Contributed Paper

Characterizing efforts to reduce consumer demand for wildlife products

Diogo Veríssimo (iD)[1,2,3]* and Anita K. Y. Wan[4]

[1]Department of Zoology, University of Oxford, Zoology Research and Administration Building, 11a Mansfield Road, Oxford, OX1 3SZ, U.K.
[2]San Diego Zoo Institute for Conservation Research, 15600 San Pasqual Valley Road, Escondido, CA, 92027, U.S.A.
[3]Oxford Martin School, Oxford Martin Programme on the Illegal Wildlife Trade, University of Oxford, 34 Broad Street, Oxford, OX1 3BD, U.K.
[4]School of Life Sciences, Sun Yat-Sen University, 135 West Xingang Road, Guangzhou, 510275, China

Abstract: *The unsustainable trade in wildlife is a key threat to Earth's biodiversity. Efforts to mitigate this threat have traditionally focused on regulation and enforcement, and there is a growing interest in campaigns to reduce consumer demand for wildlife products. We aimed to characterize these behavior-change campaigns and the evidence of their impacts. We searched peer-reviewed and grey literature repositories and over 200 institutional websites to retrieve information on demand-reduction campaigns. We found 236 campaigns, mainly in the grey literature. Since the 1970s, the number of campaigns increased, although for over 15% a start date could not be found. Asia was the primary focus, although at the national level the United States was where most campaigns took place. Campaigns most often focused on a single species of mammal; other vertebrates groups, with the exception of sharks, received limited attention. Many campaigns focused on broad themes, such as the wildlife trade in general or seafood. Thirty-seven percent of campaigns reported some information on their inputs, 98% on strategies, 70% on outputs, 37% on outcomes (i.e., changes in the target audience), and 9% on impacts (i.e., biological changes or threat reduction). Information on outcomes and impacts was largely anecdotal or based on research designs that are at a high risk of bias, such as pre- and postcampaign comparisons. It was unclear whether demand-reduction campaigns had direct behavioral or biological impacts. The lack of robust impact evaluation made it difficult to draw insights to inform future efforts, a crucial part of effectively addressing complex issues, such as the wildlife trade. If demand-reduction campaigns are to become a cornerstone of the efforts to mitigate the unsustainable trade in wildlife, conservationists need to adopt more rigorous impact evaluation and a more collaborative approach that fosters the sharing of data and insights.*

Keywords: behavior change, demand reduction, illegal, impact evaluation, outreach, social marketing, trafficking, wildlife trade

Caracterización de los Esfuerzos para Reducir la Demanda de Productos de Fauna

Resumen: *El mercado insostenible de fauna es una amenaza importante para la biodiversidad del planeta. Los esfuerzos por mitigar esta amenaza se han enfocado tradicionalmente en la regulación y la ejecución, además de que existe un interés creciente por campañas para reducir la demanda del consumidor por productos de fauna. Buscamos caracterizar estas campañas de cambios de comportamiento y las evidencias de su impacto. Buscamos en los depósitos de literatura gris y aquella revisada por colegas y en más de 200 sitios web institucionales para recuperar información sobre las campañas de reducción de demanda. Encontramos 236 campañas, principalmente en la literatura gris. Desde la década de 1970, el número de campañas incrementó, aunque para más del 15% no pudimos encontrar una fecha de inicio. Asia era el foco principal, aunque a nivel nacional fueron los Estados Unidos en donde se realizaron más campañas. La*

mayoría de las campañas se enfocaban en una sola especie de mamífero; los demás grupos de vertebrados, excepto los tiburones, recibieron atención limitada. Muchas campañas se enfocaron en temas generales, como el mercado de fauna en general o en la industria de los mariscos. El 37% de las campañas reportaron algo de información sobre sus entradas, el 98% sobre sus estrategias, el 70% sobre su producción, el 37% sobre sus resultados (es decir, cambios en el público diana), y el 9% sobre sus impactos (es decir, cambios biológicos o reducción de la amenaza). La información sobre los resultados y los impactos en su mayoría fue anecdótica o basada en diseños de investigación que tienen un alto riesgo de sesgo, como las comparaciones pre- y post-campaña. No quedó claro si las campañas de reducción de demanda tuvieron impactos biológicos o de comportamiento directos. La falta de una evaluación fuerte de los impactos dificulta la obtención de conocimiento para informar los esfuerzos en el futuro, una parte crucial para tratar con eficiencia temas complejos, como el mercado de fauna. Si se busca que las campañas de reducción de demanda se conviertan en una piedra angular de los esfuerzos para mitigar el mercado insostenible de fauna, los conservacionistas necesitan adoptar una evaluación más rigurosa de los impactos y una estrategia más colaborativa que promueva la distribución de datos y conocimiento.

Palabras Clave: cambios en el comportamiento, evaluación de impacto, ilegal, mercado de fauna, mercado negro, mercadotecnia social, participación, reducción de la demanda

摘要: 非可持续性的野生动植物贸易是地球生物多样性的主要威胁。为减缓这种威胁, 传统的策略侧重于监管和执法方面, 而今越来越关注于减少消费者野生动植物产品需求的活动。我们旨在描述这些针对行为改变的活动及其影响的实证。为检索有关减少需求的活动信息, 我们搜索了同行评议文献、灰色文献资料库, 及 200 多个机构网站.从中我們共检索到 236 项活动, 主要来自于灰色文献。尽管有超过 15% 的活动无法找到确切的开始日期, 但自 20 世纪 70 年代以来, 此类活动的数量仍是有所增加的。虽然在国家层面上美国是大多数活动的举办地, 但亚洲是主要的焦点地区。活动通常针对某一哺乳动物; 除鲨鱼外, 其他脊椎动物受到的关注有限。很多活动所侧重的主题比较宽泛, 例如总体的野生动植物贸易或海产食物。有 37% 的活动报告了其投入情况、98% 的介绍了其策略、70% 的总结了其产出情况、37% 的说明了其效果 (例如目标受众的改变)、以及 9% 的报告了其影响 (例如生物上的变化或威胁的减少)。关于效果和影响的信息主要是轶事的或基于具有高度偏差风险的研究设计, 例如活动前后的比较。目前尚不清楚在行为或生物方面减少需求的活动是否具有直接影响。由于缺乏强有力的影响评估, 难以为未来的措施提供见解, 而这正是有效解决诸如野生动植物贸易等复杂问题的关键部分。如果减少需求的活动要成为减缓野生动植物非可持续性贸易的基石, 保护主义者需要采取更严格的影响评估和加强协作的方法, 来促进数据和见解的共享。

