Characterizing efforts to reduce consumer demand for wildlife products

Diogo Veríssimo[1,2,3*], Anita K. Y. Wan [4*]

[1] Department of Zoology, University of Oxford, Oxford, United Kingdom

[2] San Diego Zoo Institute for Conservation Research, Escondido, CA, USA

[3] Oxford Martin School, Oxford Martin Programme on the Illegal Wildlife Trade, University of Oxford, Oxford, United Kingdom

[4] Independent researcher, Tai Po, New Territories, Hong Kong, China.

Corresponding author

Diogo Veríssimo

Department of Zoology, University of Oxford,

Zoology Research and Administration Building

11a Mansfield Road

Oxford OX1 3SZ

Email: verissimodiogo@gmail.com

Running head: Demand reduction

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/cobi.13227.

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

Article Impact Statement: Effort to reduce demand for wildlife products has increased, but there is high uncertainty about the impact of past campaigns.

Abstract

The unsustainable trade in wildlife is a key threat to the Earth's biodiversity. Whilst efforts to mitigate this threat have traditionally focused on regulation and enforcement, there is a growing interest in campaigns aimed at reducing consumer demand for wildlife products. In this article, we aim to characterize these behaviour change campaigns and the existing evidence around their impacts. We found a total of 236 campaigns, mainly in the grey literature. The number of campaigns grew over the last decades, although for more than 15% of the campaigns a start date was not available. The Asian continent was the primary focus, although at the national level the USA was where most campaigns took place. As expected, campaigns most often focused on single species of mammal, with other vertebrates groups, with the exception of sharks, receiving limited attention. Many campaigns also focused on broad themes such as the wildlife trade in general or seafood. Regarding evaluation, we found that 37% of campaigns reported some information on their inputs, 98% on strategies, 70% on outcomes, 37% on outcomes and 9% on impacts. Information on outcomes and impacts was largely anecdotal or based on research designs that are at a high risk of bias, such as pre-post comparisons. At present, it is challenging to know whether demand reduction campaigns are having a direct behavioural or biological impact. The lack of robust impact evaluation also makes it difficult to draw learning 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 practices as well as a more collaborative approach that fosters the sharing of both data and learning insights.

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 rhino horn and elephant ivory, involving also large numbers of invertebrate animals, plants and fungi (Blundell & Mascia 2005). Yet, beyond its substantial and often irreversible impact on biodiversity, the illegal wildlife trade can have profound impacts on the health, economic development and governance of societies in both source and consumer countries (Haenlein et al. 2016; Karesh et al. 2005; Rosen & Smith 2010).

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

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 labelled in a variety of ways, from awareness-raising and environmental education to human-centred design and social marketing (Olmedo et al. 2017; Wallen & Daut 2018). Yet, our understanding of what these different outreach efforts have achieved remains limited. This stems from unclear goals, mismatches between campaign activities and the behaviours that to be influenced and lack of rigorous project monitoring and evaluation (Greenfield 2015; Olmedo et al. 2017; Veríssimo et al. 2018b). As a result, claims of success by implementers are often received with scepticism (Roberton 2014). These challenges are not unique to efforts around the wildlife trade, being felt 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 (Curzon & Kontoleon 2016; Ferraro & Pattanayak 2006), often along a logic model (Figure 1) that linked project inputs and activities to outcomes and impact (Rissman & Smail 2015). Under this type of approach, it is often assumed that if the indicator improves, an intervention is being effective, while if the indicator worsens, the opposite is true (Ferraro 2009).

While 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 2009; Ferraro & Pattanayak 2006). To evaluate impact we need 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 an unobservable counterfactual scenario, (i.e., what would have happened without said intervention), which cannot be observed, and can only be inferred 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 against biases and confounders (Curzon & Kontoleon 2016; Ferraro 2009).

In the case of behaviour change efforts aimed at reducing demand for illegal or unsustainable wildlife products, we lack even a basic understanding of what efforts have taken place globally, what has been achieved thus far and what lessons have been learned. In this article, we aim to (1) characterise the temporal, spatial distribution and biological focus of demand reduction campaigns; (2) synthesize the evidence of project monitoring and evaluation of those campaigns; and (3) generate learning insights to inform future conservation efforts.

