




# A framework for evaluating biodiversity mitigation metrics

Kevin Bracy Knight , Emily S. Seddon, Theodore P. Toombs

Received: 15 April 2019/Revised: 10 September 2019/Accepted: 14 September 2019/Published online: 12 October 2019

**Abstract** Biodiversity mitigation is a cornerstone of applied conservation. Mitigation encompasses a suite of practices, ranging from planned avoidance of impacts to creation of new natural habitats. Accurate and appropriate metrics quantifying impacts to natural systems and the effectiveness of restoration are necessary to measure the success of different mitigation strategies. Because effective mitigation requires adequate metrics, we developed a Biodiversity Metrics Framework to assist practitioners and policy makers in assessing biodiversity mitigation metrics. Based on Noss' Hierarchy of Biodiversity, the Scorecard highlights the mismatch between scientifically defined best practices and metrics required by policy. The Framework may serve a vital role in standardizing and validating mitigation projects into the future.

**Keywords** Biodiversity metrics · Biodiversity offsets · Conservation · Design · Mitigation

## INTRODUCTION

Given increasing impacts on ecosystems around the world (Ngo et al. 2019) and limited resources for conservation (Evans et al. 2012; Waldron et al. 2013), it is critical that conservation actions are effective and cost-efficient. Whether the goal of a conservation program is to preserve remaining natural resources or restore them, measuring the quantity and quality of the species, the ecosystem, or the natural resources is a necessary tool in determining efficacy. It is also important to identify which human activities

are impacting these natural resources and how much negative effect they have. Therefore, it is valuable to understand the type and amount of conservation actions necessary to ameliorate or reverse the anthropogenic impacts upon the system (Bull et al. 2017).

This is particularly important in the case in biodiversity mitigation (BM) programs (Maron et al. 2012; Gonçalves et al. 2015). The process of BM involves measuring anthropogenic impacts and addressing them through conservation action. However, a BM program may be less effective at achieving conservation goals (Maron et al. 2012, 2016) if it ineffectively measures impacts or restoration, either by requiring metrics that poorly assess biodiversity (Business and Biodiversity Offsets Programme (BBOP) 2012; Brownlie et al. 2013; Pilla 2014) or by failing to properly implement these measurements.

To address the need for better outcomes in the BM process, conservation science has created increasingly rigorous systems that define when and how conservation actions are applied. The Mitigation Hierarchy (International Finance Corporation 2012; M Ekstrom et al. 2015) is one example of such a system that has become a cornerstone in conservation policy worldwide (Arlidge et al. 2018). It defines sequential conservation actions that range from avoiding all impacts whenever possible, to minimizing necessary impacts to the system, and in cases where impacts cannot be avoided, implementing biodiversity offsets (BO) which recreate lost biodiversity in another location. However, while the Mitigation Hierarchy defines what types of conservation actions are appropriate, it does not define what metrics are appropriate to measure biodiversity impacts and conservation outcomes.

The goal of many BM programs is to achieve a “no net loss” of biodiversity (CITE). Whether or not a BM program achieves this goal depends, in part, on the metrics

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s13280-019-01266-y>) contains supplementary material, which is available to authorized users.

used to measure impacts and restoration actions (Gordon et al. 2011; Habib et al. 2013; van Teeffelen et al. 2014; Bull et al. 2017; Gamarra and Toombs 2017). Poorly measuring impacts or restoration may lead to incorrect management choices or, in a worst-case scenario, failure of the program to achieve its goal (Maron et al. 2012; Gonçalves et al. 2015; Van Bochove et al. 2016; Gourevitch et al. 2018). While the consequences of poorly chosen metrics may be severe in any type of BM program, they may be particularly so in the case of biodiversity offsetting (BO). The intent of BO programs is to create new resources, or restore damaged resources (habitat, ecosystems, etc.) as direct recompense for destroying or damaging these resources in another location. Improper measurement has great potential for harm under BO as areas of land may be reduced in function or converted completely and newly restored areas may not match or exceed the lost function (Maron et al. 2016). In part due to this, BO has a mixed to poor record of success. Studies regularly show that BO frequently fails to achieve the goal of “no net loss” when one area is impacted and another restored (Brownlie et al. 2013; Bezombes et al. 2019; zu Ermgassen et al. 2019).

Failure to achieve “no net loss” in mitigation is often attributed to either a failure of governments to mandate a suitably comprehensive conservation plan, planning for long-term viability, or a failure to develop a plan that comprehensively measures the correct elements of biodiversity within the focal system to assure the desired result (Bull et al. 2013). Although BO programs are a part of mitigation policy or practice in more than 45 countries, the number of transactions remains low (Madsen et al. 2011). Based on our experience with designing BO programs, and through communications with government agencies and practitioners around the world, the lack of strong policies with clear goals and required metrics that assure success is one reason for the limited success of these programs. This lack of clearly defined metrics also hinders researchers’ and practitioners’ ability to compare observed results from programs with similar conservation goals.

However, this failure to consistently develop adequate metrics for BM programs is not surprising. Ecosystems are notoriously complex, consisting of thousands of interacting biotic and abiotic processes (Pimm 1984; Cadenasso et al. 2006; Duffy et al. 2007). Determining which elements of any ecosystem may be particularly important to that system’s continuing healthy function or which are particularly impacted by human stressors is critical to any conservation action, particularly in the case of more invasive forms mitigation such as BO (Van Bochove et al. 2016).

