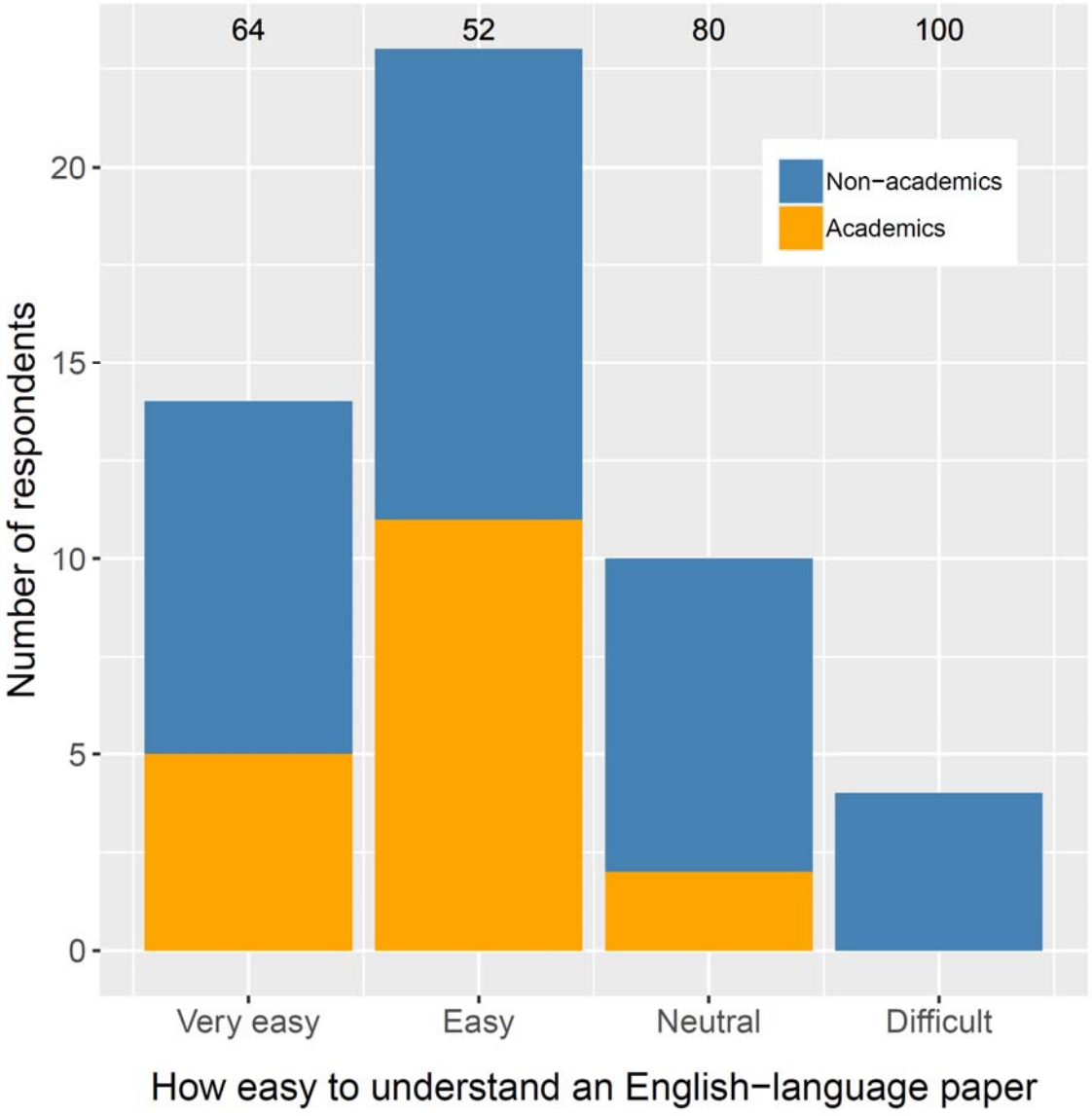
782



783

784 **Fig. S3. The number of years the report authors have been involved in conservation (either in**
785 **on-ground management, research, or policy advice, or any combination).** Data were collected
786 from 51 report authors in 35 countries/territories.