Methods

We define demand reduction campaigns as outreach interventions aimed to voluntarily change actual behaviour of current or potential consumers of wildlife products or their derivatives (Burgess 2016; Olmedo et al. 2017; Veríssimo et al. 2012). We therefore only considered consumer-focused efforts, and not those with, for example, an enforcement or policy focus. We defined wildlife as including all animals and plants. In cases where multiple interventions overlapped geographically and shared implementing or funding agencies, those of 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. These searches used general keywords associated with wildlife trade, as well as species-specific terms involving 43 CITES listed species often associated with the wildlife trade (see Supplementary Material - Table S1). We also included the Bluefin Tuna (*Thunnus thynnus*), a high profile species that has repeatedly faced stiff resistance to CITES Appendix listing, despite being highly threatened by commercial exploitation. All searches were conducted using both 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. Due to 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 (see Supplementary Material - Table S2). Lastly, we visited websites of NGOs, professional groups and funding organizations with a track record of funding or implementing demand reduction campaigns. We obtained an initial list through the catalogue 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, with the ultimate goal of locating information on additional demand reduction initiatives (See Supplementary Material - Table S3).

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, outputs, outcomes and impact (Figure 1). We classified the inputs as quantified information, whether 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 behaviours. Lastly, 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 better understand trends regarding the distribution of interventions across time, space and topic, as well as 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 (see Supplementary Material – Table S4), all but four 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, and/or lack of response from the implementing organization. The information on campaigns was not equally distributed across the various literature sources (χ^2 (2, 236) = 103.8, $p < 0.001$), with grey literature significantly more represented than peer-reviewed documents (see Supplementary Material – Table S5). The campaigns were also not led in equal proportions by different organisation types (χ^2 (3, 233) = 467.1, $p < 0.001$). These were significantly more often led by non-governmental organizations (NGOs, 85%), and significantly less from independent organizations (e.g. universities), governments and intergovernmental organizations (see Supplementary Material – Table S6).

Campaigns were not distributed equally across time (χ^2 (4, 197) = 176.2, $p < 0.001$). There were significantly more campaigns since the turn of the century (see Supplementary Material – Table S7), with this period accounting for 72% of all campaigns in this review (Figure 2). Yet, for more than 15% of campaigns we could not determine the year of implementation. We also found that campaigns were not equally distributed in space (χ^2 (5, 185) = 153.5, $p < 0.001$). Excluding international efforts, audiences in North America (20%) and particularly Asia (37%) were targeted by significantly more campaigns (Figure 3), while Latin America, Europe and Oceania had significantly fewer campaigns (see Supplementary Material – Table S8). On a national level, campaigns were most common in the USA (18%), China (9%) and Vietnam (6%).

Demand reduction efforts were found either species-focused, or with a wider scope that often targeted single or multiple genera, based on topical criteria such as a product category (Figure 4). Campaigns were not equally distributed across major taxa for species efforts (χ^2 (4, 140) = 123.43, $p < 0.001$). Mammals were significantly more represented while birds, reptiles and plants were significantly less represented (see Supplementary Material – Table S9). In terms of specific animal groups, most attention focused on sharks (11%), elephants (7%) and rhinos (7%) (see Supplementary Material – Table S10). 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 (χ^2 (4, 236) = 255.5, $p < 0.001$; Figure 5). Data was significantly less available regarding campaign inputs (37%), outcomes (25%) and impacts (9%), while significantly more was on strategies (98%) and outputs (70%; see Supplementary Material – Table S11). This overall trend is consistent over time, however largely driven the large number of campaigns taking place in the last decades (see Supplementary Material – Figure S1). Campaigns applied various research designs for impact evaluation. Simple Pre-post comparisons were one of the most used designs, with 26 (44%) campaigns reporting outcomes and five (24%) reporting impacts using it (Figure 6). Time-series designs, with data collection occurring at multiple periods before and after interventions, were the most commonly reported impact data, with 43% of interventions that reported impacts using 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, with other sources from 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 behavioural outcomes were reported (χ^2 (2, 236) = 2.14, p = 0.343). The majority of behavioural indicators were self-reported (41%) and only five campaigns (5%) reported observations of direct change of the behaviour of interest (see Supplementary Material – Table S12). Only anecdotal evidence was presented by 15 (25%) regarding outcomes, and two (10%) campaign regarding impact (Figure 6). Furthermore, only two (3%) campaigns reported relevant information with comparable data and variability estimates that allowed for the calculation of effect sizes related to behavioural outcomes, with none related to impact.

Discussion

The heightened visibility of the wildlife trade as a threat to both biodiversity and livelihoods has increased the effort placed on demand reduction campaigns. Yet, we have little knowledge of 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 gaps in our knowledge will be key to 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.