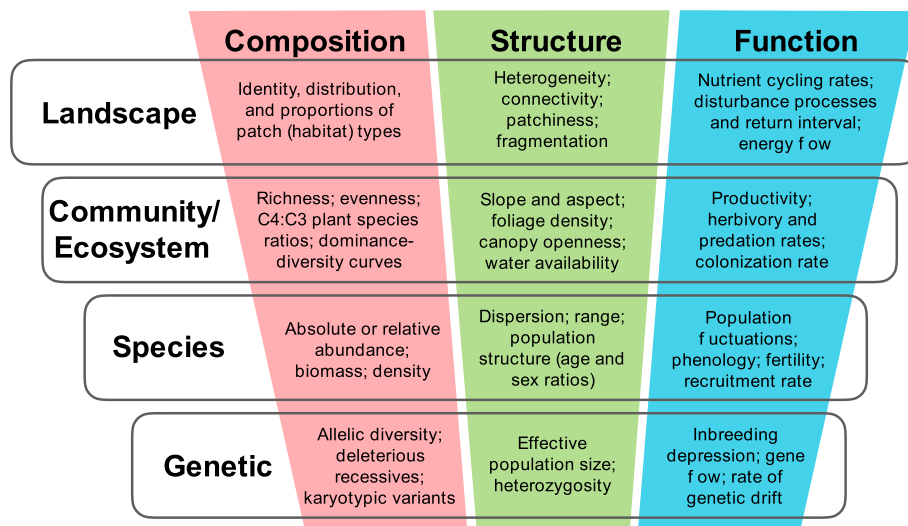
We see a clear need for a common language describing the metrics required by mitigation policy that addresses known elements of the focal systems (e.g., species richness, genetic diversity, or soil permeability).

To address this need, we propose a Biodiversity Metrics Framework (henceforth the “Framework”). We specifically designed the Framework to serve the following two purposes. First, the Framework is intended to aid in creating mitigation policy and programs. In this application, the Framework serves as a tool to ensure known conservation issues within the focal system (e.g., inbreeding depression, soil pH necessary for survival) are addressed by required metrics. Second, the Framework may be used as a tool for evaluating conservation policies, both determining the degree to which known conservation issues are addressed by a policy and facilitating easy comparison between policies with respect to metrics employed. We believe that the Framework, if regularly implemented in creation and review of mitigation policies, may become another cornerstone tool in conservation programs, which in turn may reduce the risk of further ecosystem degradation and species extinction.

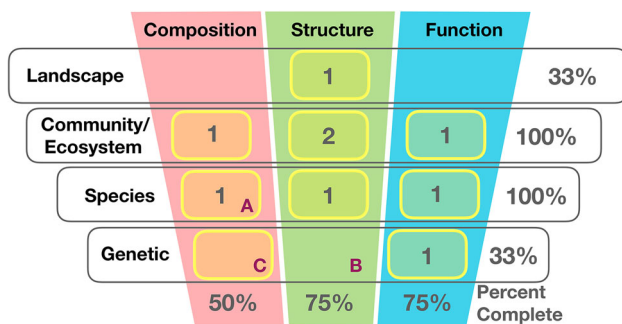
## THE BIODIVERSITY METRICS FRAMEWORK

The Framework is based upon Noss’ Hierarchy of Biodiversity (Fig. 1). Noss’ Hierarchy is well-known (cited by more than 1400 published works as of 2019) and provides a scaffold describing the three attributes of biodiversity (composition, structure, and function; Jerry and Franklin 1993) at four different scales ranging from genetic to regional inter-ecosystem (Noss 1983). Noss’ Hierarchy also provides indicators corresponding to these elements of biodiversity and even suggests common techniques to measure the various elements of biodiversity within an ecosystem. For example, a measure of structure at the regional/landscape-scale (the top, center cell in Fig. 1) may include spatial heterogeneity (an example of an indicator) conducted by applying spatial analysis to remotely sensed data (i.e., satellite or aerial imagery).

The Framework itself consists of two components: The Biodiversity Scorecard (“Scorecard,” see Fig. 2) and the Definitions & Descriptions (“D&Ds,” see Fig. 3). The Scorecard is an organizational tool and serves as a visual introduction to a description of a mitigation program. The D&Ds follow the Scorecard and describe the mitigation program, highlighting which metrics are required by policy and information from published scientific literature describing specific components of biodiversity that are critical to conserving the focal system. The D&Ds are divided into sections that correspond to, and serve as the basis for, each cell of the Scorecard. The user would first read the Scorecard for a quick review of the comprehensiveness of the mitigation policy and then read the D&Ds to gain a more in-depth understanding. We discuss the creation and interpretation of both of these components below, beginning with the D&Ds, as the Scorecard uses information derived from them.