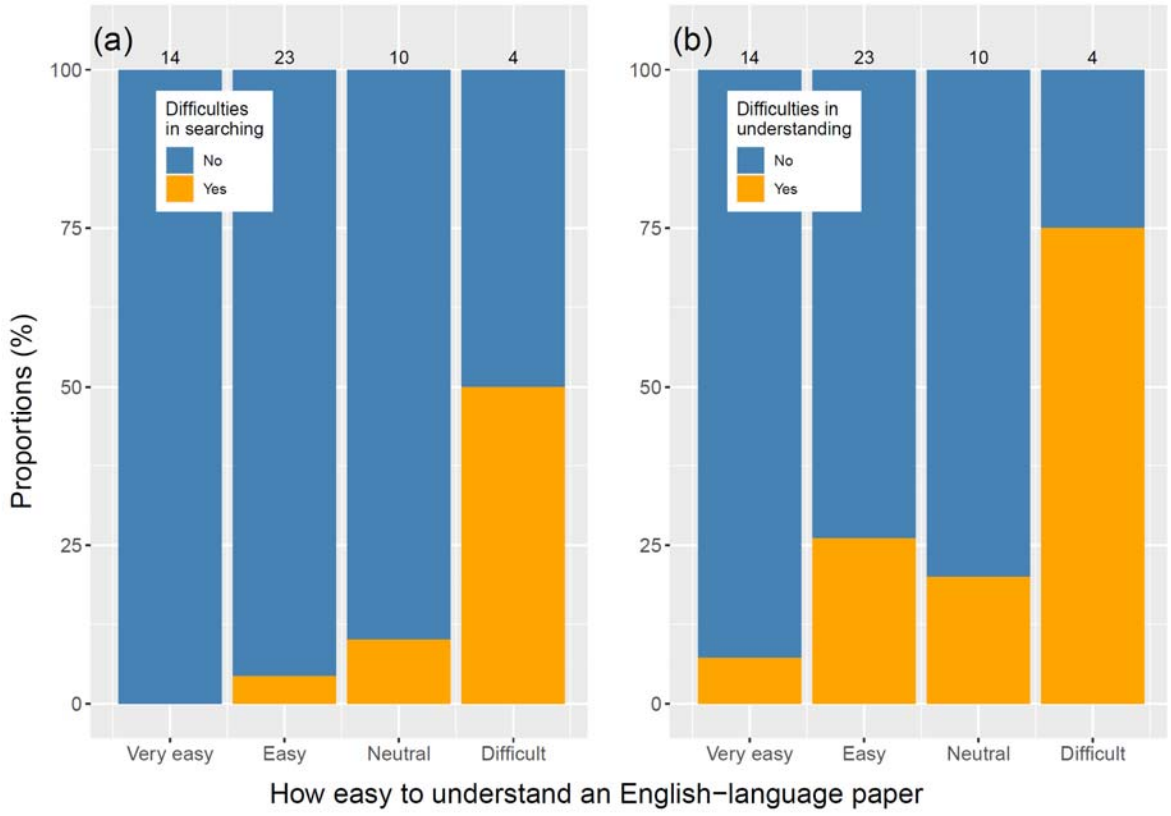
787



788

789 **Fig. S4. Self-reported English proficiency of the 51 report authors in 35 countries/territories**. The
790 report authors were asked to answer how easy it is for them to read and understand the full text of an
791 English-language peer-reviewed paper on biodiversity conservation, based on five options: “Very
792 easy”, “Easy”, “Neutral”, “Difficult” and “Very difficult”. Note that no authors selected “Very
793 difficult”, which therefore is excluded from this figure. Orange indicates answers by academics (i.e.,
794 those who chose “Academic institution or university” in Question 1 of Supplementary Text S1) and
795 blue by all others. Numbers above bars are the percentage of non-academic survey respondents in each
796 category of English proficiency.