Characterising demand reduction campaigns

In this study we use a broad definition of demand reduction intervention that focuses on the ultimate goal of influencing consumer action to reduce threats to biodiversity. We recognise that there is great heterogeneity amongst 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 behavioural sciences. Examples include the inconsistent use of disciplinary labels, such as social marketing, and confounding disciplines that aim to influence human behaviour, as in conservation education, with intervention planning tools like theory of change or campaign goals such as awareness-raising (Greenfield 2015; Wallen & Daut 2018). Improving the standardisation of the use of social and behavioural 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 dataset (Figure 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 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, while certainly not capturing all campaigns, allow us to understand 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 rhino horn (Figure 3). However, at the national level, most demand efforts were implemented in the USA, reflecting perhaps where many NGOs implementing demand reduction campaigns are based, and thus where their donor-base is located.

Single species were unsurprisingly the most popular topic, with mammals being the most frequent target of these efforts (Figure 4). This is not a surprise, as there is a wealth of evidence around the societal bias towards mammals (Martín-López et al. 2008). The high number of campaigns focusing on fish was unexpected, although it should be noted that the catalogue on demand reduction efforts for sharks compiled by Heller (2015), together with the increased attention on shark fin trade may have influenced this result. The very limited number of campaigns targeting plants should also be highlighted, as they are a key biological group threatened by the wildlife trade (Hinsley et al. 2015).

Assessing the available evidence

The evidence around campaign implementation was very limited, with only about a third reporting any information on campaign inputs (Figure 5), some of which was largely anecdotal. This is problematic as 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 changes when it comes to reporting outcomes, with only a quarter of campaigns reporting any evidence (Figure 5), and a quarter of this being anecdotal (Figure 6). It is also important to mention that even when systematic data was collected, the majority of campaigns focused on indicators such as knowledge or attitudes, which are often poor proxies of the behavioural 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 behaviour focused mostly on self-reported indicators, which have been shown to also be poor predictors of actual behaviour (Kormos & Gifford 2014). This means that it is currently difficult to understand how effective demand reduction campaigns have been in achieving behaviour change. The same could be said about the impacts of demand reduction campaigns on biodiversity, with less than 9% reporting data on biological impact (Figure 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. 2018a). 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 any behavioural changes will lead to impacts on biodiversity.

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

Another frequently used experimental design used was a controlled-post, where those exposed to an intervention were compared with those unexposed, 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. 2018b). 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 to can also be seen even in fields such as international development, where paradigm shifts around impact evaluation are already taking place (Baylis et al. 2015). While the movement for evidence-based conservation has been gaining momentum, it is critical that conservation scientists and practitioners see behaviour change initiatives as being as much in need of scientific rigour as any other part of conservation practice. There are already some encouraging signs of this, with recent studies suggesting a growing concern with impact evaluation amongst conservation practitioners (Curzon & Kontoleon 2016; McKinnon et al. 2015).

Improving the evidence base

While 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, amongst others, high natural outcome variability, 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). In specific context to the wildlife trade, further complications arise from intricate consumptive influences with multifaceted drivers of demand, delayed responses for long-term behaviour change, and an ever-adaptable industry environment 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, while less demanding in terms of time and resources, are still able to identify credible counterfactuals. One option would be the use of BACI quasi-experimental designs where the choice of control units is based on clearly selected variables, chosen by expert elicitation and/or secondary data (Veríssimo et al. 2018a). 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 is used to allow for spurious changes to be excluded and, at the same time, that statistical power analysis is used to ensure sample sizes are large enough to detect changes expected from an intervention. Another challenge is the measurement of change in human behaviour, which is a crucial outcome for many conservation interventions (Veríssimo 2013). Measurements of change in sensitive behaviours 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 behavioural indicators, which are unsystematically related to actual behaviour, even with non-sensitive behaviours (Kormos & Gifford 2014). This limitation can be tackled by triangulating the results using other independent data sources (e.g., market data), as well as 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

While obstacles remain to ensuring that impact evaluation is seen as a specialist field of study, requiring specific knowledge and skills, there are also serious non-technical 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. 2015), there may be instances where the limited resources available, or the existence of similar pre-existing evaluations, dictate that the most reasonable option is to scale back the investment in evaluation (Mascia et al. 2014).

An additional limitation that precludes impact evaluation from generating learning insights, and in that way improving how conservation is practiced, 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, as they not only have the power to dictate short-term change in the data sharing requirements of their grantees, but 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 established by TRAFFIC (http://www.changewildlifeconsumers.org/) will be key to further the exchange of experiences and learning insights among conservationists.

The current emphasis placed on demand reduction provides a unique opportunity to make these behavioural 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 (Baylis et al. 2015; Ferraro & Pattanayak 2006; Keene & Pullin 2011; Sutherland et al. 2004). 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 behaviour, 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

DV acknowledges support of the David H. Smith Conservation Research Fellowship. DV is an Oxford Martin Fellow at the Oxford Martin School Programme for the Illegal Wildlife Trade.