**Fig. 1** Noss' Hierarchy of Biodiversity (adapted from Noss 1990). The red, green, and blue columns each represent one of the three primary attributes of biodiversity (Franklin et al. 1981). Rows represent the scales from the largest (Landscape) to the smallest (Genetic). Cells (the nexus of a row and column) are each one "element" of biodiversity within the hierarchy as described by Noss. Indicators of each element (samples from the original Noss manuscript) appear in each cell. For in-depth descriptions of each element as well as recommended methods for assessing them, see Noss (1990)



**Fig. 2** An example application of the Biodiversity Metrics Scorecard. Row and column definitions are adapted from Noss' hierarchy of biodiversity (Noss 1990) and define nested elements of biodiversity ranging in scale from genetic to landscape (inter-ecosystem). Numbers within the cells describe the number of metrics that assess that particular aspect of biodiversity. Summary cells (rightmost column and bottom row) describe how comprehensively the metrics prescribed by the offsetting policy/program address the ecosystem. Summary cells list the percent of the cells within the respective row or column that have any metric required by policy. Yellow highlighted cells indicate elements of biodiversity identified in literature review of published research as critical for the focal species or system. Mismatch (cells with yellow highlights, but no metric) indicates areas where conservation targets are not explicitly addressed by metrics employed by the mitigation program. The purple letters "A," "B," and "C" are reference for the manuscript only and not part of the Scorecard

## Definitions and descriptions

The purpose of the Definitions and descriptions (D&Ds) is to summarize the best available science regarding the conservation of the focal system, and the link metrics required by the policy to known impacts on the system. Each

subsection of the D&Ds focuses on one element (one cell in the Biodiversity Scorecard) in Noss' hierarchy (see Fig. 3 as an example of a D&D for the species-level composition cell in the Scorecard). Each of these subsections list (A) threats or impacts to the focal system, and (B) metric(s) required by the policy or plan to measure each threat or impact as part of mitigation. While we recommend in-depth descriptions of each of these later in a policy document, each subsection of the D&D should be concise with numbered lists (see Fig. 3 and Appendix S1 for examples) linking threats to the focal system and metrics that measure these impacts or actions related to these threats. For example, if the first enumerated threat to the focal species is extirpation through spread of disease when the population becomes too dense, the first listed metric employed should correspond to this threat. Relevant, peer-reviewed, scientific evidence for both the nature of the threat and the viability of the metric as a tool to measure it should be cited for each enumerated point. While metrics described in the D&Ds may directly measure the element in question (e.g., a field count of the population as a means of measuring abundance), they may also be indirect, or proxy, assessments. As an example of an indirect metric, the total wintering monarch butterfly population in North America is estimated based on a number of hectares of forest that are occupied during their winter torpor rather than a rigorous count of each individual. Here, the justification for this indirect assessment should be included in the cited scientific literature in the D&D for species-level composition.

The process of creating the D&Ds is the same whether it is for the purposes of creating a policy document,

COMPOSITION	LEVEL: SPECIES
Absolute or relative abundance; frequency; importance or cover value; biomass; density	
<b>A. THREATS OR IMPACTS:</b>	
<ol style="list-style-type: none"> <li>1. Utah prairie dog population size: viable populations are the primary goal of the mitigation practices as declines are reported range-wide (USFS 2015).</li> <li>2. Sylvatic plague: High population density is identified as increasing risk of plague-related extirpations among Utah prairie dog colonies (Barnes 1993).</li> </ol>	
<b>B. METRIC(S) REQUIRED:</b>	
<ol style="list-style-type: none"> <li>1. Utah prairie dog abundance: annual population surveys are required both for impact and offset locations. Although the Utah Prairie Dog Final revised recovery plan (1991) defines survey methodology, it provides no scientific justification for any methods as being appropriate for survey of this species.</li> <li>2. Utah prairie dog density: population density is <u>not</u> directly assessed as part of metrics and is not part of credit calculation. However, overall size of the colony is calculated as part of the mitigation program, so density <u>could</u> be easily be calculated given the metrics of abundance and colony size.</li> </ol>	

**Fig. 3** Example of subsection of “Definitions & Descriptions” (D&D) section. The header of this subsection refers to the cell (intersection of row and column) of the Biodiversity Scorecard (Fig. 2). The description (subtext within the header) is text describing elements of biodiversity adapted from (Noss 1990). The alphabetically listed sections (A & B) allow the reader to identify known threats or impacts to the system (A), and which metrics are required by policy to address the impacts in a mitigation program (B). Peer-reviewed evidence for the threats/impacts particularly important to the focal system and the efficacy of required metrics for assessing each threat should be cited as well. For clarity, each cell of the Biodiversity Scorecard should have a subsection with the D&D regardless of whether or not there are metrics addressing that element of the hierarchy. A complete D&D for Utah Prairie Dog (provided for example only) with instructions is available in Appendix S1

reviewing a policy, or comparing several policies. In each case, we recommend the author to first perform a literature review of the focal system, particularly focused upon peer-reviewed articles to assure information is based upon the best possible science. If the focal system has a paucity of information, the literature review should include the closest possible analogous systems (e.g., closely related species, similar hydrology to that of the focal system, but in a different geographic location, etc.). If using analogs, the author should provide notes in the manuscript (i.e., the policy or review of policy) justifying the use of these analogs and stating potential differences between the focal system and the chosen analogous systems. We recommend highlighting and annotating all literature used in this review process to match that of the relevant subsection of the D&D and Scorecard (e.g., use “Species-Level Composition” or “C2” as annotation for highlighted text in published scientific texts). This technique may also be employed when creating a Framework as a review of an existing policy as a way to annotate the policy document, highlighting both the scientific evidence of impacts and also the metrics required by the policy to measure mitigation of these impacts.