关键词: 行为改变, 减少需求, 非法, 影响评估, 普及, 社会营销, 贩卖, 动植物贸易

Introduction

The illegal or unsustainable trade in wildlife is increasingly being recognized as a global threat to biodiversity by conservationists and policy makers (Rosen & Smith 2010). Following the globalization of international trade links, this multibillion-dollar industry has expanded in recent decades (Warchol 2004). The illegal and unsustainable trade in wildlife reaches well beyond well-known wildlife commodities, such as rhinoceros horn and elephant ivory, to large numbers of invertebrate animals, plants, and fungi (Blundell & Mascia 2005). Yet, beyond its substantial and often irreversible impact on biodiversity, illegal wildlife trade can have profound impacts on the human health, economic development, and governance of societies in source and consumer countries (Karesh et al. 2005; Rosen & Smith 2010; Haenlein et al. 2016).

Traditionally, responses to wildlife trafficking have focused on regulation and enforcement, both of which try to tackle the supply side of the trade (Veríssimo et al. 2012; Challender & MacMillan 2014). However, in recent years the emphasis on demand-side management has increased, which aims to reduce the market value of illegal wildlife products by getting consumers to voluntarily change their purchasing behavior (Ayling 2016; Wallen & Daut 2018). At the same time, conservationists have stepped up their consumer-research efforts (e.g., Hinsley et al. 2015; Shairp et al. 2016; Megias et al. 2017) with the goal of better understanding the complex social, cultural, and economic contexts of wildlife-product use.

This increased focus on consumers and demand has led to the launch of numerous campaigns aimed at influencing consumers of wildlife products. These campaigns have been labeled in a variety of ways, from awareness raising and environmental education to human-centered design and social marketing (Olmedo et al. 2018; Wallen & Daut 2018). Yet, understanding of what these different outreach efforts have achieved remains limited. This stems from unclear goals, mismatches between campaign activities and the behaviors that to be influenced, and lack of rigorous project monitoring and evaluation (Olmedo et al. 2018; Veríssimo et al. 2018b; Greenfield & Veríssimo 2019). As a result, claims of success by implementers are often received with skepticism (Roberton 2014). These challenges are not unique to efforts related to wildlife

trade and exist across other areas of biodiversity conservation (Miteva et al. 2012; Roe et al. 2015).

Early this century, the drive for improved project monitoring and evaluation began to gain momentum in conservation practice (Stem et al. 2005). This movement focused mostly on documenting trends in selected indicators (Ferraro & Pattanayak 2006; Curzon & Kontoleon 2016), often along a logic model (Fig. 1) that linked project inputs and activities to outcomes (i.e., changes in the knowledge, attitudes, or behavior of the target audience) and impact (i.e., biological changes or threat reduction) (Rissman & Smail 2015). Under this type of approach, it is often assumed that if the indicator improves, an intervention is being effective, and if the indicator worsens, the opposite is true (Ferraro 2009).

Although project monitoring has become widely adopted across the conservation sector (Rissman & Smail 2015), this type of exercise is not enough to measure the impact on a program or intervention (Ferraro & Pattanayak 2006; Ferraro 2009). To evaluate impact one needs to answer the question, does the intervention work better than no intervention at all (or an alternative) (Ferraro 2009)? This requires assessing the extent to which change can be attributed to a specific project or intervention, rather than to potential biases or confounding factors (Rosenbaum 2010). Thus, impact evaluation is at its core about making inferences about a counterfactual scenario (i.e., what would have happened without said intervention) that can be inferred only indirectly (Ferraro & Pattanayak 2006).

Impact evaluation research can be experimental or observational (Rosenbaum 2010). The key difference is that in the former, the researcher controls the units (e.g., villages or individuals) assigned for treatment or control status, whereas the latter studies are common when random assignment of said treatment is not possible for practical or ethical reasons (Margoluis et al. 2009). Such observational studies are typically divided between qualitative case studies and quantitative quasi experiments (Rosenbaum 2010). There has been particular emphasis in the literature on the need for more experimental and quasi-experimental evaluations in biodiversity conservation as these methods are considered the best way to mitigate biases and confounders (Ferraro 2009; Curzon & Kontoleon 2016).

In the case of behavior-change efforts aimed at reducing demand for illegal or unsustainable wildlife products, what has been undertaken globally, what has been achieved , and what lessons have been learned are relatively unknown. We aimed to characterize the temporal and spatial distribution and biological focus of demand-reduction campaigns; synthesize evidence of project monitoring and evaluation of those campaigns; and generate learning insights to inform future conservation.

Methods

We defined "demand-reduction campaigns" as outreach interventions to get people to voluntarily change their current or potential behavior as consumers of wildlife products or their derivatives (Veríssimo et al. 2012; Burgess 2016; Olmedo et al. 2018). We therefore considered only consumer-focused efforts and not those with, for example, an enforcement or policy focus. We defined "wildlife" as all animals and plants. In cases where multiple interventions overlapped geographically and shared implementing or funding agencies, those with the same conservation goals were grouped and considered as a single campaign.

To understand how much information is available on demand-reduction campaigns implemented to date, we conducted searches on multiple online databases and platforms. From September 2016 to March 2017, we conducted searches examining all results on Thomson Reuters Web of Science, Scopus, and the first 100 results on Google Scholar. In these searches, we used general keywords associated with wildlife trade and species-specific terms related to 43 species listed with CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora) that are often associated with the wildlife trade (Supporting Information). We also included the bluefin tuna (*Thunnus thynnus*), a high-profile species that has repeatedly faced stiff opposition to CITES listing, despite being highly threatened by commercial exploitation. All searches were conducted in English and Chinese (traditional and simplified) keywords.

To improve our ability to capture grey literature, we used Google Search to examine the first 50 hits for each search. Because of the abundance of unrelated content in these searches, which made them less cost effective, we used a restricted set of general keywords related to wildlife trade and demand reduction (Supporting Information). Finally, we visited websites of nongovernmental organization (NGOs), professional groups, and funding organizations with a track record of funding or implementing demand-reduction campaigns. We obtained an initial list through the catalog of demand-reduction campaigns compiled by Sharif (2014) and then used the institutional websites of the identified organizations and snowball sampling to find other institutions working in this area. The goal of this effort was to locate information on additional demand-reduction initiatives (Supporting Information).