Supporting Information

Data of 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 in the this review (Appendix S3), the results of the chi-square residual analysis (Appendix S4) and a description of the topics on which the campaigns reviewed 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 M. Hall, T. Wyatt, N. South, A. Nurse, G. Potter, and J. Maher, editors. Greening Criminology in the 21st Century: contemporary debates and future directions in the study of environmental harm. Routledge, Oxford, UK ; New York, USA.

Baylis, K., J. Honey-Rosés, J. Börner, E. Corbera, D. Ezzine-de-Blas, P. J. Ferraro, R. Lapeyre, U. M. Persson, A. Pfaff, and S. Wunder. 2015. Mainstreaming impact evaluation in nature conservation. Conservation Letters.

Blundell, A. G., and M. B. Mascia. 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, D. W., and D. C. MacMillan. 2014. Poaching is more than an enforcement problem. Conservation Letters **7**:484-494.

Curzon, H. F., and A. Kontoleon. 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, P. J. 2009. Counterfactual thinking and impact evaluation in environmental policy. New Directions for Evaluation **2009**:75-84.

Ferraro, P. J., and S. K. Pattanayak. 2006. Money for nothing? A call for empirical evaluation of biodiversity conservation investments. PLoS biology **4**:e105.

Greenfield, S. J. 2015. What is the role of social marketing in consumer demand reduction of illegal wildlife products? An investigation of ivory and rhino horn campaigns in China and Viet Nam. Masters Thesis, Cambridge University, UK.

Haenlein, C., T. Maguire, and K. Somerville. 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, USA. Available online at https://escholarship.org/uc/item/3k230030.

Hinsley, A., D. Verissimo, and D. L. Roberts. 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., S. Stolton, N. Dudley, and R. James. 2009. Data credibility: what are the “right” data for evaluating management effectiveness of protected areas? New Directions for Evaluation **2009**:53-63.

Karesh, W. B., R. A. Cook, E. L. Bennett, and J. Newcomb. 2005. Wildlife trade and global disease emergence. Emerg Infect Dis **11**:1000-1002.

Keene, M., and A. S. Pullin. 2011. Realizing an effectiveness revolution in environmental management. Journal of Environmental Management **92**:2130-2135.

Kennedy, E. H., T. M. Beckley, B. L. McFarlane, and S. Nadeau. 2009. Why we don't" walk the talk": Understanding the environmental values/behaviour gap in Canada. Human Ecology Review:151-160.

Khandker, S., B. Koolwal, and H. Samad. 2009. Handbook on impact evaluation. World Bank, Washington DC. USA.

Kongsted, H. C., and M. Konnerup. 2012. Are more observational studies being included in Cochrane Reviews? BMC Research Notes **5**:570-570.

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

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

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

Mascia, M. B., S. Pailler, M. L. Thieme, A. Rowe, M. C. Bottrill, F. Danielsen, J. Geldmann, R. Naidoo, A. S. Pullin, and N. D. Burgess. 2014. Commonalities and complementarities among approaches to conservation monitoring and evaluation. Biological Conservation **169**:258-267.

McKinnon, M. C., M. B. Mascia, W. Yang, W. R. Turner, and C. Bonham. 2015. Impact evaluation to communicate and improve conservation non-governmental organization performance: the case of Conservation International. Phil. Trans. R. Soc. B **370**:20140282.

Megias, D. A., S. C. Anderson, R. J. Smith, and D. Veríssimo. 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.

Miteva, D. A., S. K. Pattanayak, and P. J. Ferraro. 2012. Evaluation of biodiversity policy instruments: what works and what doesn't? Oxford Review of Economic Policy **28**:69-92.

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

Olmedo, A., V. Sharif, and E. Milner-Gulland. 2017. Evaluating the design of behaviour change interventions: A case study of rhino horn in Vietnam. Conservation Letters.

Pullin, A. S., W. Sutherland, T. Gardner, V. Kapos, and J. E. Fa. 2013. Conservation priorities: identifying need, taking action and evaluating success. Key topics in conservation biology **2**:3-22.

Redford, K. H., and A. Taber. 2000. Writing the Wrongs: Developing a Safe-Fail Culture in Conservation. Conservation Biology **14**:1567-1568.

Rissman, A. R., and R. Smail. 2015. Accounting for results: how conservation organizations report performance information. Environmental management **55**:916-929.

