Bull, Joseph W., Strange, Niels, Smith, Robert J. and Gordon, Ascelin (2020)
Reconciling multiple counterfactuals when evaluating biodiversity conservation impact in social-ecological systems. Conservation Biology . ISSN 0888-8892.

Reconciling multiple counterfactuals when evaluating biodiversity conservation impact in social-ecological systems

Abstract:	When evaluating the impact of a biodiversity conservation intervention, a 'counterfactual' is needed, as true experimental controls are typically unavailable. Counterfactuals are possible alternative system trajectories in the absence of an intervention and comparing observed outcomes against the chosen counterfactual allows the impact (change attributable to the intervention) to be determined. Since counterfactuals are hypothetical scenarios, and by definition never occur, they must be estimated. Sometimes there may be many plausible counterfactuals, given that they can include multiple drivers of biodiversity change, and be defined on a range of spatial or temporal scales. Here we posit that, by definition, conservation interventions always take place in social-ecological systems (SES; ecological systems integrated with human actors). Evaluating the impact of an intervention within an SES therefore means taking into account the counterfactuals assumed by different human actors. Use of different counterfactuals by different actors will give rise to perceived differences in the impacts of interventions, which may lead to disagreement about its success or the effectiveness of the underlying approach. Despite that there are biophysical biodiversity trends, it is often true that no single counterfactual is definitively the 'right one' for conservation assessment, so multiple evaluations of intervention efficacy could be considered justifiable. Therefore, we propose the need to calculate a quantity termed the sum of perceived differences, which captures the range of impact estimates associated with different actors within a given SES. The sum of perceived differences gives some indication how closely actors within an SES agree on the impacts of an intervention. We illustrate the concept of perceived differences using a set of global, national and regional case studies. We discuss options for minimising the sum, drawing upon literatures from conservation science, psychology, behavioural economics, management and finance.

Reconciling multiple counterfactuals when evaluating biodiversity conservation impact in social-ecological systems

Abstract

When evaluating the impact of a biodiversity conservation intervention, a 'counterfactual' is needed, as true experimental controls are typically unavailable. Counterfactuals are possible alternative system trajectories in the absence of an intervention and comparing observed outcomes against the chosen counterfactual allows the impact (change attributable to the intervention) to be determined. Since counterfactuals are hypothetical scenarios, and by definition never occur, they must be estimated. Sometimes there may be many plausible counterfactuals, given that they can include multiple drivers of biodiversity change, and be defined on a range of spatial or temporal scales. Here we posit that, by definition, conservation interventions always take place in social-ecological systems (SES; ecological systems integrated with human actors). Evaluating the impact of an intervention within an SES therefore means taking into account the counterfactuals assumed by different human actors. Use of different counterfactuals by different actors will give rise to perceived differences in the impacts of interventions, which may lead to disagreement about its success or the effectiveness of the underlying approach. Despite that there are biophysical biodiversity trends, it is often true that no single counterfactual is definitively the 'right one' for conservation assessment, so multiple evaluations of intervention efficacy could be considered justifiable. Therefore, we propose the need to calculate a quantity termed the *sum of perceived differences*, which captures the range of impact estimates associated with different actors within a given SES. The *sum of perceived differences* gives some indication how closely actors within an SES agree on the impacts of an intervention. We illustrate the concept of perceived differences using a set of global, national and regional case studies. We discuss options for minimising the sum, drawing upon literatures from conservation science, psychology, behavioural economics, management and finance.

Conservation Biology

Introduction

The effectiveness of attempts to conserve the biodiversity of ecosystems – and associated contributions towards human wellbeing – has become an increasingly pressing topic over recent decades. A 'counterfactual' is necessary when quantitatively evaluating the ecological impact of a biodiversity conservation intervention (Ferraro & Hanauer, 2014). Counterfactuals are a type of 'reference scenario', capturing an alternative possible trajectory of a dynamic system in the absence of a given intervention. The 'impact' of the intervention is the change attributable to the intervention, measured as the difference between the actual observed trajectory (the 'outcome') and the predicted counterfactual (Ferraro & Pattanayak, 2006). Numerous counterfactuals can be reasonably specified for most systems relevant to conservation, as it is possible to select from a range of drivers of system change for potential inclusion within the counterfactual (Maron et al., 2018), notwithstanding that counterfactuals can also be specified over various spatiotemporal scales (Bull et al., 2014). Crucially counterfactuals are, by definition, a scenario that does not occur, so they can never be directly observed and monitored – and there is often no single 'correct' counterfactual, but rather various counterfactuals of differing plausibility. This is even true to an extent for 'control' sites used in quasi-experimental methods, as subjective decisions have to be made when choosing such sites to reduce the impacts of confounding factors (e.g. Wiik et al., 2019). Thus, though such control sites may give a good approximation to what would have happened at the 'treatment' sites without the intervention, they are still open to some interpretation.