From all the sources described above we collected a number of campaign descriptors, namely start year, topical focus, country of implementation, research design used for data collection, and type of data collected for campaign evaluation. Regarding the latter, we classified the available information according to a basic impact evaluation logic model, divided into inputs; strategies;
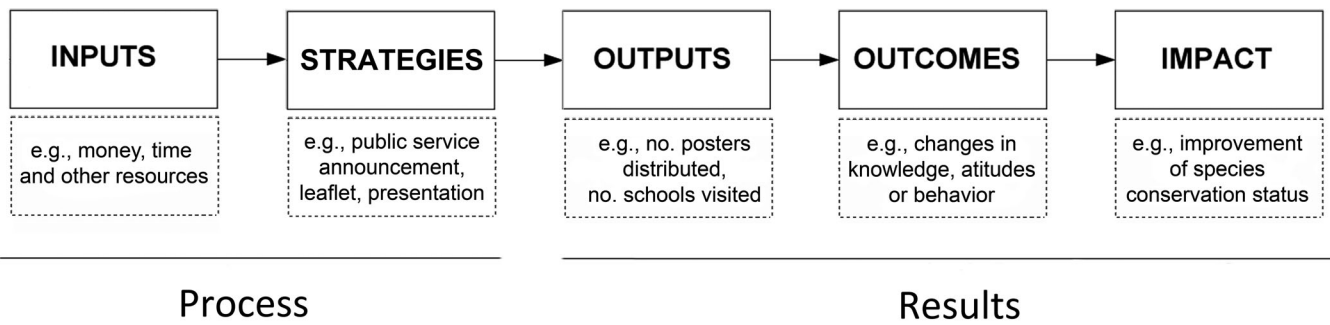
Process | Results

Figure 1. General logic model followed to classify the types of monitoring and evaluation information reported for different outreach campaigns to reduce the demand for wildlife products. Adapted from Margoluis et al. (2009).

outputs; and outcomes and impact (Fig. 1). We classified the inputs as quantified information, partial or complete, on human and financial resources used for campaign implementation, such as on personnel size and grants. Under strategies we considered descriptions of activities or products used as part of the campaign. Regarding outputs we included information on the implementation and use of the previously described strategies, such as measurable data on audience reached by a public service announcement (PSA), individuals trained and recruited for programs, and media reports. For outcomes we considered evidence of specific changes in the target audience, in terms of knowledge, attitudes, or behaviors. For campaign impacts we included biological changes in the target natural resource or species or reduction in conservation threats, such as a reduction in mortality. Where possible, we tried to calculate the effect size of the change reportedly generated by the campaign.

We used chi-square tests to examine trends in the distribution of interventions across time, space, and topic and variations in intervention characteristics. Where statistically significant differences from a random distribution were found, post hoc analyses of the standardized residuals were conducted to determine which individual categories significantly contributed to those differences (Sharpe 2015). Bonferroni corrections were used to adjust for multiple comparisons (Sharpe 2015).

Results

A total of 236 demand-reduction campaigns for wildlife products were identified (Supporting Information), all but 4 originating from the searches in English. A potential further 46 initiatives were identified but no information about them was available, due to either inaccessible literature sources, broken internet links, or lack of response from the implementing organization. The information on campaigns was not equally distributed across literature sources ($\chi^2 = 103.8$, df = 2, $n = 236$, $p < 0.001$); grey literature was significantly more represented than peer-reviewed documents (Supporting Information). The cam-
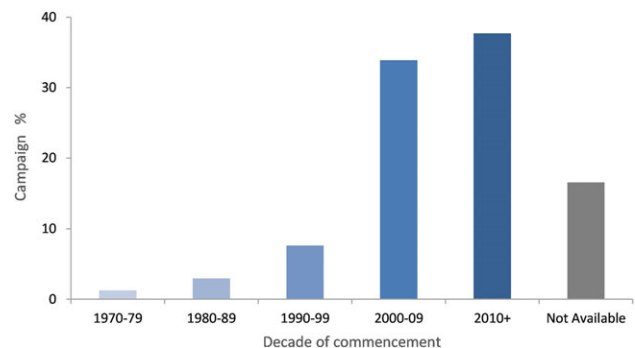

Figure 2. Demand-reduction campaigns for wildlife products by decade in which the campaign commenced.

paigns were also not led in equal proportions by different types of organizations ($\chi^2 = 467.1$, df = 3, $n = 233$, $p < 0.001$). Significantly more NGOs (85%) than independent organizations (e.g., universities), governments, and intergovernmental organizations (Supporting Information) led campaigns.

Campaigns were not distributed equally across time ($\chi^2 = 176.2$, df = 4, $n = 197$, $p < 0.001$). There were significantly more campaigns since the turn of the century (Supporting Information). This period accounted for 72% of all campaigns in this review (Fig. 2). Yet, for more than 15% of campaigns, we could not determine the year of implementation. Campaigns were not equally distributed in space ($\chi^2 = 153.5$, df = 5, $n = 185$, $p < 0.001$). Excluding international efforts, audiences in North America (20%) and particularly Asia (37%) were targeted by significantly more campaigns (Fig. 3) than audiences in Latin America, Europe, and Oceania (Supporting Information). On a national level, campaigns were most common in the United States (18%), China (9%), and Vietnam (6%).