Roberton, S. I. 2014. Has Demand for Rhino Horn Truly Dropped in Vietnam? National Geographic Blog. Avaialbe at http://voices.nationalgeographic.com/2014/11/03/has-demand-for-rhino-horn-truly-dropped-in-vietnam/.

Roe, D., F. Booker, M. Day, W. Zhou, S. Allebone-Webb, N. A. Hill, N. Kumpel, G. Petrokofsky, K. Redford, and D. Russell. 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, G. E., and K. F. Smith. 2010. Summarizing the evidence on the international trade in illegal wildlife. EcoHealth **7**:24-32.

Rosenbaum, P. R. 2010. Design of observational studies. Springer, New York, USA.

Shairp, R., D. Veríssimo, I. Fraser, D. Challender, and D. MacMillan. 2016. Understanding Urban Demand for Wild Meat in Vietnam: Implications for Conservation Actions. PloS one **11**.

Sharif, V. 2014. Analysis of conservation initiatives aimed at reducing demand for traded wildlife in China and Vietnam. Stop Ivory. London, UK. Available at: http://stopivory.org/resources/.

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

Sills, E. O., D. Herrera, A. J. Kirkpatrick, A. Brandão Jr, R. Dickson, S. Hall, S. Pattanayak, D. Shoch, M. Vedoveto, and L. Young. 2015. Estimating the impacts of local policy innovation: the synthetic control method applied to tropical deforestation. PloS one **10**:e0132590.

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

Sutherland, W. J., A. S. Pullin, P. M. Dolman, and T. M. Knight. 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., A. Bianchessi, A. Arrivillaga, F. C. Cadiz, R. Mancao, and K. Green. 2018a. Does it work for biodiversity? Experiences and challenges in the evaluation of social marketing campaigns. Social Marketing Quarterly **24**:18-34.

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

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

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

Warchol, G. L. 2004. The transnational illegal wildlife trade. Criminal justice studies **17**:57-73.

Wright, L. L., and D. A. Lake. 2015. Basics of Research for the Health Professions. Chapter 3.7 Internal Validity. Available at http://www.pt.armstrong.edu/wright/hlpr/text/3.7.intval.htm.

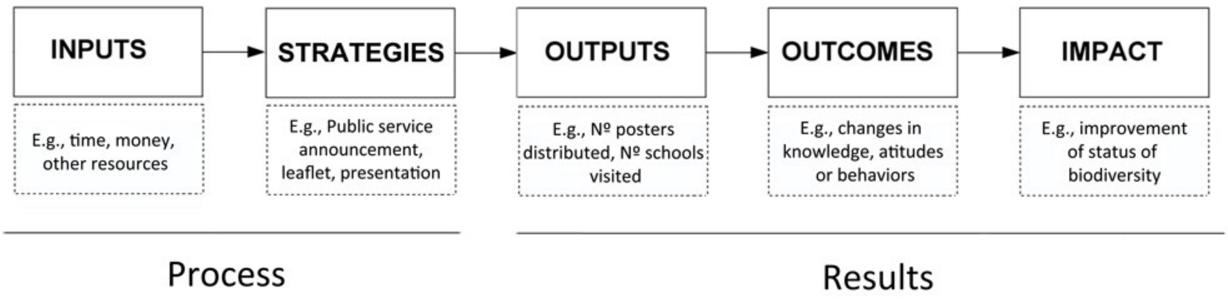

Figure 1 – General logic model followed to classify the types of monitoring and evlauation information reported for different outreach campaigns to reduce the demand for wildlife products. Adapted from Margoluis et al. (2009).