Typically, conservationists are interested in a counterfactual representing the alternative ecological trajectory of a system, which is often influenced by multiple anthropogenic activities beyond the intervention in question (Ferraro & Pattanayak, 2006; Maron et al., 2018). But – while the literature on impact evaluation is accumulating rapidly, along with tools for implementation – impact evaluation is still rarely carried out in practice (Wiik et al., 2019). Moreover, conservation interventions do not take place in purely ecological systems – they take place in social-ecological systems (SES; Berkes & Folke, 1998), i.e. ecological systems integrated with social systems consisting of human actors. Thus, it is critical to go beyond purely *ecological* counterfactuals when evaluating the impact of a conservation intervention, and consider interlinked social systems (Alagona et al., 2012; Maron et al., 2018). This adds considerable challenges, as different actors may subjectively assume different

2

57   counterfactuals are most relevant when judging impact (e.g. whether to use a counterfactual at the

58   spatial scale of the project or the landscape; Bull et al., 2014), due to factors such as unconscious

59   biases (Tversky & Kahneman, 1974), temporal starting point (Pauly, 1995), or assumptions made

60   about processes driving change. This is important as the choice of counterfactual not only alters

61   perceptions about intervention success, but also potentially the actions of stakeholders (Bull et al.,

62   2014). Therefore, it is insufficient to consider counterfactuals on a purely ecological basis when

63   judging conservation impact; also deserving attention are the ecological counterfactuals associated

64   with differing interpretations from relevant stakeholders within the SES. We term this set of possible

65   social-ecological counterfactuals – i.e. ecological counterfactuals derived from varying stakeholder

66   perceptions – a *family* of counterfactuals.

67

68   Here, our objective is to further formalise the use of counterfactuals for evaluating biodiversity

69   outcomes within an SES. Biodiversity is not the only property of an SES which might require

70   conservation interventions, but it is our focus. We develop an exploratory conceptual framework,

71   illustrated (although not formally tested) with case studies. In particular, we focus on differences of

72   interpretation during quantitative evaluation of the ecological impact of conservation interventions.

73   Approaches for qualitative evaluation exist (Sutherland et al., 2018); however, as the latter do not

74   require quantitative impact estimates, we do not explore them further here. We do, though, explore

75   how an actor's precise choice of counterfactual arises from their personal 'reference frame'. We go on

76   to explore what approaches exist for minimising divergence in personal reference frames and thus the

77   choice of counterfactual, to avoid conflicting perceptions of conservation impact in SES. Ultimately,

78   we aim to enable better impact evaluation by considering not only biophysical outcomes against

79   counterfactuals, but also how multiple stakeholders view the plausibility of different counterfactuals.

80

81   **Towards families of counterfactuals**

82   We start with a conservation intervention in a hypothetical ecosystem. We do consider 'conservation

83   interventions' to be targeted at biodiversity, but framed more broadly around 'people and nature'

84   (emphasizing "the importance of cultural structures and institutions for developing sustainable and

85   resilient interactions between human societies and the natural environment"; Mace, 2014). By

86   definition, an intervention being implemented means that: (a) at least one human actor has the

3

87   potential to influence that biodiversity; (b) at least one human actor must be affected by that influence

88   on biodiversity; and, (c) at least one human actor must be responsible for the intervention. In the

89   simplest case, (a), (b) and (c) describe the same actor; for instance, if the SES is a farm, the single

90   actor might be a farmer creating habitat for declining bird species. Note the logic that, for any system

91   within which a conservation intervention is taking place, there must be at least one relevant human

92   actor – that is, *every* system involving a conservation intervention must be an SES.

93

94   We focus on social perceptions of the ecological impact of an intervention, a necessary precursor to

95   evaluating its social *impacts*; but though social impacts require similar treatment (e.g. Davidson,

96   2013), they are beyond scope here. So, to evaluate the impact of the intervention we then specify a

97   counterfactual ("a causal effect of a program is only defined with respect to a well-defined alternative";

98   Ferraro & Hanauer, 2014). The actor anticipates an alternative biodiversity trend that would have

99   taken place without the intervention. But biodiversity measurement is open to interpretation (Purvis &

100   Hector, 2000), and sub-components of biodiversity are often ascribed wildly different weightings

101   dependent upon the actor (Baylis et al., 2016; Bull & Maron, 2016; Pearson, 2016). Compounding this

102   is the difficulty in determining which components of biodiversity are relevant when going beyond static

103   considerations – which depends upon the chosen spatiotemporal scale (Bull et al., 2014). Finally, the

104   'shape' of the counterfactual trend will be heavily influenced by actor's expectations (Ferraro &

105   Hanauer, 2014). Yet in the simplest case of our hypothetical SES, the single actor is free to select

106   whatever counterfactual they choose – and the perceived impact of the intervention is the difference

107   between the observed biophysical outcome and that counterfactual (Ferraro & Hanauer, 2014).