Based on topical criteria, such as a product category, demand-reduction efforts were either species focused or had a wide scope that often targeted single or multiple genera (Fig. 4). Campaigns were not equally distributed across major taxa for species efforts ($\chi^2 = 123.43$, df = 4, $n = 140$, $p < 0.001$). Mammals were significantly more
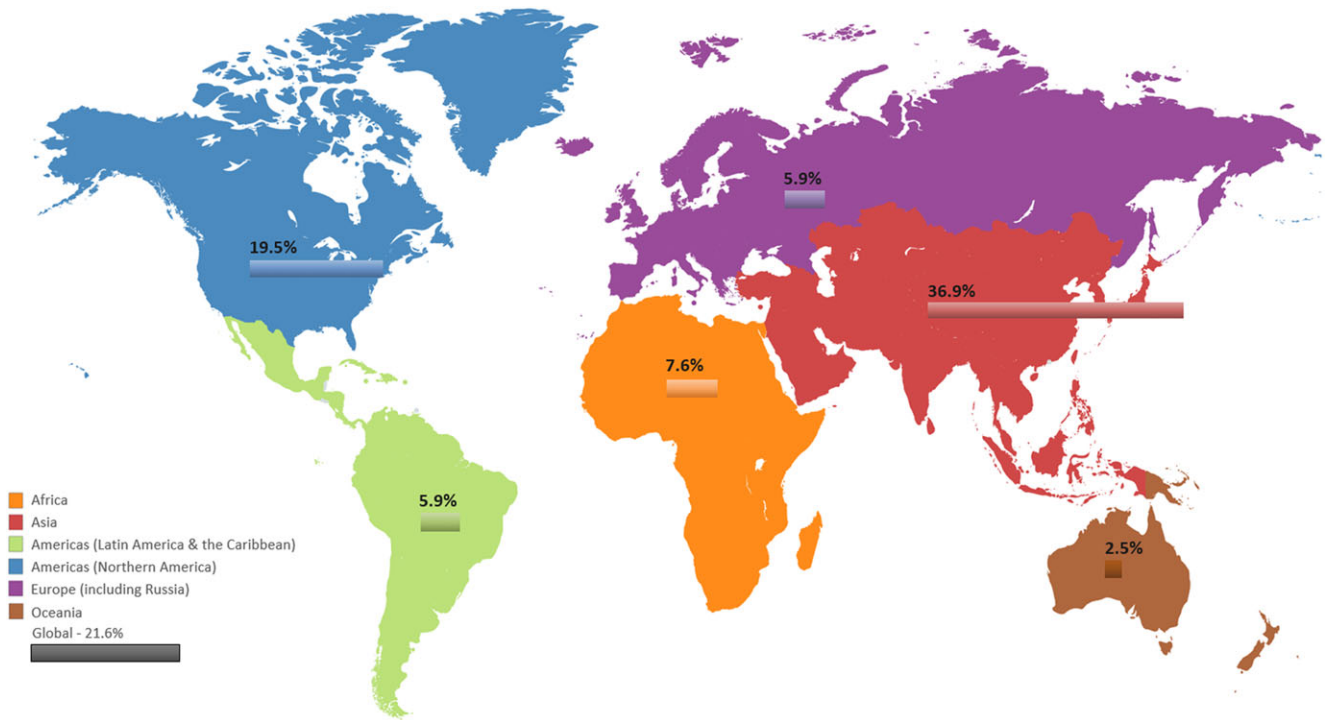
Figure 3. Spatial distribution of wildlife-demand reduction campaigns (based on United Nations Geoscheme 2018). Global refers to demand-reduction campaigns that targeted audiences across multiple continents.
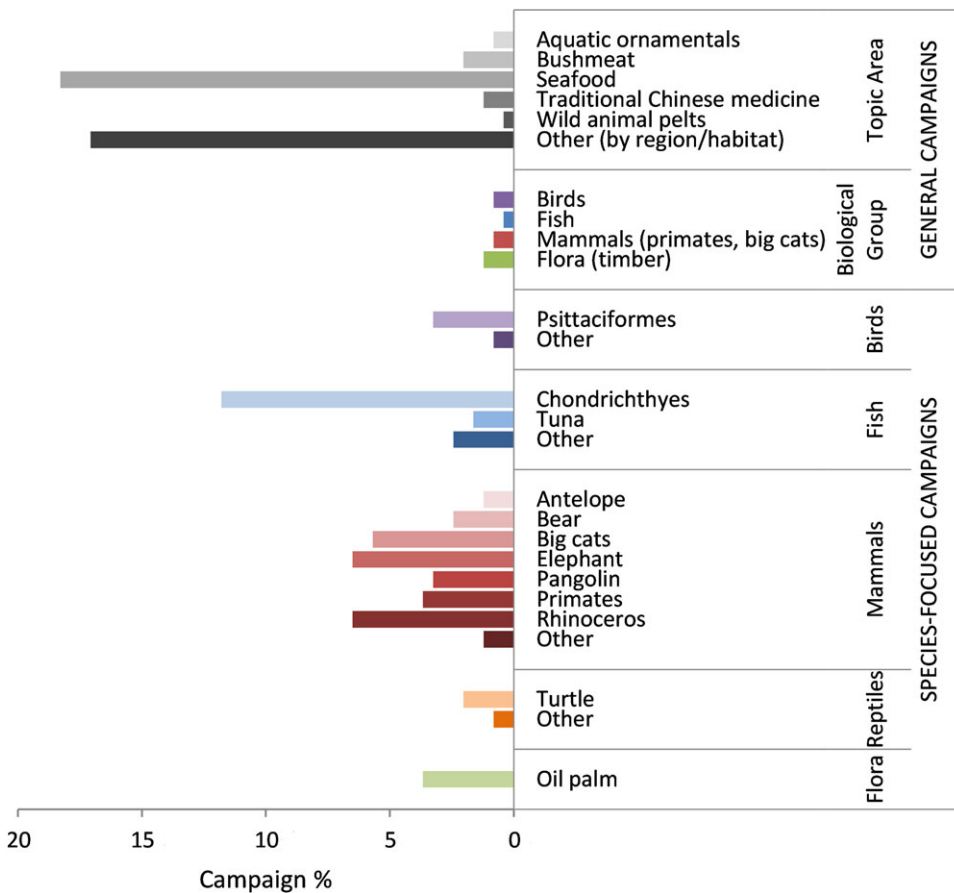

Figure 4. Percentage of demand-reduction campaigns focused on traded species or subject based on record frequency. Records that specified targeted efforts on multiple individual species are split and considered individual counts for species analysis (n = 246). Species groups with <3 campaign records are grouped and labeled "other" per biological group (see Supporting Information).
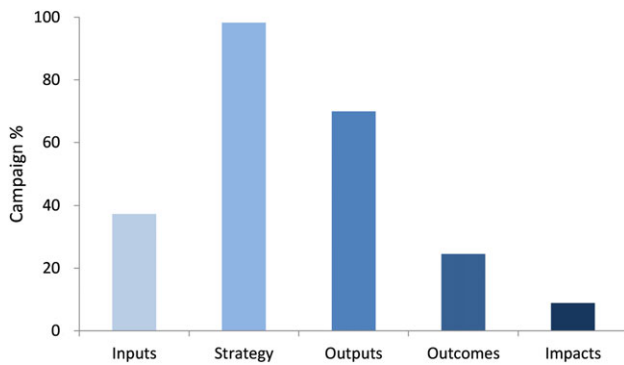
Figure 5. Types of project monitoring and evaluation information reported in campaigns to reduce demand for wildlife products (outcomes, changes in the target audience regarding, for example, knowledge, attitudes or behavior; impacts, biological changes or reduction to conservation threats).

represented than birds, reptiles, and plants (Supporting Information). In terms of specific animal groups, the greatest focus was on sharks (11%), elephants (7%), and rhinoceroses (7%) (Supporting Information). Notwithstanding, about 27% of campaigns focused broadly on products such as seafood, traditional Chinese medicine, and palm oil.