797



798

799 **Fig. S5. English-language barriers encountered by report authors across their self-reported**
800 **English proficiency levels**. The proportion of report authors who (a) experienced difficulties in
801 searching (n = 51) and (b) understanding (n = 51) English-language literature for their report because
802 the source was written in English, and its association with their self-reported English proficiency
803 (based on five options: “Very easy”, “Easy”, “Neutral”, “Difficult” and “Very difficult” to read and
804 understand the full text of an English-language peer-reviewed paper on biodiversity conservation: note
805 that no authors selected “Very difficult”, which therefore is excluded from this figure). Numbers
806 above bars are the number of survey respondents in each category of English proficiency.

807

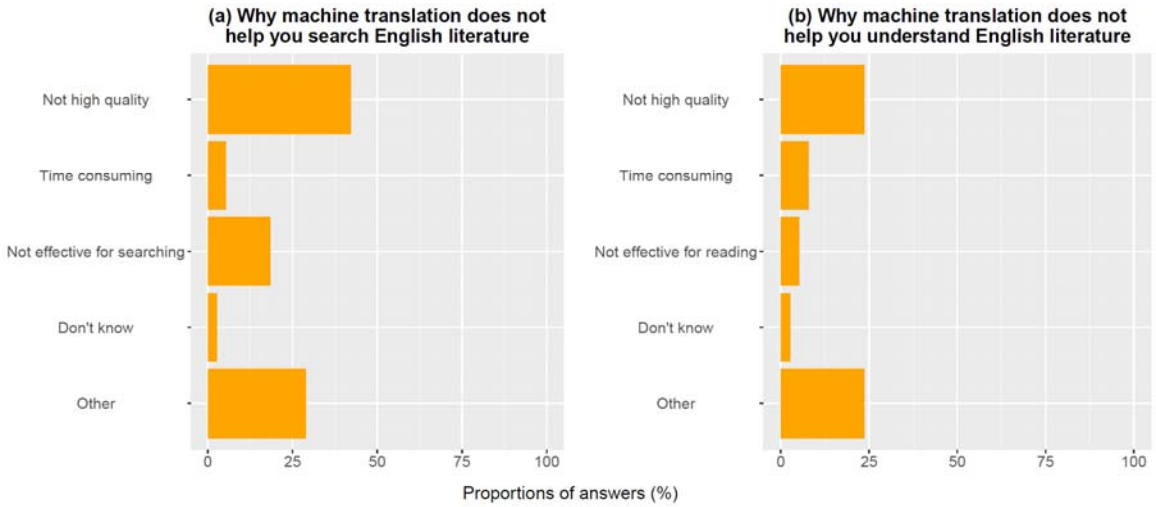


Fig. S6. Reasons why machine translation does not help report authors (a) search or (b) understand English-language literature. Answers were collected from (a) 38 and (b) 26 report authors who answered either "No" or "Unsure" to Questions 23 (*Do you think that machine translation helps you search relevant English-language literature for your report?*) and 25 (*Do you think that machine translation helps you understand relevant English-language literature for your report?)* in Supplementary Text S1, respectively (shown in Fig. 4c).