108   Continuing our farmer example, the farmer might compare outcomes to the counterfactual scenario in

109   which they had not created new habitat (acknowledging multiple plausible futures).

110

111   Consider a more complex hypothetical scenario with *two* actors: one carrying out activities that reduce

112   biodiversity, and the other implementing conservation interventions. Assuming it is relevant for both

113   actors to monitor impact, each specifies a counterfactual and measures the outcome. There is a

114   wealth of reasons why their choice of counterfactual might differ – but even if their choice is the same

115   in theory, expectation or uncertainty might mean that the precise trajectory of their chosen

116   counterfactual diverges. The perceived impact for Actor 1 is the difference between the observed

117   biophysical outcome and Actor 1's chosen counterfactual, and for Actor 2 it is the difference between

118   the same biophysical outcome and Actor 2's chosen counterfactual. If the counterfactuals are

119   different, there is consequently a difference between perceived impact for Actor 1 and Actor 2, which

120   – whatever the actual trajectory – equals the difference in their chosen counterfactuals (assuming

121   they use the same measure of outcome; Fig. 1, Appendix 1). For a real example see Maron et al.

122   (2015), who find Australian policymakers (Actor 1) using different counterfactual rates to those

123   calculated by researchers (Actor 2) for 'biodiversity offset' conservation interventions (Fig. 1, and

124   below); or analogously, multiple reference levels in climate change mitigation (Griscom et al., 2009).

125

126   Finally, consider the general case of three or more actors within the hypothetical SES, who all

127   evaluate the impact of an intervention, but at least some specify different counterfactuals. This

128   resultant set of counterfactuals is the aforementioned *family of counterfactuals*. For the SES as a

129   whole, consider the total difference in perceived impact across the family of counterfactuals for the full

130   diversity of stakeholders: the larger it is, the greater the difference in perceived impact for multiple

131   actors, which may imply that some actors are not satisfied with the intervention. We propose

132   calculating a *sum of perceived differences* for the family of counterfactuals, which is the sum of the

133   magnitude of the difference between the counterfactuals assumed by every actor and every other

134   actor within the system (see Appendix 1 for our mathematical formulation, and justification). Note: the

135   sum could conceivably weight the counterfactuals assumed by various individuals differently, e.g. to

136   account for different uncertainties, or for uneven power dynamics – we return to this later in the

137   article. Since no counterfactual is definitively 'correct', but rather is chosen on the basis of actors'

138   value judgments, the sum of perceived differences is necessary to capture the impact of a

139   conservation intervention in an SES. Importantly, the sum incorporates the counterfactuals actually

140   used by actors without differentiating between those counterfactuals that are more or less plausible;

141   but applying this framework *does* make the counterfactuals used more transparent, facilitating

142   discussion around plausibility.

143

144   Actual biodiversity trends are, in principle, objective and biophysical. Yet, because the result of

145   conservation interventions will be scrutinised by multiple different actors who each have their own

146   personal counterfactual, which will often diverge, even robustly monitored interventions could lead to

147   wide-ranging and at times conflicting interpretations of efficacy (Pearson, 2016). Equally, since robust

148   evaluation of interventions must inextricably be linked to the initial design of the interventions

149   themselves (Bull et al., 2014), the challenge of divergence in counterfactuals is important when

150   setting conservation objectives. Concerning the sustainability of interventions, it would be preferable

151   to minimise the sum of perceived differences for an SES – implying that most actors use

152   approximately the same counterfactual. Is it possible to reduce the sum of perceived differences? To

153   answer, we first consider the 'reference frame' within which the counterfactual is specified.

154

155   **Counterfactuals and reference frames**

156   Again, counterfactuals are a type of *reference scenario* (Maron et al., 2018). Reference scenarios are

157   specified within a *reference frame* (or 'frame of reference'), and any number of reference scenarios

158   can be specified within one reference frame (Bull et al., 2014). In Appendix 2 we formally describe a

159   reference frame, and in Appendix 3 discuss usage of the term in different scientific disciplines. To a

160   greater or lesser extent across different disciplines, a reference frame is typically composed of: (i) the

161   relevant 'parameter space' capturing all possible variables of interest, particularly the subset of

162   interest to the relevant 'observer'; and, (ii) the observer themselves, including the social values

163   through which that observer views the chosen parameters. Unlike other disciplines in the natural

164   sciences, which do so *alongside* rather than as *part* of the reference frame, we also include: (iii) the

165   specific *coordinate system* used to make measurements within the parameter space, chosen by the

166   observer, which determines the spatiotemporal scale of the reference frame (see Fig. A2.1). A

167   coordinate system is a set of numbers used to specify quantitative information (e.g. an object's

168   location in space; Appendix 3). In contrast to measurement of physical change in disciplines where

169   standardised units exist, a coordinate system requires specification when assessing quantitative

170   change in biodiversity as it is an imprecise term, and so there are many different ways in which

171   'biodiversity' could be measured (Purvis & Hector, 2000).