Overall, our results show that information was not available equally across the different stages of the evaluation logic model ($\chi^2 = 255.5$, df $= 4$, $n = 236$, $p < 0.001$) (Fig. 5). Data were significantly less available regarding campaign inputs (37%), outcomes (25%), and impacts (9%), whereas significantly more were on strategies (98%) and outputs (70%) (Supporting Information). This overall trend was consistent over time, however largely driven the large number of campaigns taking place in the last decades (Supporting Information). Campaigns applied various research designs for impact evaluation. Simple pre- and postcampaign comparisons were one of the most used designs; 26 (44%) campaigns reporting outcomes and 5 (24%) reporting impacts used it (Fig. 6). Time-series designs, with data collection occurring at multiple periods before and after interventions, were the most commonly reported impact data; 43% of interventions that reported impacts used it. Based on data from 44 records where some evaluation was carried out, campaigns mainly relied on questionnaire surveys (82%) to obtain systematic measures of outcomes; other sources were structured interviews (36%), direct observations (25%), and market data (5%).

In terms of outcome indicators, most campaigns used multiple measures, but showed no significant difference in how frequently knowledge, attitudes, and behavioral outcomes were reported ($\chi^2 = 2.14$, df $= 2$, $n = 236$, $p = 0.343$). The majority of behavioral indicators were self-reported (41%) and only 5 campaigns (5%) reported

observations of direct change of the behavior of interest (Supporting Information). Only anecdotal evidence was presented by 15 (25%) regarding outcomes, and 2 (10%) campaign regarding impact (Fig. 6). Two (3%) campaigns reported relevant information with comparable data and variability estimates that allowed for the calculation of effect sizes related to behavioral outcomes, but neither reported information related to impact.

Discussion

The heightened visibility of the wildlife trade as a threat to biodiversity and livelihoods has increased the effort placed on demand-reduction campaigns. Yet, there is little information on how this effort is being distributed across time, space, and taxa, as well as limited evidence around the evaluation of efforts to date, indicating major uncertainty about their impact. Filling these knowledge gaps will help ensure effort is allocated to the regions, topics, and species that most benefit from it and help derive learning insights that can support the design of future demand-reduction interventions.

Characterizing Demand-Reduction Campaigns

We applied a broad definition of "demand-reduction intervention" that focuses on the ultimate goal of influencing consumer action to reduce threats to biodiversity. We recognize there is great heterogeneity among the approaches used in the context of reducing demand for wildlife products and a further lack of clarity in the use of terminology from the social sciences and behavioral sciences. Examples include the inconsistent use of disciplinary labels, such as social marketing, and confounding disciplines that aim to influence human behavior, as in conservation education, with intervention planning tools like theory of change or campaign goals such as awareness raising (Greenfield & Veríssimo 2019; Wallen & Daut 2018). Improving the standardization of the use of social and behavioral science terminology should therefore be a key initial goal to improve the way conservationists communicate.

Despite our extensive search effort, we found it generally challenging to retrieve information about demand-reduction interventions. There were multiple instances where even basic information about campaigns was not available. One example was date of implementation, which was missing for more than 15% of the campaigns in our data set (Fig. 2). This is partially explained by the large number of campaigns that are only documented in the grey literature (Curzon & Kontoleon 2016), which is by nature less structured and thus harder to search. Nonetheless, this difficulty in retrieving basic information is also a result of lack of public access to reports and other internal project documents that are often seen as
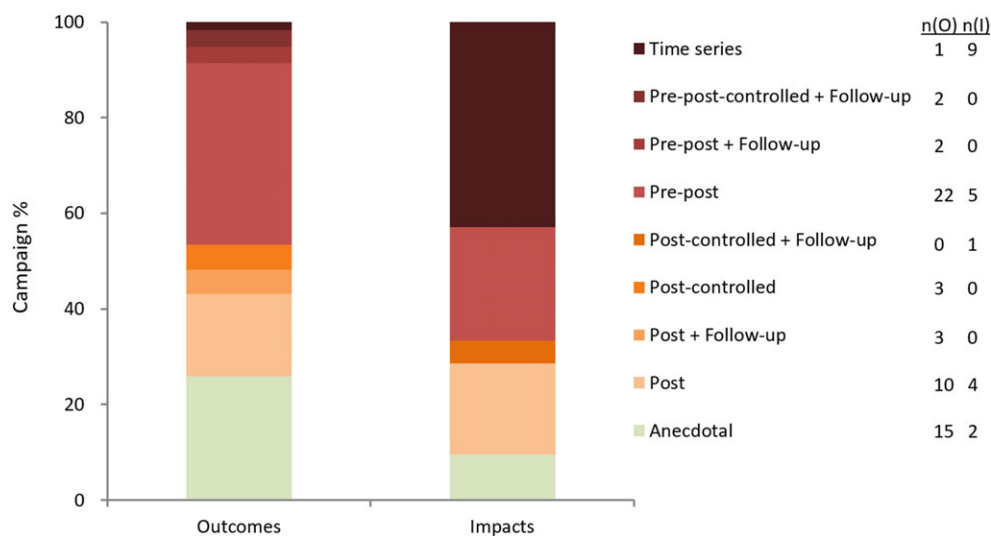
Figure 6. Research design of wildlife demand-reduction campaigns that reported outcomes (i.e., outcomes, changes in the target audience regarding, for example, knowledge, attitudes or behavior) and impacts (i.e., biological changes or reduction to conservation threats) (pre-, refers to data collected prior to the intervention; post-, refers to data collected after the intervention was completed; controlled, refers to the existence of a comparison unit that; follow-up, refers to data additional collection conducted after the intervention was completed, beyond that done immediately post intervention).

proprietary by the implementation institutions (Keene & Pullin 2011). The variety of sources searched and the emphasis on grey literature, allow us to be confident that our results, although we did not capture all campaigns, reveal broad patterns in the effort around demand-reduction campaigns for wildlife products.