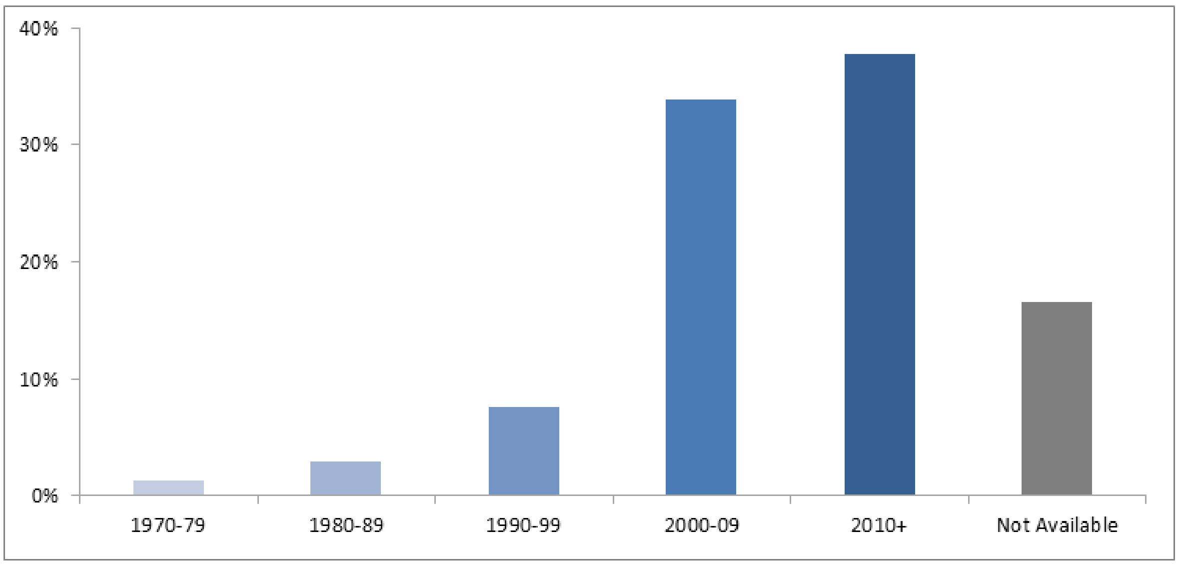

Figure 2 – Proportion of demand reduction campaigns for wildlife products by year of commencement.

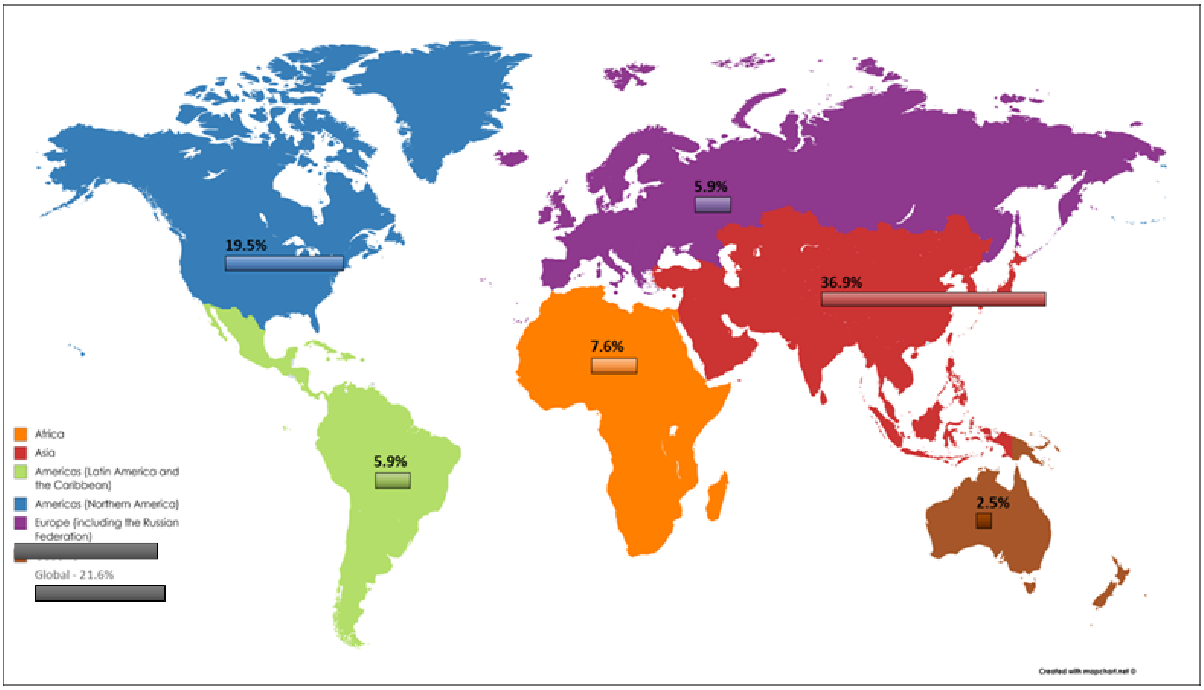

Figure 3 – Spatial distribution of wildlife demand reduction campaigns by macro-geographical region (based on the United Nations Geoscheme 2018). Global refers to demand reduction campaigns that targeted audiences across multiple continents.