172

173   A counterfactual (or indeed any change) cannot be comprehensively specified in the absence of a

174   reference frame: the observer, the parameters they are interested in, and the scale and coordinate

175   system they are using must all be defined first (Fig. A2.1) even if in practice these are typically implicit

176   or assumed (Bull et al., 2014). The reason for discussing reference frames here is that it is differences

177   in the actors' underlying reference frames that give rise to the family of counterfactuals. To expand,

178   reference frames used for evaluating conservation interventions vary depending upon the actor

179   carrying out the evaluation – because different actors have different personal values built into their

180   reference frame when constructing a counterfactual (e.g. Bull & Maron, 2016) which additionally may

181   change over time, or because actors make different assumptions about the appropriate parameter

182   space and scale for evaluation (e.g. Pauly, 1995). Minimising the sum of perceived differences in an

183   SES therefore requires influencing underlying reference frames.

184

185   Reference frames can be *divergent* (e.g. actors using different scales) and even *conflicting* (e.g.

186   observers making directly contradictory assumptions about value ascribed to biodiversity

187   components) (Hahn et al., 2014). The idea of conflicting reference frames is not new – Schön & Rein

188   (1994) discuss divergent and conflicting reference frames in policy design respectively as

189   'disagreements' (resolvable by examining facts) versus 'controversies' (which tend to be intractable).

190   In conservation, it is well established that focusing upon different parameters or coordinate systems

191   can cause conflicting assessments on efficacy (e.g. Naidoo et al., 2008).

192

193   **Illustration**

194   We now illustrate the preceding concepts using examples across different spatial scales. Multi-scalar

195   illustration is important as counterfactual evaluation is relevant and necessary for conservation

196   interventions on any spatial scale (Ferraro & Pattanayak, 2006). Note that in testing our proposals

197   more thoroughly – and certainly in implementing them – the counterfactuals used would be based on

198   extensive empirical data collected from different actors.

199

200   *International: influence of Aichi Target 11*

201   We start at the largest scale, treating the global biosphere as our SES. Aichi Target 11 is associated

202   with the Convention on Biological Diversity (CBD) (https://www.cbd.int/sp/targets/): these targets, set

203   in 2010, were intended to be met by 2020, and there are numerous interested stakeholders seeking to

204   understand the impact of setting them (see Butchart et al., 2010). Aichi Target 11 states that, "By

205   2020, at least 17 per cent of terrestrial and inland water, and 10 per cent of coastal and marine areas,

206   especially areas of particular importance for biodiversity and ecosystem services, are conserved

207   through effectively and equitably managed, ecologically representative and well connected systems of

208   protected areas and other effective area-based conservation measures, and integrated into the wider

209   landscapes and seascapes". Here, for conceptual clarity, we focus on a subset of protected areas

210   (marine areas), tracked via the World Database on Protected Areas (WDPA).

211

212   Observations seemingly indicate progress towards Target 11, evidenced by the ongoing trend in

213   marine protected area (MPA) coverage (Fig. 2). But the process of setting the Aichi Targets can be

214   viewed as an attempt to stimulate specific actions across the 196 countries Party to the CBD; as

215   such, it is reasonable to ask whether there has been any associated change in conservation trends

216   since 2010. So, imagine one actor characterises the immediately pre-2010 trend in MPA coverage as

217   linear, with the post-2010 projection of that linear trend taken as their counterfactual. Comparing the

218   latter against the observed trend for MPA coverage post-2010, this first actor treats the difference

219   between the two as resulting from action stimulated by setting Target 11 (Fig. 2). Though not a

220   particularly sound statistical analysis, it represents a plausible interpretation of the data.

221

222   A second actor might instead argue that a non-linear trend-line fits the 18-year dataset better,

223   suggesting the MPA network is on a longer-term exponential growth trajectory. This would imply that

224   the counterfactual scenario in the absence of the CBD Strategic Plan was also the observed outcome,

225   insinuating (correctly or otherwise) that setting Target 11 had no influence on net MPA outcomes. The

226   perceived difference between actors one and two, regarding the degree to which Target 11 has

227   stimulated additional growth in MPA coverage, is clear – despite their using the same biophysical

228   outcome, metric and dataset (Fig. 2).

229

230   *Regional: Australian state biodiversity offset policy*

231   A more pertinent application is in exploring the actions taken towards meeting global conservation

232   policy targets at regional scales; not least because it is on such scales that conservation interventions

233   typically act, and which ecosystem responses can be monitored over reasonable timescales. So, we

234   turn to the aforementioned example of Australian state-level 'biodiversity offset' policies. These

235   policies require that biodiversity losses via clearance of certain native habitats, as a result of

236   economic development activities, are fully compensated through biodiversity gains on sites with

237   comparable habitat elsewhere. These gains can take the form of 'averted losses' (i.e. protection of a

238   habitat that prevents otherwise near-certain degradation), and as such the counterfactual scenario

239   used to evaluate impacts is critical (Maron et al., 2015).