In the example provided (Fig. 3), we give one subsection of a D&D for the Utah prairie dog (*Cynomys parvidens*). A complete Framework for this mitigation program is provided in the supplemental to this publication (Appendix S1). Literature review of risks for this species identified two conservation concerns for Species-level Composition: population size and population density. Here, the United States Forest Service (2015) finds that the species’ total population continues to decline and that local populations are declining below minimum viability thresholds. This threat is enumerated as “1” in the “Threats or Impacts” section of the D&D. The metric required by the offset policy to address this threat is to perform annual population surveys at both the impact and offset sites. This metric corresponds to the first threat in the section above, and therefore also is listed as “1.” The second threat identified in the literature review for this species relevant to this D&D subsection is sylvatic plague. This disease is associated with high population densities of prairie dogs (the disease is primarily spread by close contact and fleas). In this case, there is no metric required as part of the mitigation program to address this threat. We note this in the line marked “2.” In addition, we state that population

density *could* be calculated for this program because both colony area and annual population surveys are required as part of this mitigation program.

### The biodiversity scorecard

The Biodiversity Scorecard (Fig. 2) is designed to be a summary table describing the metrics required to mitigate the focal system. As described below, each cell (the nexus of a row and column) represents an element of biodiversity within the focal system. Highlighted areas and numbers in the cells of the Scorecard refer to information in the D&Ds. As such, we recommend that the Scorecard precedes the D&Ds in any policy document (or formal review of policy or comparison of policies).

As with Noss' hierarchy, the Scorecard is composed of three columns and four rows of cells describing elements of biodiversity. The three columns represent the three primary attributes of ecosystems: composition, structure, and function (Jerry and Franklin 1993). The four rows describe the focal system at a range of scales from within a species (genetic and population scale) to the interactions of patches of ecosystems within a landscape. Each cell (the nexus of a row and column) represents one element of Noss' hierarchy. For example, the cell in Fig. 2 labeled "A" describes the species-level composition of the focal system. Any number within a cell (in the case of the "A" cell in Fig. 2, "1") describes the number of metrics required by the conservation policy to measure this element. Cells highlighted in yellow (for example, in Fig. 2 cell "A," but not "B") represent elements of biodiversity identified by a literature review as of particular conservation significance for the focal system. For example, the primate species Golden Lion Tamarin (*Leontopithecus rosalia*) is particularly vulnerable to extinction due to inbreeding depression (Dietz et al. 2000). When creating a Scorecard for a conservation program focused on this species of Tamarin, an author would place a yellow highlight in the Function cell at the Genetic scale, because Noss (1990) classifies inbreeding depression as this type of element. Also, see Noss (1990) for in-depth descriptions of each of the elements and biodiversity indicators that describe them.

In addition to describing both the elements of biodiversity considered important for the conservation of the focal system (yellow highlighted cells) and elements that policy requires measurement of (numbers within cells), the Scorecard provides an accounting of the comprehensiveness of the metrics required by the policy. The bottom row and rightmost column of the Scorecard summarize the percent of Noss' elements that are addressed by the required metrics. These values describe the number of cells in a given row (scale level) or column (attribute) as a percent of total cells. While any cell may have more than

one metric addressing it, and thus, a number higher than one (e.g., if population counts of two different species are required by a policy, the species-scale composition cell would have the number "2" recorded), we intend the summary rows and columns to illuminate the comprehensiveness of the mitigation policy with respect to the breadth and depth of an ecosystem. This is because one noted cause of poorer than expected outcomes in BM is the failure to account for the complexities of the focal system when taking action (Bull et al. 2013; Gelcich et al. 2017). This is particularly the case when engaging in offsetting practices that involve the creation or restoration of ecosystems or habitat, such as wetland mitigation or other similarly intensive practices (Brown and Veneman 2001; Robb 2002; BenDor 2009; Vaissière and Levrel 2015).

A reader interpreting the Scorecard should first note the yellow highlighted cells showing the elements of biodiversity identified by published research as particularly valuable in conserving the system. Then, the user should note the yellow highlighted cells that lack a number inside the cell (e.g., the cell marked "C" in Fig. 2). These are instances where research has identified an element of biodiversity as noteworthy in the conservation of the system that the policy does not measure. Alternately, there may be instances where a cell contains a number and yet no yellow highlight. This indicates that a metric, or multiple metrics if the number is greater than one, required by the mitigation policy does not correspond to an element of biodiversity identified by peer-reviewed literature as particularly valuable to the focal system. In either case, a mismatch between identified risks for the focal system (yellow highlights) and metrics required to assess the risks (numbers within the cells) may indicate that a mitigation policy is at risk of decreased success. Because the information provided within the Scorecard itself is limited, an evaluation of a mitigation policy should begin with reading it, rather than end with it. Evaluation should then continue with reviewing the D&Ds and finally any in-depth descriptions of metrics (if needed) and guides for application as part of a mitigation project.