Our data showed an upward temporal trend, which will be sustained at least until the end of the current decade. This suggests conservation practitioners are increasingly seeing demand-reduction campaigns as a way of mitigating the illegal wildlife trade, although this trend could also reflect, for example, an increase in online availability of documentation for more recent campaigns due to increased internet use. Asia was the continent with the most campaigns, a result that is not surprising given the dominant role of Asian countries as destination markets for high-profile wildlife products such as shark fin, elephant ivory, or rhinoceros horn (Fig. 3). However, at the national level, most demand efforts were implemented in the United States, reflecting perhaps where many NGOs implementing demand-reduction campaigns are based and thus where their donor-base is located.

Single species were the most popular topic, and mammals were the most frequent target of these efforts (Fig. 4). This was not a surprise because there is a wealth of evidence around human bias toward mammals (Martín-López et al. 2008). The large number of campaigns focusing on fish was unexpected, although the catalog on demand-reduction efforts for sharks compiled by Heller (2015) and the increased attention on shark-fin trade may have influenced this result. The very lim-

ited number of campaigns targeting plants should also be highlighted because they are a key biological group threatened by the wildlife trade (Hinsley et al. 2015).

Assessing Available Evidence

The evidence around campaign implementation was very limited; only about one-third reported any information on campaign inputs (Fig. 5), some of which was largely anecdotal. This is problematic because having a robust understanding of the resources invested in these interventions is key to understanding the trade-offs between intervention costs and benefits. Encouragingly, nearly all interventions reported some information about the strategies used, although the mostly anecdotal nature of this information made it impossible to understand, for example, what strategies are most common. We believe information on strategies was widely reported likely because it is one of the easiest to communicate to external audiences, such as donors and membership, while being cheap to collect and not institutionally sensitive. Regarding campaign outputs, these were reported by more than two-thirds of the campaigns, which is an encouraging sign when it comes to campaign implementation but does not allow for understanding campaign impacts.

The situation changed when it came to reporting outcomes. Only one-quarter of campaigns reported any evidence (Fig. 5), and one-quarter of the evidence was anecdotal (Fig. 6). Even when systematic data were collected, the majority of campaigns focused on indicators such as knowledge or attitudes, which are often poor proxies

of the behavioral changes demand-reduction campaigns aim to achieve (Kennedy et al. 2009). This situation is further exacerbated by the fact that those campaigns that collected data on behavior focused mostly on self-reported indicators, which are also poor predictors of actual behavior (Kormos & Gifford 2014). This means that it is currently difficult to determine how effective demand-reduction campaigns have been in achieving behavior change. The same could be said about the impacts of demand-reduction campaigns on biodiversity, with <9% reporting data on biological impact (Fig. 5). This information gap, also seen in other environmental fields, is likely a result of pragmatic factors, such as the high cost of data collection and the logistics needed to collect data on wildlife populations that can be continents away from relevant consumers, together with the technical complexities of establishing causality across large spatial scales (Rissman & Smail 2015; Veríssimo et al. 2018*a*). These limitations mean that it may not be possible or cost-effective for all campaigns to monitor biological impacts, although in that case it is key not to assume that behavioral changes will lead to impacts on biodiversity.

Our results also show that the campaigns often used research designs likely to have important limitations when it comes to internal validity (Wright & Lake 2015). For example, uncontrolled pre- and postcampaign comparisons did not account for time-varying factors, assuming no other event of relevance to the outcomes being considered occurred between the beginning and end of the campaign (Khandker et al. 2009). Yet, when campaigns lasted for several months or even years, this was most often an implausible assumption. These are important limitations that have led, for instance, many systematic review authors to only include evidence from quasi-experimental studies with an independent control group (Kongsted & Konnerup 2012).

In another frequently used experimental design those exposed to an intervention were compared with those who were not exposed after the intervention took place. However, these studies are vulnerable to selection bias, which can make treatment and control groups incomparable (Khandker et al. 2009). This is seen with mass media interventions such as PSAs, where allocations to the experimental or control group are most often conditional to respondents choosing to be exposed to the intervention (Veríssimo et al. 2018*b*). Considering the expectation that those with an interest in wildlife are more likely to listen to a PSA focused on wildlife conservation, in addition to being more likely to recall it, this would positively skew the comparison between treatment and control simply due to their initial composition.

The limitations highlighted above are not different from those faced in other areas of conservation science and occur even in fields such as international development, where paradigm shifts around impact evaluation are already taking place (Baylis et al. 2016). While the movement for evidence-based conservation has been gaining momentum, it is critical that conservation scientists and practitioners see behavior change initiatives as being as much in need of scientific rigor as any other part of conservation practice. There are already some encouraging signs of this; results of recent studies suggest a growing concern with impact evaluation among conservation practitioners (McKinnon et al. 2015; Curzon & Kontoleon 2016).

Improving the Evidence Base

Although the shortcomings highlighted above must be addressed, under penalty of the limited resources available for biodiversity conservation being used ineffectively, it must be recognized that impact evaluation in the context of biodiversity conservation can be very challenging. Barriers include, among others, large variability in natural outcome, long time lags between intervention and ecological response, programs with multiple interventions, complex spillover effects (e.g., due to species movement), and the large spatial scales of environmental processes (Ferraro 2009; Hockings et al. 2009; Rissman & Smail 2015). Specific to the wildlife trade, further complications arise from intricate consumptive influences with multifaceted drivers of demand, delayed responses for long-term behavior change, and an ever-adaptable industry that constantly challenges the effectiveness of management actions (Ayling 2016).

This complexity means that it is often difficult to implement best practices in terms of impact evaluation, leading to calls for a focus on developing and improving minimum standards instead (Curzon & Kontoleon 2016). These should focus on designs that, although less demanding in terms of time and resources, are still able to identify credible counterfactuals. One option is to use before after control impact (BACI) quasi-experimental designs in which the choice of control units is based on clearly selected variables, expert elicitation, or secondary data (Veríssimo et al. 2018*a*). Another option, where long-term data exists on an outcome of interest, would be the use of synthetic counterfactuals to systematically and transparently select control units by focusing on similarity in outcomes before the intervention (Sills et al. 2015). All impact evaluation in biodiversity conservation should also include a conceptual mapping tool (e.g., theory of change), that reveals the assumptions made in terms of causal links and identifies potential confounders for the outcomes being measured (Ferraro 2009).