240

241   For six offset policies across five Australian states, Maron et al. compared the counterfactual rate of

242   habitat loss assumed by the regulator (the 'crediting baseline' in Fig. 1) against a counterfactual

243   calculated by researchers based on the proxy of 'recent observed deforestation'. These can be

244   considered the counterfactuals assumed by Actors A1 and A2 respectively, and clearly the difference

245   is substantial for most states. Using our mathematical formulation (Appendix 1), we can calculate the

246   sum of perceived differences between the two actors for each state separately (Table 1). In doing so,

247   we show that even the normalised sum varies by 1–2 orders of magnitude between states, and South

248   Australia is a potential outlier (although this may be an artefact of offset data availability for the state;

249   Maron et al., 2015). In isolation, these figures do not substantially advance our understanding of the

250   SES, but if compared to similar statistics calculated for a wide range of regional policy interventions

251   elsewhere (an empirical application of our framework), they could indicate the relative degree of

252   disagreement over policy outcomes.

253

254   *Local: species protection in Uzbekistan*

255   The previous examples consider two actors, but the framework can be extended to any number, and

256   for different biodiversity components. Consider protection efforts on a subnational scale for an IUCN

257   Red List species, the Critically Endangered saiga antelope *Saiga tatarica*. In remote northwest

258   Uzbekistan, the 'Ustyurt' saiga population has declined in recent years, and consequently attracted

259   intensive conservation efforts; but the region is also experiencing increasing economic development

260   activity (Bull et al., 2013). New development projects typically seek to mitigate any impacts that might

261   exacerbate extinction risk for saigas, and so a counterfactual scenario is needed for evaluating the

262   success of mitigation measures relative to the observed trend in saiga population numbers.

263

264   We compare five possible counterfactual population trends for Uzbek saigas over 15 years against

265   the observed population trend. Three counterfactuals relate to known population trends for distinct

266   saiga populations of historically comparable sizes (Association for the Conservation of Biodiversity of

267   Kazakhstan, unpublished data), and two are hypothetical but realistic counterfactuals (extirpation, and

268   population expansion without die-offs) (Fig. 3). From initially close alignment, the counterfactuals

269   (which we ascribe to five different actors, A1 – A5) diverge substantially over time. This can be

270   tracked using the annual normalised sum of perceived differences across the family of counterfactuals

271   (Fig. 3); showing the dramatic variation possible in terms of interpreting the impact of protection

272   efforts.

273

274   If divergent families of counterfactuals in SES do undermine interventions, it is a problem which needs

275   resolution. One solution would be to make the divergence explicit, requiring actors to 'agree to

276   disagree' (Biggs et al., 2017). But next, alternatively, we consider opportunities to minimize the

277   summed difference across a family of counterfactuals via resolving conflicting frames.

278

279   **Minimising perceived differences**

280   A starting point is to *understand* the relevant actors' reference frames. Extensive literature exists on

281   stakeholder analysis in relationship to natural resource management (Reed et al., 2009; Cummings et

282   al., 2018). Determining reference frames through stakeholder analysis would involve uncovering the

283   perceptions of different actors, associated social discourses, and relationships between actors (Reed

284   et al., 2009; Baynham-Heard et al., 2018). One aspect that needs consideration when implementing

285   our approach is the *power* dynamics between actors (i.e. differentiation between actors' ability or

286   capacity to influence outcomes). This involves assessing who is represented and who is left out of the

287   decision-making, and whether the reference frames of some actors be given greater weighting (e.g.

288   Smith et al., 2010). This is not the same as *weighting* for power dynamics when performing our

289   proposed calculations: that is something we avoid, since our framework relates specifically to

290   differences in counterfactuals *chosen* by various actors, rather than actors' ability to *act* upon those

291   differences. Power dynamics are certainly important to our approach if they result in certain

292   stakeholders being completely excluded from consideration, such that their perspective is not

293   incorporated into the sum of perceived differences; a risk that requires careful treatment. Conversely,

294   consideration of the specific nuance of power dynamics – albeit likely an important next step for

295   research on this topic – is more crucial to discussions around steps taken to *resolve* such differences.

296   Though we begin to explore the latter in this article, the issue of power dynamics deserves further

297   attention in its own right.

298

299   Different authors have explored options for using understanding of actors' perspectives to design

300   effective conservation (Battista et al., 2018; Cummings et al., 2018; Cinner, 2018); here we do so

301   based upon our formalised structure for reference frames (Appendix 2). So – having isolated different

302   actors' reference frames, we can structure possible approaches to conflict resolution in terms of the

303   three key components (Fig. A2.1; Table 2), each of which we discuss below.

304

305   <u>Conflict in the physical component</u>

306   The physical component of the reference frame (i.e. measurable biophysical quantities, whether biotic

307   or abiotic; Appendix 2) is inevitably associated with development of conservation objectives.