### APPLYING THE FRAMEWORK

We envision two main applications for the Framework: as a tool aiding in the creation of policy documents and as a means of evaluating the comprehensiveness of the metrics required by a policy or comparing policies. The primary use for the Biodiversity Framework, as a tool that organizes mitigation policy documents in a standardized fashion, is intended to summarize the often necessarily complex policy documents. While the Framework is not intended to replace a comprehensive policy document, it is a means of



concisely summarizing the often necessarily complex policy documents describing the mitigation metrics required in a policy. Additionally, in utilizing the Scorecard portion of the Framework, policy creators may identify flaws in the mitigation policy before it is adopted. A mismatch between vulnerabilities of the focal system and metrics (either a case where research identifies a particular vulnerability and there are no metrics required to assess it in the policy, or if a metric is required that does not address a vulnerability) may indicate risk that the policy may not meet conservation goals. As previously stated, a leading cause of mitigation programs failing to achieve goals is a lack of comprehensive metrics that address both short- and long-term vulnerabilities to a system (Bull et al. 2013; Gonçalves et al. 2015; Maron et al. 2016). By using the Framework as a tool to create policy, mismatches may be identified, and hopefully rectified, before policy is adopted. Alternately, if a mismatch is not addressed requiring additional metrics, the policy could flag these mismatches and suggest means of addressing them in the future if desired mitigation outcomes are not being achieved.

The second application of the Framework is as a means of assessing existing mitigation policies. Reviews of the relative success or failure of mitigation policies are common. Although poor mitigation outcomes are frequently attributed to poor choice of metrics for the focal system (at least as a contributing factor), no tool exists to readily identify which metrics may have been poorly chosen or which particular vulnerabilities of the focal system were not addressed by metrics. Because the Scorecard considers the composition, structure, and function of a focal system and assesses the comprehensiveness of these attributes represented in metrics outlined in mitigation policy, it may predict long-term consequences for other elements of the focal ecosystem that may not have been addressed by metrics. For example, consider an offsetting program that relocates a threatened arboreal primate species to a newly restored area of habitat. If the policy requires population counts of the species (the Species-level Composition element on the Scorecard) but neglects to measure canopy connectedness (Community-Ecosystem-level Structure), the species may establish in the short-term, but ultimately become extirpated at that location. Lastly, it is likely of interest to compare mitigation policies for the same, or similar, focal systems. Whether the same focal system is being mitigated by different policies (for example, one species found in multiple countries) or two different, but similar focal systems, a tool for evaluating metrics may provide a valuable first step in comparison. Comparing the Scorecards for two different programs with different success rates may suggest that a metric employed in one program absent from another may be particularly critical to success, or at least a point worthy of further investigation.

## DISCUSSION

We began this research as an investigation of which metrics were used to measure biodiversity impacts and conservation outcomes in BO programs around the world. Specifically, we set out to document the comprehensiveness of these metrics with respect to Noss' Hierarchy of Biodiversity, hoping to identify potential biases in required metrics that may lead to poorer than expected outcomes in this particularly sensitive type of mitigation. In spite of the fact that BO programs are now in use or nascent in at least 45 countries (Madsen et al. 2011), policy documents defining the BO programs in these countries rarely define concise metrics.

As one example of how the Framework may add clarity and transparency, consider the biodiversity offset program at the Rio Tinto QIT Minerals Site in Madagascar. Rio Tinto has partnered with IUCN and other NGO organizations to develop and help administer this program, and it is often considered a model of success for a private company volunteering to mitigate the considerable impacts of mining through and offset program (Temple et al. 2012; Bidaud et al. 2015). Documentation describing methods for this program provides programmatic goals (e.g., “no net loss”) and describes how offsetting “currency” is calculated by assessing quality multiplied by quantity or land area as a percent of the total remaining geographic range of focal species (Temple et al. 2012), but it does not provide clear methods or tools to assess whether these metrics are adequate for the goals of the program. If the Framework were to have been used in the case of Rio Tinto, it could fill this gap. For example, the Scorecard would provide an overall sense of whether the metrics were comprehensive enough to address the conservation goal, and the D&D would provide the details of how the metrics were derived, and the peer-reviewed scientific research justifying the need for each metric. It may be that these considerations were made during the development of the Rio Tinto program, but since they were not clearly presented and made available to the public through any documentation, it makes it difficult for others to evaluate the program. The Framework addresses this problem by clearly organizing all the necessary the information within a sound and foundational science framework (i.e., Noss 1990). This is particularly valuable for conservation-focused agencies within the Malagasy government or NGOs concerned with the long-term viability of the species and ecosystems impacted by this program as it may shed light on particularly useful metrics in conservation, or those that fall short.

Without a standardized system of review and comparison of the metrics employed in BM, it is unclear whether successes and failures can be attributed to poorly executed policies or poorly designed ones (Curran et al. 2015).

Further, both the lack of clarity in policies and poor success in achieving offsetting goals were consistently stated in our communications with practitioners as reasons why BO was not more readily employed.

The Biodiversity Metrics Framework elucidates which biodiversity components *are* and *are not* assessed in biodiversity mitigation programs, identifying mismatches between the stated goals of policy and the metrics employed to measure them. By doing so, the Framework may flag accidental oversights that could lead to inaccurate measurements of program outcomes. Since the Framework is grounded in Noss' Hierarchy, it can provide scientific justification for the metrics required by mitigation in an organized fashion that can be applied to any BM policy.