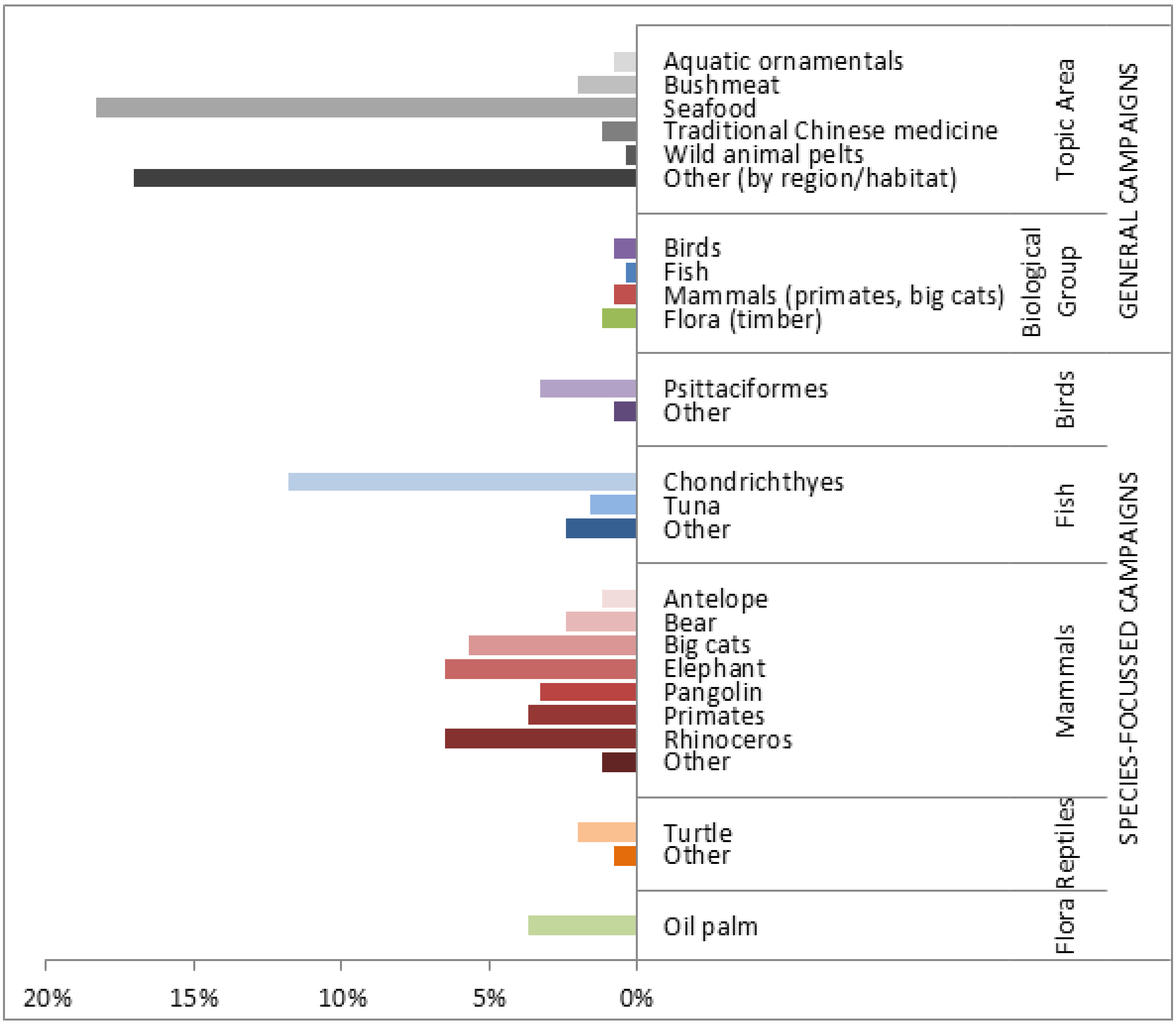

Figure 4 – Campaign focus by traded species or subject, based on record frequency. Records that specified targeted efforts on multiple individual species were split and considered as individual counts for species analysis (n=246). Species groups that had less than three records of campaigns were grouped and labelled under “Other” per biological group (see Supplementary Material – Table S4).