308   Consequently, resolution of conflict in the physical component of the reference frame relates strongly

309   to conservation objectives and targets.

310

311   Structured decision-making (SDM; an approach designed to "systematically incorporate participant

312   values, objectives and knowledge in decision-making"; Addison et al., 2013) can help develop

313   objectives for natural resource use, given competing actor values (e.g. Robinson et al., 2016).

314   Further, despite involving explicit expression of conflicting reference frames, SDM can strengthen

315   consensus amongst groups during decision-making (Priem et al., 1995) such as design and

316   evaluation of interventions. Indeed, experts typically perform better at making decisions based on

317   well-structured discussions with peers rather than alone (Burgman et al., 2011). Incorporating

318   monitoring and evaluation into SDM is well established (Lyons et al., 2008). Considering these

319   factors, SDM provides a practical platform for using group consensus to reduce conflicts between

320   stakeholders with different reference frames.

321

322   When there are intractable differences between stakeholders, consensus may be impossible.

323   International policy decisions can be seen as negotiated decisions constructed through multi-actor

324   interactions (Daniels et al., 2012). In such cases, e.g. the development of quantifiable environmental

325   targets, Maxwell et al. (2015) propose allowing room for manoeuvre, promoting ambiguity in targets

11

326   with the goal of building trust, cooperation and consensus. This is akin to allowing ambiguity in

327   specification of the parameter space for the physical component of conflicting frames, to promote

328   convergence in social components of those reference frames. An obvious disadvantage is that

329   ambiguity in the definition of any component of the reference frame makes rigorous evaluation

330   impossible – so the approach sacrifices short-term measurability to enable long-term resolution.

331

332   <u>Conflict in the social component</u>

333   More diverse groups often frame likely outcomes more accurately (Sutherland & Burgman, 2015). But

334   this is not only relevant to experts and decision-makers: Addison et al. (2013) suggest using

335   participatory SDM to increase broader social acceptance of conservation objectives. That is, through

336   participatory exercises featuring diverse non-experts, interventions can be designed that integrate

337   multiple social reference frames. Such reference frames typically involve several layers of complexity:

338   deeply held values, important interests, local knowledge, and socially embedded conflicts among

339   them (Daniels et al., 2012). Similarly, participatory modelling is designed to integrate a diversity of

340   perspectives (i.e. social reference frames) through e.g. using role playing games (Jones et al., 2007)

341   – again, for both specialist and non-specialist participants. In seeking consensus across

342   heterogeneous groups of actors, there are many formal processes available e.g. stakeholder

343   engagement, scenario planning (Cummings et al., 2018), and workshops designed to facilitate

344   alignment between individual social values (Kenter et al., 2015).

345

346   Beyond participatory mechanisms, extensive literature concerns the degree to which personal

347   reference frames can be modified, including for conservation (Iftekahr & Pannell, 2015). Consider

348   'anchoring' – when estimating numerical outcomes, individuals generally make minor adjustments to

349   some initial value to yield their answer. The initial (anchoring) value is linked to the individual's

350   personal cognitive frame, and influenced by the formulation of the question itself (Tversky &

351   Kahneman, 1974). Consequently, the extensive theoretical basis underlying the construction of

352   personal reference frames could be leveraged to support resolution of conflicting reference frames –

353   for instance, by purposefully creating common, appropriate anchoring points for evaluating

354   conservation interventions by a range of actors (e.g. the quantitative extent of remaining habitat

355   evaluated as 'acceptable'; see Cinner, 2018).

356

357 The power of setting expectations for assessment is also relevant in terms of qualitative presentation

358 of interventions; policies are likely deemed more acceptable when presented without loaded terms.

359 Again, the presentation of interventions can thus be used proactively to help set personal reference

360 frames amongst actors and reduce potential conflicts; albeit with associated ethical considerations

361 (e.g. Rothschild, 2000). Alternatively, personal reference frames might be open to modification

362 through management of appropriate incentives. Whilst, prosaically, these could be the provision of

363 financial incentives, they could also be 'intrinsic' incentives related to personal desires or values.

364 Some stakeholders might be better motivated by e.g. attachment to the land than by financial

365 incentives (Reddy et al., 2017).

366

367 Of course, substantial components of personal reference frames are deeply held and cannot be

368 modified. Individuals distinguish between 'sacred' and 'secular' values (Tetlock et al., 2000). This

369 forms part of an actor's personal reference frame, and they may consider it unacceptable for a

370 conservation intervention to exchange one type of value for the other (Daw et al., 2015). In that case,

371 the solutions would be to explicitly identify sacred values held by different actors and ensure that they

372 are not jeopardised by interventions. Such an approach – requiring an iterative process through which

373 trust and cooperation is built between key parties on a seemingly intractable issue – involves

374 negotiating shared belief structures (Cummings et al., 2018), e.g. squaring different 'mental models'

375 held by stakeholders (Biggs et al., 2017). Alternatively, the intervention could be redesigned around

376 secular values, for which examples in the literature go back decades (Cummings et al., 2018).