To be most effective, we believe the Framework should be used in conjunction with other mitigation planning tools such as the Mitigation Hierarchy (ten Kate et al. 2004; BBOP and UNEP 2015), the Metrics Decision Tree (Gamarra et al. 2018), and the site selection protocol described in Kiesecker et al. (2009) to increase accountability and the likelihood of successful conservation. Each of these other tools may prove critical to the success of a conservation program: The Mitigation Hierarchy provides guidance as to what type of conservation action is preferred to best assure “no net loss.” The Metrics Decision Tree defines which general types of metrics may be applicable given available data on a focal species. Kiesecker's site selection tool identifies target areas based on landscape metrics. The benefit of including the Framework into this BM is a means of organizing and explicitly clarifying which metrics are appropriate for the focal system based upon the best available science. This clarity is necessary to make sure that biodiversity is appropriately and consistently measured under a policy and is a necessary component for each of the mitigation planning tools described above.

We present the Biodiversity Metrics Framework as a working draft of this tool that will hopefully be improved as it is applied. It is limited to addressing elements of biodiversity, rather than a comprehensive assessment of the effectiveness of a policy. As such, the Framework does not measure socioeconomic factors such as cost-effectiveness, support by stakeholders, or a myriad of other metrics that might be used to assess a mitigation policy. It also does not assess compliance, which is another considerable challenge in achieving mitigation goals (Walker et al. 2009; van Teeffelen et al. 2014; Lindenmayer et al. 2017). However, the Framework serves to standardize, and make more transparent, one portion of the creation and evaluation of mitigation policies.

Better BM driven by better accounting of losses and gains is increasingly necessary in the face of continuing human impacts and clear indications that current efforts at ameliorating these impacts have failed (Barnosky et al.

2011; Alvarado-Quesada et al. 2014; Ceballos et al. 2015). Our intent in developing the Framework is to advance the ongoing conversation among the scientific and practitioner communities focused on the improvement of metrics and BM policy. The Framework offers a tangible example of how such a science-based tool could contribute to a more comprehensive assessment of biodiversity, and improve the effectiveness and legitimacy of BM.

**Acknowledgements** This work was supported, in part, by the Marie-Jose and Henry R Kravis Foundation. We thank Joseph Bull, Eric Anderson, David Wolfe, and Scott Chiavacci for feedback on biodiversity metrics and the thoughtful reviewers for feedback on manuscript drafts.

## REFERENCES

- Alvarado-Quesada, I., L. Hein, and H.P. Weikard. 2014. Market-based mechanisms for biodiversity conservation: A review of existing schemes and an outline for a global mechanism. *Biodiversity and Conservation*. <https://doi.org/10.1007/s10531-013-0598-x>
- Arlidge, W.N.S.S., J.W. Bull, P.F.E.E. Addison, M.J. Burgass, D. Gianuca, T.M. Gorham, C.D.S.C. Jacob, N. Shumway, et al. 2018. A global mitigation hierarchy for nature conservation. *BioScience*. <https://doi.org/10.1093/biosci/biy029>.
- Barnosky, A.D., N. Matzke, S. Tomiya, G.O.U. Wogan, B. Swartz, T.B. Quental, C. Marshall, J.L. McGuire, et al. 2011. Has the Earth's sixth mass extinction already arrived? *Nature* 471: 51–57. <https://doi.org/10.1038/nature09678>.
- BBOP & UNEP. 2015. A cross-sector guide for implementing the Mitigation Hierarchy: Executive summary and overview.
- BenDor, T. 2009. A dynamic analysis of the wetland mitigation process and its effects on no net loss policy. *Landscape and Urban Planning* 89: 17–27. <https://doi.org/10.1016/j.landurbplan.2008.09.003>.
- Bezombes, L., C. Kerbiriou, and T. Spiegelberger. 2019. Do biodiversity offsets achieve no net loss? An evaluation of offsets in a french department. *Biological Conservation* 231: 24–29. <https://doi.org/10.1016/j.biocon.2019.01.004>.
- Bidaud, C., M. Hrabanski, and P. Meral. 2015. Voluntary biodiversity offset strategies in Madagascar. *Ecosystem Services* 15: 181–189. <https://doi.org/10.1016/j.ecoser.2015.02.011>.
- Van Bochove, J.-W., R. Mitchell, J. Pilgrim, L. Bennun, and J. Ekstrom. 2016. *Integrating IFC's Performance Standard 6 into impact assessment*. Vol. 6.
- Brown, S.C., and P.L.M. Veneman. 2001. Effectiveness of compensatory wetland mitigation in Massachusetts, USA. *Wetlands* 21: 508–518.
- Brownlie, S., N. King, and J. Treweek. 2013. Biodiversity tradeoffs and offsets in impact assessment and decision making: can we stop the loss? *Impact Assessment and Project Appraisal* 31: 24–33. <https://doi.org/10.1080/14615517.2012.736763>.
- Bull, J.W., K.B. Suttle, A. Gordon, N.J. Singh, and E.J. Milner-Gulland. 2013. Biodiversity offsets in theory and practice. *Oryx* 47: 369–380. <https://doi.org/10.1017/S003060531200172X>.
- Bull, J.W., S.P. Lloyd, and N. Strange. 2017. Implementation gap between the theory and practice of biodiversity offset multipliers. *Conservation Letters* 00: 1–14. <https://doi.org/10.1111/conl.12335>.
- Business and Biodiversity Offsets Programme (BBOP). 2012. *Biodiversity Offset Design Handbook—updated*: 112.