Supplementary Discussion

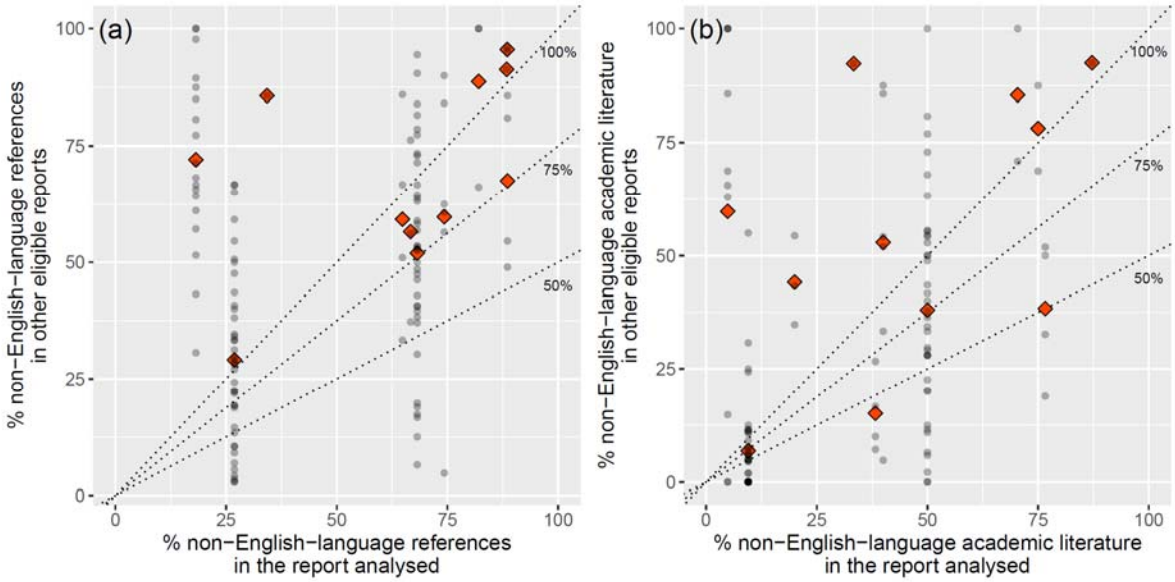
To investigate the use of references written in different languages, we selected the most relevant report in each country/territory based on pre-defined criteria (i.e., 37 reports in total). However, considering the possibility that the selected report for each country/territory may not be representative of other eligible reports in the country/territory in terms of English-language literature usage, we conducted a sensitivity analysis, focusing on 11 countries/territories (Czech Republic, France, Guatemala, Hungary, Japan, Paraguay, Romania, Russia, Senegal, Slovakia, and Taiwan) where we had also recorded the number, type and language of references cited in other eligible reports that were not used in the analysis (the number of such reports ranged from one to 40, with a median of four).

The proportion of non-English-language literature cited varied greatly among reports within each country (Fig. S7). However, in six out of 11 countries, the mean proportion of non-English-language references (academic and grey literature combined) cited in other eligible reports was higher than the proportion in the report used in the analysis (i.e., red diamonds are above the line labelled "100%" in Fig. S7a). In the remaining five countries, the mean proportion of non-English-language references in other eligible reports was lower than that in the report used in the analysis but only by up to 25% (i.e., red diamonds are above the line labelled "75%" in Fig. S7a).

Similarly, the mean proportion of non-English-language academic literature cited in other eligible reports was higher than the proportion in the report used in the analysis in seven out of 11 countries, lower but only by up to 25% in two countries, lower but by within 50% in a country, and lower by more than 50% in one country (Fig. S7b).

This result indicates that in most countries, the proportion of non-English-language literature cited in the reports used in the analysis does not necessarily overestimate the proportions in other eligible reports. In over half of the countries investigated here, the prevalence of non-English-language literature was even higher in other eligible reports. In most of the other countries, the mean proportion of non-English-language literature in other eligible reports was only slightly lower than, and within 25% of, the proportion in the report used in the analysis.

We thus believe that our main conclusion that the use of non-English-language literature in national biodiversity assessments is widespread, is not sensitive to the choice of reports used in the analysis.



845

846 **Fig. S7. The difference in the proportion of non-English-language literature cited between the**
847 **reports used in the analysis and other eligible reports in each of the 11 countries of focus.** The
848 comparison of proportions of (a) non-English-language references (academic and grey literature
849 combined) and (b) non-English-language academic literature. Grey dots represent values in each
850 report and red diamonds represent the mean value in each country.