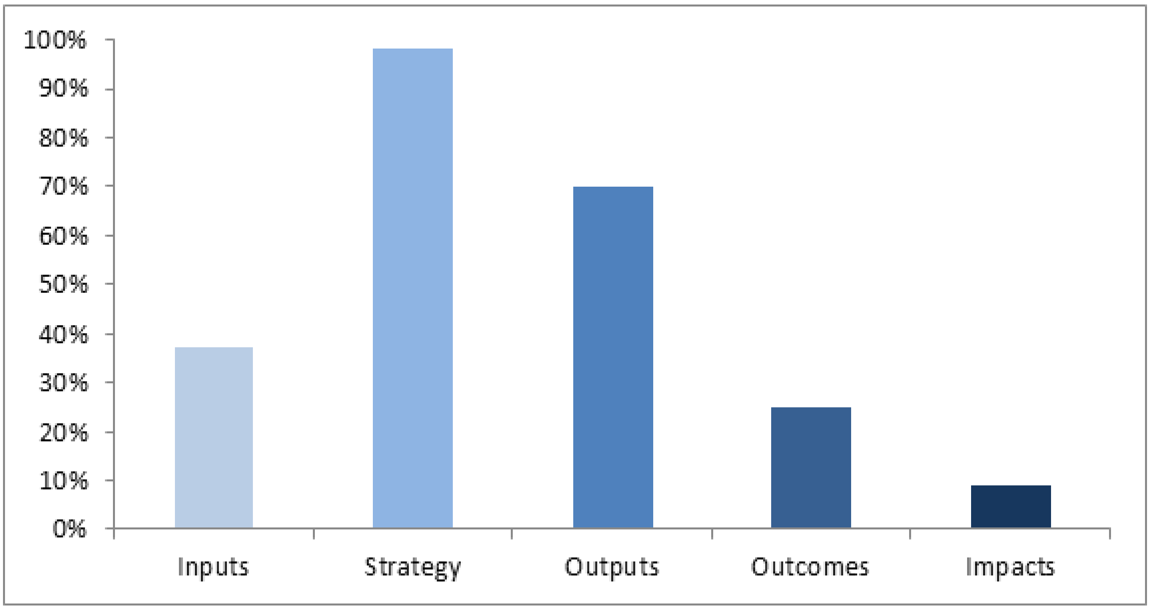

Figure 5 – Types of project monitoring and evaluation information reported on demand reduction campaigns for wildlife products

Accepted Article

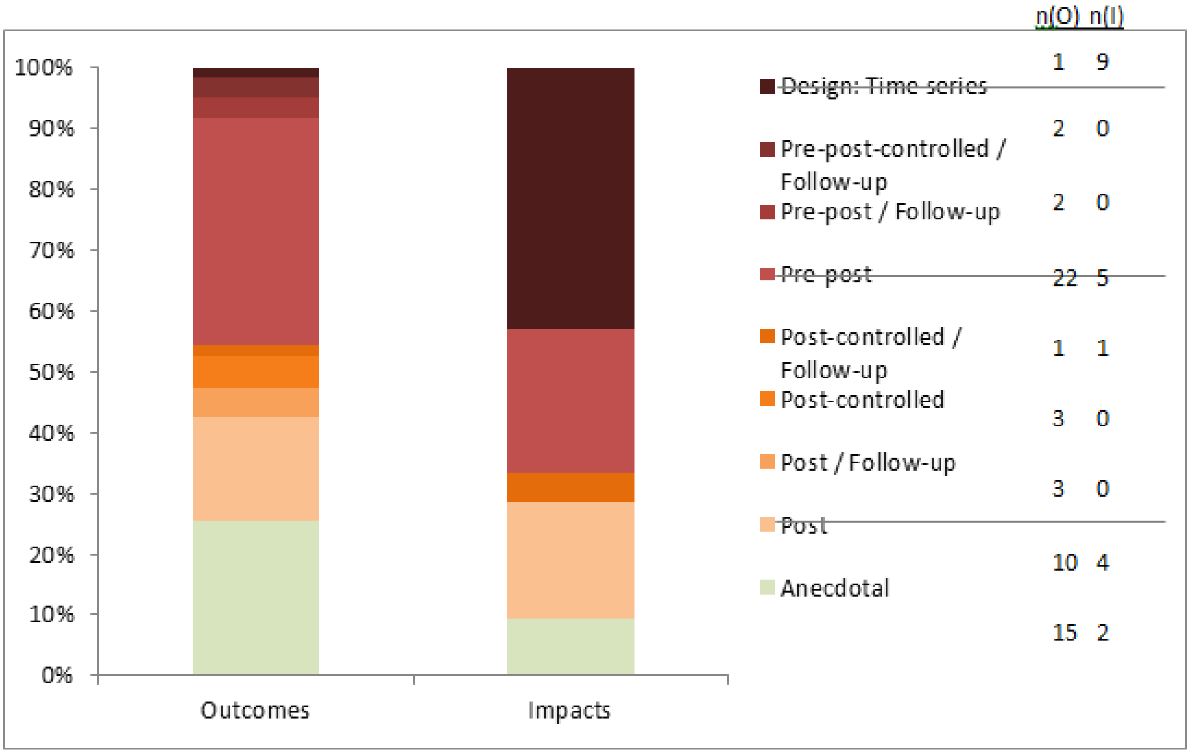

Figure 6 – Research design of wildlife demand reduction campaigns that reported outcomes and impacts