- Cadenasso, M., S.T.A. Pickett, and M. Grove. 2006. Dimensions of ecosystem complexity: Heterogeneity, connectivity, and history. *Ecological Complexity* 3: 1–12. <https://doi.org/10.1016/j.ecocom.2005.07.002>.
- Gamarra, M.J.C., and T.P. Toombs. 2017. Thirty years of species conservation banking in the U.S.: comparing policy to practice. *Biological Conservation* 214: 6–12. <https://doi.org/10.1016/j.biocon.2017.07.021>.
- Gamarra, C., J. Maria, J.P. Lassoie, J.C. Milder, M.J.C. Gamarra, J.P. Lassoie, and J.C. Milder. 2018. Accounting for no net loss: A critical assessment of biodiversity offsetting metrics and methods. *Journal of Environmental Management* 220: 36–43. <https://doi.org/10.1016/j.jenvman.2018.05.008>.
- Ceballos, G., P.R. Ehrlich, A.D. Barnosky, A. García, R.M. Pringle, and T.M. Palmer. 2015. Accelerated modern human—Induced species losses: entering the sixth mass extinction. *Sciences Advances* 1: 1–5. <https://doi.org/10.1126/sciadv.1400253>.
- Curran, M., S. Hellweg, and J. Beck. 2015. The jury is still out on biodiversity offsets: Reply to Quéfier et al. *Ecological Applications* 25: 1741–1746. <https://doi.org/10.1890/14-2363.1>.
- Dietz, J. M., A. J. Baker, and J. D. Ballou. 2000. Demographic evidence of inbreeding depression in wild golden lion tamarins. In *Genetics, Demography and Viability of Fragmented Populations*, ed. A. G. Young and G. M. E. Clarke, 203–212. Conservation Biology. Cambridge: Cambridge University Press. <https://doi.org/10.1017/cbo9780511623448.014>.
- Duffy, J.E., B.J. Cardinale, K.E. France, P.B. McIntyre, E. Thébault, and M. Loreau. 2007. The functional role of biodiversity in ecosystems: Incorporating trophic complexity. *Ecology Letters* 10: 522–538. <https://doi.org/10.1111/j.1461-0248.2007.01037.x>.
- Ekstrom, J.M., L.A. Bennun, and R. Mitchell. 2015. *A cross-sector guide for implementing the mitigation hierarchy*. Cambridge, UK.
- Evans, D.M., P. Barnard, L.P. Koh, C.A. Chapman, R. Altwegg, T.W.J. Garner, M.E. Gompper, I.J. Gordon, et al. 2012. Funding nature conservation: Who pays? *Animal Conservation* 15: 215–216. <https://doi.org/10.1111/j.1469-1795.2012.00550.x>.
- Franklin, J.F., W. Denison, A. Mckee, and F. Swanson. 1981. *Ecological characteristics of old-growth Douglas-fir forests*. Portland, OR. <https://doi.org/10.2737/PNW-GTR-118>.
- Gelcich, S., C. Vargas, M.J. C. J.C. Castilla, and C.J. Donlan. 2017. Achieving biodiversity benefits with offsets: Research gaps, challenges, and needs. *Ambio* 46: 184–189. <https://doi.org/10.1007/s13280-016-0810-9>.
- Gonçalves, B., A. Marques, A.M.V.D.M. Soares, and H.M. Pereira. 2015. Biodiversity offsets: From current challenges to harmonized metrics. *Current Opinion in Environmental Sustainability* 14: 61–67. <https://doi.org/10.1016/j.cosust.2015.03.008>.
- Gordon, A., W.T. Langford, J.A. Todd, M.D. White, D.W. Mullerworth, and S.A. Bekessy. 2011. Assessing the impacts of biodiversity offset policies. *Environmental Modelling & Software* 26: 1481–1488. <https://doi.org/10.1016/j.envsoft.2011.07.021>.
- Gourevitch, J., C.C. Nicholson, L.J. Sonter, A.J. Schwartz, T.H. Ricketts, N.K. Singh, L.L. Richardson, I. Koh, et al. 2018. Biodiversity offsets may miss opportunities to mitigate impacts on ecosystem services. *Frontiers in Ecology and the Environment* 16: 143–148. <https://doi.org/10.1002/fee.1781>.
- Habib, T.J., D.R. Farr, R.R. Schneider, and S. Boutin. 2013. Economic and ecological outcomes of flexible biodiversity offset systems. *Conservation Biology* 27: 1313–1323. <https://doi.org/10.1111/cobi.12098>.
- International Finance Corporation. 2012. IFC Performance Standards on Environmental and Social Sustainability.
- Jerry, F., and Franklin. 1993. Preserving biodiversity: Species, ecosystems, or landscapes? *Ecological Applications* 3: 202–205.
- Kiesecker, J.M., H. Copeland, A. Pocewicz, N. Nibbelink, B. McKenney, J. Dahlke, M. Holloran, and D. Stroud. 2009. A framework for implementing biodiversity offsets: Selecting sites and determining scale. *BioScience* 59: 77–84. <https://doi.org/10.1525/bio.2009.59.1.11>.
- Lindenmayer, D.B., M. Crane, M.C. Evans, M. Maron, P. Gibbons, S. Bekessy, and W. Blanchard. 2017. The anatomy of a failed offset. *Biological Conservation* 210: 286–292. <https://doi.org/10.1016/j.biocon.2017.04.022>.
- Madsen, B., N. Carroll, D. Kandy, and G. Bennett. 2011. 2011 Update: State of biodiversity markets. Forest Trends’s Ecosystem Marketplace. Washington D.C.
- Maron, M., R.J. Hobbs, A. Moilanen, J.W. Matthews, K. Christie, T.A. Gardner, D.A. Keith, D.B. Lindenmayer, et al. 2012. Faustian bargains? Restoration realities in the context of biodiversity offset policies. *Biological Conservation* 155: 141–148. <https://doi.org/10.1016/j.biocon.2012.06.003>.
- Maron, M., C.D. Ives, H. Kujala, J.W. Bull, F.J.F. Maseyk, S. Bekessy, A. Gordon, J.E.M. Watson, et al. 2016. Taming a wicked problem: Resolving controversies in biodiversity offsetting. *BioScience*. <https://doi.org/10.1093/biosci/biw038>.
- Ngo, H. T., M. Guèze, J. Agard Trinidad, A. Arneith, P. Balvanera, K. Brauman, R. Watson, I. Baste, et al. 2019. *Global assessment report on biodiversity and ecosystems*. New York.
- Noss, R.F. 1983. A regional landscape approach to maintain diversity. *BioScience* 33: 700–706. <https://doi.org/10.2307/1309350>.
- Noss, R.F. 1990. Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology* 4: 355–364. <https://doi.org/10.1111/j.1523-1739.1990.tb00309>.
- Pilla, E. 2014. Towards the development of metrics for no net loss of biodiversity in Peru. *Inter-American Development Bank*. <https://doi.org/10.1006/bijl.1996.0096>.
- Pimm, S.L. 1984. The complexity and stability of ecosystems. *Nature* 307: 321–326. <https://doi.org/10.1038/307321a0>.
- Robb, J.T. 2002. Assessing wetland compensatory mitigation sites to aid in establishing mitigation ratios. *Wetlands* 22: 435–440.
- Temple, H. J., S. Anstee, J. Ekstrom, J. D. Pilgrim, J. Rabenantoandro, J.-B. Ramanamanjato, F. Randraitafika, and M. Vincellette. 2012. *Forecasting the path towards a Net Positive Impact on biodiversity for Rio Tinto QMM. IUCN and Rio Tinto Technical Series No. 2*. ISBN: 978-2-8317-1441-7
- ten Kate, K., J. Bishop, R. Bayon, K. ten Tate, J. Bishop, and R. Bayon. 2004. Biodiversity offsets: Views, experience, and the business case. *Diversity*. ISBN: 2-8317-0854-0.
- United States Department of Agriculture. 2015. *Common Questions & Answers BLM-USFS Greater Sage-Grouse Conservation Plans*.
- Vaissière, A.C., and H. Levrel. 2015. Corrigendum to “Biodiversity offset markets: What are they really? An empirical approach to wetland mitigation banking” [ECOLEC 110 (2015) 81–88]. *Ecological Economics* 114: 242. <https://doi.org/10.1016/j.ecolecon.2015.04.007>.
- van Teeffelen, A.J.A., P. Opdam, F. Wätzold, F. Hartig, K. Johst, M. Drechsler, C.C. Vos, S. Wissel, et al. 2014. Ecological and economic conditions and associated institutional challenges for conservation banking in dynamic landscapes. *Landscape and Urban Planning* 130: 64–72. <https://doi.org/10.1016/j.landurbplan.2014.06.004>.
- Waldron, A., A.O. Mooers, D.C. Miller, N. Nibbelink, D.W. Redding, and T.S. Kuhn. 2013. Targeting global conservation funding to limit immediate biodiversity declines. *Proceedings of the National Academy of Sciences* 110: 1–5. <https://doi.org/10.5061/dryad.p69t1>.
- Walker, S., A.L. Brower, R.T.T.T. Stephens, and W.G. Lee. 2009. Why bartering biodiversity fails. *Conservation Letters* 2: 149–157. <https://doi.org/10.1111/j.1755-263X.2009.00061.x>.