851

852 **Supplementary Data S1 (separate file)**. **List of 333 biodiversity assessment reports identified in**
853 **37 countries/territories.** The explanations of column names are as follows: Country/territory:
854 country/territory where the report was published, Non-English title: report title in the non-English
855 language, English title: report title in English (if available), Publication language: the language of
856 publication, Used in analysis: YES for the 37 reports used in the analysis, Organisation(s) that
857 edited/published the report: organisations that edited or published the report, Publication year:
858 publication year, Topic: broad topic covered by the report, Citing non-English language references:
859 whether the report cited at least one non-English-language reference, Citing at least 15 references:
860 whether the report cited at least 15 references in total, URL: link to the report.
861
862 **Supplementary Data S2 (separate file). List of 37 biodiversity assessment reports used for the**
863 **analysis, with the numbers of references by category and language.** The explanations of column
864 names are as follows: Country/territory: country/territory where the report was published, Language:
865 the language of publication, Report name: report title in the non-English language, English-language
866 academic literature: the number of English-language academic literature cited, English-language grey
867 literature: the number of English-language grey literature cited, non-English-language academic
868 literature: the number of non-English-language academic literature cited, non-English-language grey
869 literature: the number of non-English-language grey literature cited, EPI: English Proficiency Index,
870 GDPpercapita: gross domestic product per capita (based on purchasing power parity, current
871 international $), Region: regions defined by the IPBES[11], Sub-region: sub-regions defined by the
872 IPBES[11], Number of authors/editors contacted: the number of the report authors/editors contacted,
873 Number of authors/editors who participated: the number of the report authors/editors who participated
874 in the survey.
875
876 **Supplementary Data S3 (separate file). List of 130 eligible reports in 11 countries, used for the**
877 **sensitivity analysis.** Details of column names are as follows: Country/territory: country/territory
878 where the report was published, Used in analysis: YES for the 11 reports used in the analysis,
879 Publication language: the language of publication, Publication year: publication year, Non-English
880 title: report title in the non-English language, English title: report title in English (if available), Topic:
881 broad topic covered by the report, Organisation(s) that edited/published the report: organisations that
882 edited or published the report, English_academiclit: the number of English-language academic
883 literature cited, English_ greylit: the number of English-language grey literature cited,

NonEnglish_academiclit: the number of non-English-language academic literature cited,

NonEnglish_greylit: the number of non-English-language grey literature cited, URL: link to the report.

Supplementary Text S1. Questionnaire survey on the use of references in different languages in national biodiversity assessment reports.

Participant Information Sheet

Survey on the use of scientific literature in domestic reports on biodiversity and its conservation

The purpose of the study

Our recent work showed that up to 35% of scientific literature on biodiversity conservation is published in languages other than English (Amano et al 2016 PLOS Biology). Nevertheless, it is still largely unknown how such non-English-language literature has been used in environmental evidence syntheses at global (e.g., in assessments conducted by the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services) and national levels (e.g., in domestic reports).

This study aims to understand the use of English-language and non-English-language scientific literature in domestic reports on biodiversity and its conservation, published in countries where English is not widely spoken. We have already identified such domestic reports on biodiversity and its conservation in various countries, including those that you authored/edited/published. We would thus be grateful if you could fill in this survey to help us understand how and why references in different languages were identified and cited in the report you authored/edited/published.

What is involved?

Participation in this study is entirely online and will take approximately 15 minutes and the survey can be undertaken at a time and place that is convenient to you.

Participation and withdrawal

Participation in this study is completely voluntary and you are free to withdraw from this study at any time without prejudice or penalty. If you wish to withdraw, simply stop completing the survey and decide not to send it back to the person who contacted you. Feel free to ask any questions

about the research (contact the project coordinator).

Risks

Participation in this study should involve no physical or mental discomfort, and no risks beyond those of everyday living. If, however, you should find any question to be invasive or offensive, you are free to omit answering or participating in that aspect of the study.