It is also key that statistical testing be used to allow for exclusion of spurious changes and that statistical power analysis be used to ensure sample sizes are large enough to detect changes expected from an intervention. Another challenge is the measurement of change in human behavior, which is a crucial outcome for many conservation interventions (Veríssimo 2013).

Measurements of change in sensitive behaviors had seen some methodological advances in the last decade, through the emergence of techniques such as the unmatched count technique or the randomized response technique (Nuno & St John 2015). However, there is still an overreliance on self-reported behavioral indicators, which are unsystematically related to actual behavior, even with nonsensitive behaviors (Kormos & Gifford 2014). This limitation can be tackled by triangulating the results based on other independent data sources (e.g., market data) and measuring revealed preferences for products whose commercial trends could be expected to correlate with the products of interest (e.g., substitutes).

Delivering the Effectiveness Revolution

Although obstacles remain to ensuring that impact evaluation is seen as a specialist field of study, requiring specific knowledge and skills, there are also serious nontechnical barriers to improving evaluation practices across conservation science (Keene & Pullin 2011). Regardless of the valuable insights impact evaluation can generate, it is also true that it is a costly undertaking, both in terms of time and resources (Pullin et al. 2013). In the absence of a donor culture where evaluation is valued (Baylis et al. 2016), there may be instances where the limited resources available, or the existence of similar preexisting evaluations, dictate that the most reasonable option is to scale back the investment in evaluation (Mascia et al. 2014).

An additional problem that precludes impact evaluation from generating learning, and in that way improving conservation practice, is a lack of information sharing. The perception by organizations and individuals that the data describing the interventions they implement is proprietary means there are important barriers to collaboration (Keene & Pullin 2011). The situation is worsened by a climate of competition between institutions and individual researchers for funding and airtime, which makes intervention outcomes strategic communication elements first and opportunities for learning second (Redford & Taber 2000).

Donors are likely to have a pivotal role in determining how evaluation is conducted because they not only have the power to dictate short-term change in the data-sharing requirements of their grantees, but they also have a strategic interest in rigorously evaluating their own impact (Keene & Pullin 2011). By promoting a culture of open information sharing, donors will have the opportunity to refocus the narrative around their projects from a false success or failure dichotomy to a culture of learning and adaptive management. Nevertheless, conservationists should also assume part of the responsibility for improving the accountability and oversight in their field. Efforts such as the Wildlife Consumer Behaviour Change Community of Practice

(http://www.changewildlifeconsumers.org/) will be key to furthering the exchange of experiences and learning insights among conservationists.

The current emphasis on demand reduction provides a unique opportunity to make these behavioral approaches a serious way of addressing the unsustainable and often illegal exploitation of wildlife. However, this opportunity will only come to fruition if there is an investment in the rigor, transparency, and accountability brought by open and systematic impact evaluation. While calls for conservationists to embrace the effectiveness revolution have become common throughout the last decade, limited progress has been achieved, particularly in comparison with fields such as public health or international development (Sutherland et al. 2004; Ferraro & Pattanayak 2006; Keene & Pullin 2011; Baylis et al. 2016). This scenario harms the ability of practitioners to clearly distinguish between what works and what does not, hampering their ability to generate learning insights to improve future demand-reduction efforts. Given how difficult it is to influence human behavior, particularly in a complex context, such as wildlife trade, conservationists can only succeed if they can benefit from the incremental learning that comes with a more transparent and rigorous evaluation of demand-reduction interventions.

Acknowledgments

Supporting Information

Data on the variation of reporting on campaign monitoring and evaluation over time (Appendix S1), a description of the searches conducted (Appendix S2), details of all interventions examined (Appendix S3), results of chi-square residual analysis (Appendix S4), and a description of the topics on which the campaign review focused (Appendix S5) are available online. The authors are solely responsible for the content and functionality of these materials. Queries (other than absence of the material) should be directed to the corresponding author.

Literature Cited

Ayling J. 2016. Transnational environmental crime: meeting future challenges through networked regulatory innovations. Pages 73–90 in Hall M, Wyatt T, South N, Nurse A, Potter G, Maher J, editors. Greening criminology in the 21st century: contemporary debates and future directions in the study of environmental harm. Routledge, Oxford, United Kingdom.

Baylis K, Honey-Rosés J, Börner J, Corbera E, Ezzine-de-Blas D, Ferraro PJ, Lapeyre R, Persson UM, Pfaff A, Wunder S. 2016.

Mainstreaming impact evaluation in nature conservation. Conservation Letters **9**:58–64.

Blundell AG, Mascia MB. 2005. Discrepancies in reported levels of international wildlife trade. Conservation Biology **19**:2020–2025.

Burgess G. 2016. Powers of persuasion? Traffic Bulletin **28**:65.

Challender DW, MacMillan DC. 2014. Poaching is more than an enforcement problem. Conservation Letters **7**:484–494.

Curzon HF, Kontoleon A. 2016. From ignorance to evidence? The use of programme evaluation in conservation: evidence from a Delphi survey of conservation experts. Journal of Environmental Management **180**:466–475.

Ferraro PJ. 2009. Counterfactual thinking and impact evaluation in environmental policy. New Directions for Evaluation **2009**:75–84.

Ferraro PJ, Pattanayak SK. 2006. Money for nothing? A call for empirical evaluation of biodiversity conservation investments. PLoS Biology **4**:e105.

Greenfield SJ, Veríssimo D. 2019. To what extent is social marketing used in demand reduction campaigns for illegal wildlife products? Insights from elephant ivory and rhino horn. Social Marketing Quarterly https://doi.org/10.1177/1524500418813543.

Haenlein C, Maguire T, Somerville K. 2016. III. Poaching, wildlife trafficking and terrorism. Whitehall Papers **86**:58–76.

Heller M. 2015. An exploration of effective messaging for shark conservation. M.A.S Capstone Report. Scripps Institution of Oceanography, University of California San Diego. Available from https://escholarship.org/uc/item/3k230030 (accessed February 10, 2018).

Hinsley A, Verissimo D, Roberts DL. 2015. Heterogeneity in consumer preferences for orchids in international trade and the potential for the use of market research methods to study demand for wildlife. Biological Conservation **190**:80–86.

Hockings M, Stolton S, Dudley N, James R. 2009. Data credibility: What are the "right" data for evaluating management effectiveness of protected areas? New Directions for Evaluation **2009**:53–63.