zu Ermgassen, S.O.S.E., J. Baker, R.A. Griffiths, N. Strange, M.J. Struebig, and J.W. Bull. 2019. The ecological outcomes of biodiversity offsets under “no net loss” policies: A global review. *Conservation Letters*. <https://doi.org/10.1111/conl.12664>.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## AUTHOR BIOGRAPHIES

**Kevin Bracy Knight** (✉) is a research scientist with the Office of the Chief Scientist at Environmental Defense Fund. His research interests include large-scale ecology, conservation prioritization, and anthropogenic impacts on natural systems.

*Address:* Environmental Defense Fund, 2060 Broadway, Suite 300, Boulder, CO 80304, USA.  
e-mail: [kbracyknight@edf.org](mailto:kbracyknight@edf.org)

**Emily S. Seddon** is an independent conservation scientist. Her key experiences include studying the effects of anthropogenic impacts on biodiversity, designing landscape-scale spatial conservation plans, and restoring native habitats to preserve ecological health and integrity.

*Address:* Loveland, CO, USA.  
e-mail: [seddon28@gmail.com](mailto:seddon28@gmail.com)

**Theodore P. Toombs** is conservation scientist within the Ecosystems Group at Environmental Defense Fund. His research interests include socio-ecology, systems thinking, and conservation.

*Address:* Environmental Defense Fund, 2060 Broadway, Suite 300, Boulder, CO 80304, USA.  
e-mail: [ttoombs@edf.org](mailto:ttoombs@edf.org)