Confidentiality and security of data

Your responses to the survey are anonymous; no identifying information will be collected. All other data will be stored on password protected computers and only members of the research team will have access to the data. Because all data is non-identifiable, it cannot be linked to individual participants and data will only be presented as summaries of overall responses. The data you provide will only be used for the specific research purposes of this study.

Benefits of your participation in the study:

The data from the survey will shed light on the role of non-English-language literature in domestic policy-making for biodiversity conservation, as well as consequences of language barriers to the use of English-written scientific knowledge in domestic policy-making.

Consent form

Please take the time to read the project information that is provided on the previous page. Your participation is voluntary, and you can choose to withdraw at any point. You will not be asked to give your name so any information you provide is completely anonymous. Should you wish to clarify any aspect of your potential participation or need more information you can also speak directly to a lead researcher before agreeing or disagreeing to take part in the evaluation.

If you understand the purpose of the research project and the nature of your involvement, then please complete the following:

☐ I have read the information provided about the research project and understand the nature of my involvement. I understand any information I provide is completely confidential. I agree to take part and understand I can withdraw at any time.

☐ I am over 18 years of age.

Ethics Clearance and Contacts

This study has been cleared in accordance with the ethical review guidelines and processes of the University of Queensland. These guidelines are endorsed by the University's Human Ethics Committee and registered with the Australian Health Ethics Committee as complying with the National Statement. You are free to discuss your participation in this study with project staff (contactable at t.amano@uq.edu.au). If you would like to speak to an officer of the University not involved in the study, you may contact the University of Queensland Ethics Officer on 07 3365 3924.

If you would like to learn the outcome of the study in which you are participating, please feel free to email t.amano@uq.edu.au and we can organise to send you a summary of the study once it is complete. You can also obtain general information on the project at:
https://translatesciences.com/.

Thank you for your participation in this study.

Dr Tatsuya Amano, ARC Future Fellow
School of Biological Sciences
The University of Queensland
Brisbane, Qld 4072
Email: t.amano@uq.edu.au

Violeta Berdejo Espinola, Senior Research Technician
School of Biological Sciences
The University of Queensland
Brisbane, Qld 4072
Email: v.berdejoespinola@uq.edu.au

Questionnaire survey on the use of scientific literature in national reports on biodiversity and its conservation

Section A

Q1. Which sector do you work in (please choose only one)?

National government ☐

State/provincial/regional government ☐

Local government ☐

Government research institution ☐

Not-for-profit organization ☐

Private business ☐

Academic institution or university ☐

Other (please describe): _____

Q2. Which aspect of biodiversity conservation does your role predominantly focus on (please choose only one)?

Policy ☐

On-ground management ☐

Research ☐

Other (please describe, e.g., if your role spans across these aspects):

_____

Q3. Please circle the (approximate) number of years you have been involved in conservation (either in on-ground management, research, or policy advice, or any combination)

<1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 >30

Q4. Please state your first language(s): here a first language is defined as "the language that you spoke most at home when you were at age 3-18".

Q5. How easy is it for you to read and understand the full text of an English-language peer-reviewed paper on biodiversity conservation?

Very easy	☐
Easy	☐
Neutral	☐
Difficult	☐
Very difficult	☐

Section B.

Q6. If you cited **non-English-language peer-reviewed literature** (i.e., papers published in academic peer-reviewed journals) in your report, please select reason(s) for citing it/them (select all that apply to the majority of the non-English-language peer-reviewed literature cited).

It was easy to find when searching for relevant literature	☐
It was easy to access (e.g., open access, hard copies available, had a journal subscription)	☐
The content was easy to understand to you and/or other editors of the report	☐
It would be easy for the readers of the report to understand	☐
It was recommended by someone else (e.g., experts in a relevant field)	☐
You already knew about it before writing the report	☐
It was from a widely-recognised source (e.g., published in a well-known academic journal)	☐
It contained information you were specifically looking for (e.g., important information on particular species in the country)	☐
Information provided in the literature was of high quality and thus reliable (e.g., based on a rigorous study)	☐
I did not cite non-English-language peer-reviewed literature in my report	☐

Other (please describe):_____

Q7. If you cited **non-English-language grey literature** (e.g., governmental/NGO reports, databases, websites, theses etc) in your report, please select reason(s) for citing it/them (select all that apply to the majority of the non-English-language grey literature cited).