Karesh WB, Cook RA, Bennett EL, Newcomb J. 2005. Wildlife trade and global disease emergence. Emerging Infectious Diseases **11**:1000–1002.

Keene M, Pullin AS. 2011. Realizing an effectiveness revolution in environmental management. Journal of Environmental Management **92**:2130–2135.

Kennedy EH, Beckley TM, McFarlane BL, Nadeau S. 2009. Why we don't "walk the talk": understanding the environmental values/behaviour gap in Canada. Human Ecology Review **16**:151–160.

Khandker S, Koolwal B, Samad H. 2009. Handbook on impact evaluation. World Bank, Washington, D.C.

Kongsted HC, Konnerup M. 2012. Are more observational studies being included in Cochrane Reviews? BMC Research Notes **5**:570–570.

Kormos C, Gifford R. 2014. The validity of self-report measures of proenvironmental behavior: a meta-analytic review. Journal of Environmental Psychology **40**:359–371.

Margoluis R, Stem C, Salafsky N, Brown M. 2009. Using conceptual models as a planning and evaluation tool in conservation. Evaluation and Program Planning **32**:138–147.

Martín-López B, Montes C, Benayas J. 2008. Economic valuation of biodiversity conservation: the meaning of numbers. Conservation Biology **22**:624–635.

Mascia MB, Pailler S, Thieme ML, Rowe A, Bottrill MC, Danielsen F, Geldmann J, Naidoo R, Pullin AS, Burgess ND. 2014. Commonalities and complementarities among approaches to conservation monitoring and evaluation. Biological Conservation **169**:258–267.

McKinnon MC, Mascia MB, Yang W, Turner WR, Bonham C. 2015. Impact evaluation to communicate and improve conservation nongovernmental organization performance: the case of Conservation International. Philosophical Transactions of the Royal Society B: Biological Sciences **370**:20140282.

Megias DA, Anderson SC, Smith RJ, Veríssimo D. 2017. Investigating the impact of media on demand for wildlife: a case study of Harry Potter and the UK trade in owls. PloS ONE **12** (e0182368) https://doi.org/10.1371/journal.pone.0182368.

Miteva DA, Pattanayak SK, Ferraro PJ. 2012. Evaluation of biodiversity policy instruments: What works and what doesn't? Oxford Review of Economic Policy **28**:69–92.

Nuno A, St John FAV. 2015. How to ask sensitive questions in conservation: A review of specialized questioning techniques. Biological Conservation **189**:5–15.

Olmedo A, Sharif V, Milner-Gulland E. 2018. Evaluating the design of behaviour change interventions: A case study of rhino horn in Vietnam. Conservation Letters **11**:e12365.

Pullin AS, Sutherland W, Gardner T, Kapos V, Fa JE. 2013. Conservation priorities: identifying need, taking action and evaluating success. Key Topics in Conservation Biology **2**:3–22.

Redford KH, Taber A. 2000. Writing the wrongs: developing a safe-fail culture in conservation. Conservation Biology **14**:1567–1568.

Rissman AR, Smail R. 2015. Accounting for results: how conservation organizations report performance information. Environmental Management **55**:916–929.

Roberton SI. 2014. Has demand for rhino horn truly dropped in Vietnam? Blog. National Geographic, Washington, D.C. Available from http://voices.nationalgeographic.com/2014/11/03/has-demand-for-rhino-horn-truly-dropped-in-vietnam/ (accessed September 20, 2017).

Roe D, Booker F, Day M, Zhou W, Allebone-Webb S, Hill NA, Kumpel N, Petrokofsky G, Redford K, Russell D. 2015. Are alternative livelihood projects effective at reducing local threats to specified elements of biodiversity and/or improving or maintaining the conservation status of those elements? Environmental Evidence **4**:22.

Rosen GE, Smith KF. 2010. Summarizing the evidence on the international trade in illegal wildlife. EcoHealth **7**:24–32.

Rosenbaum PR. 2010. Design of observational studies. Springer, New York.

Shairp R, Veríssimo D, Fraser I, Challender D, MacMillan D. 2016. Understanding urban demand for wild meat in Vietnam: implications for conservation actions. PLOS ONE **11** (e0134787) https://doi.org/10.1371/journal.pone.0134787.

Sharif V. 2014. Analysis of conservation initiatives aimed at reducing demand for traded wildlife in China and Vietnam. Stop Ivory, London. Available from http://stopivory.org/resources/ (accessed September 20, 2017).

Sharpe D. 2015. Your chi-square test is statistically significant: Now what? Practical Assessment, Research & Evaluation **20**:1–20.

Sills EO, Herrera D, Kirkpatrick AJ, Brandão Jr A, Dickson R, Hall S, Pattanayak S, Shoch D, Vedoveto M, Young L. 2015. Estimating the impacts of local policy innovation: the synthetic control method applied to tropical deforestation. PloS ONE **10** (e0132590) https://doi.org/10.1371/journal.pone.0132590.

Stem C, Margoluis R, Salafsky N, Brown M. 2005. Monitoring and evaluation in conservation: a review of trends and approaches. Conservation Biology **19**:295–309.

Sutherland WJ, Pullin AS, Dolman PM, Knight TM. 2004. The need for evidence-based conservation. Trends in Ecology & Evolution **19**:305–308.

Veríssimo D. 2013. Influencing human behaviour: an underutilised tool for biodiversity management. Conservation Evidence **10**:29–31.

Veríssimo D, Bianchessi A, Arrivillaga A, Cadiz FC, Mancao R, Green K. 2018*a*. Does it work for biodiversity? Experiences and challenges in the evaluation of social marketing campaigns. Social Marketing Quarterly **24**:18–34.

Veríssimo D, Challender D, Nijman V. 2012. Wildlife trade in Asia: start with the consumer. Asian Journal of Conservation Biology **1**:49–50.

Veríssimo D, Schmid C, Kimario FF, Eves HE. 2018*b*. Measuring the impact of an entertainment-education intervention to reduce demand for bushmeat. Animal Conservation **21:**324–331.

Wallen KE, Daut E. 2018. The challenge and opportunity of behaviour change methods and frameworks to reduce demand for illegal wildlife. Nature Conservation **26:**55.

Warchol GL. 2004. The transnational illegal wildlife trade. Criminal Justice Studies **17:**57–73.

Wright LL, Lake DA. 2015. Basics of research for the health professions. Internal validity. Available from http://www.pt.armstrong.edu/wright/hlpr/text/3.7.intval.htm (accessed September 20, 2017).