377

378 <u>Conflict in coordinate systems</u>

379 One solution to conflict caused by using different coordinate systems is accepting the need to track

380 multiple indicators (e.g. Butchart et al., 2010). Approaches that combine multiple indicators are used

381 widely elsewhere (e.g. finance; Engle & Gallo, 2006). Where this is not satisfactory – say, due to the

382 increased resource requirements – an alternative is again participatory approaches, facilitating

383 consensus between actors on coordinate systems rather than parameter spaces.

384

385   Conflict arising from differences in the spatial or temporal scale for evaluation might arise because

386   scales are implicitly assumed by different actors (Bull et al., 2014). An explicit statement of the scale,

387   as part of counterfactual construction, is therefore one straightforward approach towards avoiding

388   conflict. A more nuanced approach would be to consider the appropriate counterfactual scenario for

389   different spatial scales at a given moment: Maron et al. (2018) recommend that the scope of the

390   reference frame be explicitly expanded to approach the largest possible 'overarching' frame. Doing so

391   is comparable with other calls in the literature to evaluate conservation interventions using very long-

392   term timescales (e.g. Willis & Birks, 2006). Finally, Pearson (2016) suggests actors construct their

393   personal reference frames for conservation based partly upon the spatial scale in question: setting the

394   scale could itself be a means for influencing personal reference frames.

395

396   **Conclusions**

397   All biodiversity conservation interventions take place within a social-ecological system (SES), and if a

398   SES contains more than one actor, this may give rise to a family of counterfactuals. The use of a

399   family of different counterfactuals by different actors will cause perceived differences in the impacts of

400   (change attributable to) an intervention, even when all actors agree on outcomes. This may lead to

401   disagreement between actors on the efficacy of interventions, undermining efforts to conserve

402   biodiversity. But, since no counterfactual can be considered definitive, multiple different impact

403   evaluations can be performed or proposed by different actors, more than one of which might be

404   considered valid. Therefore, we have developed the basis for evaluating a *sum of perceived*

405   *differences* between actors (defined mathematically in Appendix 1), and explored approaches for

406   minimizing perceived differences. An important next step would be to test the conceptual framework

407   we have developed here against extensive empirical data: not only in terms of the value of our

408   proposed sum in different SES and associated implications, but also in terms of finding the limits to

409   which counterfactuals might be considered 'valid'. Our work provides an exploratory theoretical basis

410   for better quantifying, understanding and ultimately managing multiple diverse perspectives on nature

411   conservation in SES.

412

413   **References**

414 Addison PFE, et al. (2013). Practical solutions for making models indispensable in conservation

415   decision-making. *Diversity and Distributions*, 19, 490–502.

416 Battista W, et al. (2018). Behaviour change interventions to reduce illegal fishing. *Frontiers in Marine*

417   *Science*, 5(403).

418 Baylis K, et al. (2016). Mainstreaming impact evaluation in nature conservation. *Conservation Letters*,

419   9(1), 58–64.

420 Baynham-Heard Z, Redpath S, Bunnefeld N, Molony T, & Keane A (2018). Conservation conflicts:

421   Behavioural threats, frames, and intervention recommendations. *Biological Conservation*, 222,

422   180-188.

423 Berkes F, & Folke C (1998), Linking social and ecological systems: management practices and social

424   mechanisms for building resilience. Cambridge University Press; Cambridge, UK.

425 Biggs D, et al. (2017). Breaking the deadlock on ivory. *Science*, 358(6369), 1378–1381.

426 Bull JW, Gordon A, Law E, Suttle KB, & Milner-Gulland EJ (2014). The importance of baseline

427   specification in evaluating conservation interventions and achieving no net loss of biodiversity.

428   *Conservation Biology*, 28(3), 799–809.

429 Bull JW, Lloyd SP, & Strange N (2017). Implementation gap between the theory and practice of

430   biodiversity offset multipliers. *Conservation Letters*, 10(6), 656–669.

431 Bull JW, & Maron M (2016). How humans drive speciation as well as extinction. *Proceedings of the*

432   *Royal Society B: Biological Sciences*, 283(1833).

433 Bull JW, Suttle KB, Singh NJ, & Milner-Gulland EJ (2013) Conservation when nothing stands still:

434   moving targets and biodiversity offsets. *Frontiers in Ecology and the Environment*, 11(4), 203-

435   210.

436 Burgman MA, et al. (2011). Expert status and performance. *PLoS 1*, 6(7), e22998.

437 Butchart SHM, et al. (2010). Global biodiversity: indicators of recent declines. Science, 328(5982),

438   1164–1168.

439 Cinner J (2018). How behavioral science can help conservation. *Science*, 362(6417), 889-890.

440 Corlett RT (2016). Restoration, reintroduction, and rewilding in a changing world. *Trends in Ecology*

441   *and Evolution*, 31(6), 453–462.