- It was easy to find when searching for relevant literature ☐
- It was easy to access (e.g., open access, hard copies available, had a journal subscription) ☐
- The content was easy to understand to you and/or other editors of the report ☐
- It would be easy for the readers of the report to understand ☐
- It was recommended by someone else (e.g., experts in a relevant field) ☐
- You already knew about it before writing the report ☐
- It was from a widely-recognised source (e.g., published in a well-known academic journal) ☐
- It contained information you were specifically looking for ☐
  (e.g., important information on particular species in the country)
- Information provided in the literature was of high quality and thus reliable ☐
  (e.g., based on a rigorous study)
- I did not cite non-English-language grey literature in my report ☐
- Other (please describe):_____

Q8. If you cited **English-language peer-reviewed literature** (i.e., papers published in academic peer-reviewed journals) in your report, please select reason(s) for citing it/them (select all that apply to the majority of the English-language peer-reviewed literature cited).

- It was easy to find when searching for relevant literature ☐
- It was easy to access (e.g., open access, hard copies available, had a journal subscription) ☐

The content was easy to understand to you and/or other editors of the report ☐

It would be easy for the readers of the report to understand ☐

It was recommended by someone else (e.g., experts in a relevant field) ☐

You already knew about it before writing the report ☐

It was from a widely-recognised source (e.g., published in a well-known academic journal) ☐

It contained information you were specifically looking for ☐
(e.g., important information on particular species in the country)

Information provided in the literature was of high quality and thus reliable ☐
(e.g., based on a rigorous study)

I did not cite English-language peer-reviewed literature in my report ☐

Other (please describe):_____

Q9. If you cited **English-language grey literature** (e.g., governmental/NGO reports, databases, websites, theses etc) in your report, please select reason(s) for citing it/them (select all that apply to the majority of the English-language grey literature cited).

It was easy to find when searching for relevant literature ☐

It was easy to access (e.g., open access, hard copies available, had a journal subscription) ☐

The content was easy to understand to you and/or other editors of the report ☐

It would be easy for the readers of the report to understand ☐

It was recommended by someone else (e.g., experts in a relevant field) ☐

You already knew about it before writing the report ☐

It was from a widely-recognised source (e.g., published in a well-known academic journal) ☐

It contained information you were specifically looking for ☐
(e.g., important information on particular species in the country)

Information provided in the literature was of high quality and thus reliable ☐
(e.g., based on a rigorous study)

I did not cite English-language grey literature in my report ☐

Other (please describe):_____

Q10. Did you encounter any difficulties when **searching for English-language literature** for your report because the source was written in English? (e.g., difficult to understand how to use a literature search engine)

Yes ☐
No ☐

If yes, please describe your difficulties:

Q11. Did you encounter any difficulties when trying to **understand English-language literature** for your report because it was written in English? (e.g., difficult to understand a paper written in English)

Yes ☐
No ☐

If yes, please describe your difficulties:

Q12. Was there any **English-language literature** that you found or knew already that looked relevant to your report but you decided not to cite because you found it difficult to understand the content written in English?

Yes ☐

No ☐

Unsure ☐

Q13. Do you think that your report could be improved (i.e., more detailed, more accurate, better quality, or better coverage) if you had used more **English-language references**?

Yes ☐

No ☐

Unsure ☐

Q14. Do you think that your report could be improved (i.e., more detailed, more accurate, better quality, or better coverage) if you had used more **non-English-language references**?

Yes ☐

No ☐

Unsure ☐

Q15. What type of review did you use to search for **non-English-language references** cited in your report? (select all that apply)

Personal knowledge ☐

(e.g., only cited references that you and/or other editors of the report knew)

Expert opinion ☐
(e.g., most references recommended by a limited number of experts
other than those who authored/edited the report)

Formal consultation process ☐
(e.g., widely asked for the identification of relevant literature
from a wider expert community)

Casual or narrative review ☐
(e.g., looked at several relevant studies and used literature cited in those studies)

Systematic review ☐
(e.g., systematically screened all relevant literature
with certain keywords in one or more literature search engine)

Other (please describe):

Q16. What type of review did you use to search for **English-language references** cited in your report? (select all that apply)

Personal knowledge ☐
(e.g., only cited references that you and/or other editors of the report knew)

Expert opinion ☐
(e.g., most references recommended by a limited number of experts
other than those who authored/edited the report)

Formal consultation process ☐
(e.g., widely asked for the identification of relevant literature
from a wider expert community)

Casual or narrative review ☐

(e.g., looked at several relevant studies and used literature cited in those studies)

Systematic review ☐

(e.g., systematically screened all relevant literature

with certain keywords in one or more literature search engine)

Other (please describe):

Section C.

Q17. If all English-language scientific literature (e.g., papers published in English-language academic journals) had **the title and abstract also available in your first language**, do you think that would have made it easier and quicker to **search and identify** relevant literature for your report?

Yes ☐

No ☐

Unsure ☐

Q18. If all English-language scientific literature (e.g., papers published in English-language academic journals) had the **title, abstract and full text also available in your first language**, do you think that would have made it easier and quicker to **search and identify** relevant literature for your report?

Yes ☐

No ☐

Unsure ☐

Q19. If all English-language scientific literature (e.g., papers published in English-language academic journals) had the **title and abstract also available in your first language**, do you think that would have made it easier and quicker to **understand** relevant literature for your report?

Yes ☐
No ☐
Unsure ☐

Q20. If all English-language scientific literature (e.g., papers published in English-language academic journals) had the **title, abstract and full text also available in your first language**, do you think that would have made it easier and quicker to **understand** relevant literature for your report?

Yes ☐
No ☐
Unsure ☐

Q21. Did you use machine translation (e.g., Google Translate) when searching and/or reading English-language literature for your report?

Always (~100%) ☐
Often (~75%) ☐
Sometimes (~50%) ☐
Rarely (~25%) ☐
Never (~0%) ☐

Q22. If you used machine translation when searching and/or reading English-language literature for your report, please state the translation service you used.

Q23. Do you think that machine translation (e.g., Google Translate) helps you **search** relevant English-language literature for your report?

Yes ☐

No ☐

Unsure ☐

Q24. If you do **NOT** think that machine translation helps you **search** relevant English-language literature, please select reason(s) for why you think so (select all that apply):

Machine translation is not high quality enough (e.g., it doesn't translate well) ☐

It is time-consuming to use ☐
(e.g., need to copy and paste relevant sentences into the service)

It cannot be effectively used for searching literature ☐
(e.g., not integrated into literature search engines)

I don't know how to use it ☐

Other (please describe):

Q25. Does machine translation (e.g., Google Translate) help you **understand** relevant English-language literature

Yes ☐

No ☐

Unsure ☐

Q26. If you do **NOT** think that machine translation helps you **understand** relevant English-language literature, please select reason(s) for why you think so (select all that apply):

Machine translation is not high quality enough (e.g., it doesn't translate well) ☐

It is time-consuming to use ☐
(e.g., need to copy and paste relevant sentences into the service)

It cannot be effectively used for reading literature ☐
(e.g., it cannot be used for hard copies)

I don't know how to use it ☐

Other (please describe):

Do you have any other comments about the use of English- or non-English-language literature for your work in biodiversity conservation and management?

Thank you very much for your time!

Please visit our website (https://translatesciences.com/) for the detail of our translatE project.