442 Cummings JW, et al. (2018). Implicit decision framing as an unrecognized source of confusion in

443   endangered species classification. Conservation Biology, DOI: 10.1111/cobi.13185.

444   Daniels SE, Walker GB, & Emborg J (2012). The unifying negotiation framework: A model of policy

445   discourse. *Conflict Resolution Quarterly*, 30, 3–31.

446   Davidson MD (2013). On the relation between ecosystem services, intrinsic value, existence value

447   and economic valuation *Ecological Economics*, 95, 171-177.

448   Daw TM, et al. (2015). Evaluating taboo trade-offs in ecosystems services and human well-being.

449   *PNAS*, 112(22), 6949–6954.

450   Engle RF, & Gallo GM (2006). A multiple indicators model for volatility using intra-daily data. *Journal

451   of Econometrics*, 131, 3–27.

452   Ferraro PJ, & Pattanayak SK (2006). Money for nothing? A call for empirical evaluation of biodiversity

453   conservation investments. *PLoS Biology*, 4(4), e105.

454   Griscom B, Shoch D, Stanley B, Cortez R, & Virgilio N (2009). Sensitivity of amounts and distribution

455   of tropical forest carbon credits depending on baseline rules. *Environmental Science & Policy*,

456   12(7), 897-911.

457   Groves CR, & Game ET (2015). Conservation planning: informed decisions for a healthier planet.

458   Roberts and Company Publishers Inc.; Colorado, USA.

459   Hahn T, Preuss L, Pinkse J, & Figge F (2014). Cognitive frames in corporate sustainability:

460   managerial sensemaking with paradoxical and business case frames. *Academy of

461   Management Review*, 39(4), 463–487.

462   Iftekhar MS, & Pannell DJ (2015). 'Biases' in adaptive natural resource management. *Conservation

463   Letters*, 8, 388–396.

464   Jones NA, et al. (2007). Evaluating participatory modelling: developing a framework for cross-case

465   analysis. CSIRO Working Paper 2008-11; Canberra, Australia.

466   Kenter JO, et al. (2015). What are shared and social values of ecosystems? *Ecological Economics*,

467   111, 86–99.

468   Lyons JE, Runge MC, Laskowski HP, & Kendall WL (2008). Monitoring in the context of Structured

469   Decision-Making and adaptive management. Journal of Wildlife Management, 72(8), 1683–

470   1692.

471   Mace, GM (2014). Whose conservation? *Science*, 345(6204), 1558-1560.

472   Maron M, Bull JW, Evans MC, & Gordon A (2015). Locking in loss: baselines of decline in Australian

473   biodiversity offset policies. *Biological Conservation*, 192, 504–512.

474   Maron M, et al. (2018). The many meanings of No Net Loss in environmental policy. *Nature*

475   *Sustainability*, 1, 19–27.

476   Marris E (2013). Charismatic mammals can help guide conservation. *Nature*,

477   doi:10.1038/nature.2013.14396.

478   Maxwell SL, et al. (2015). Being smart about SMART environmental targets. *Science*, 347(6226),

479   1075–1076.

480   Naidoo R, et al. (2008). Global mapping of ecosystem services and conservation priorities. *PNAS*,

481   105(28), 9495–9500.

482   Pauly D (1995). Anecdotes and the shifting baseline syndrome of fisheries. *Trends in Ecology and*

483   *Evolution*, 10(10), 430.

484   Pearson RG (2016). Reasons to conserve nature. *Trends in Ecology and Evolution,* 31(5), 366–371.

485   Poiani KA, Goldman RL, Hobson J, Hoekstra JM, & Nelson KS (2011). Redesigning biodiversity

486   conservation projects for climate change: examples from the field. *Biodiversity and*

487   *Conservation*, 20(1), 185–201.

488   Priem RL, Harrison DA, & Muir NK (1995). Structured conflict and consensus outcomes in group

489   decision making. *Journal of Management*, 21(4), doi:10.1177/014920639502100406.

490   Purvis A, & Hector A (2000). Getting the measure of biodiversity. *Nature*, 405(6783), 212–219.

491   Reddy SMW, et al. (2016). Advancing conservation by understanding and influencing human

492   behaviour. *Conservation Letters*, 10(2), 248–256.

493   Reed MS, et al. (2009). Who's in and why? A typology of stakeholder analysis methods for natural

494   resource management. *Journal of Environmental Management*, 90, 1933–1949.

495   Robinson KF, et al. (2016). Structured decision making as a framework for large-scale wildlife harvest

496   management decisions. *Ecosphere*, 7(12), e01613.

497   Rothschild ML (2000). Ethical considerations in support of the marketing of public health issues.

498   *American Journal of Health Behavior*, 24(1), 26-35.

499   Schön A, & Rein M (1994). Frame reflection: toward the resolution of intractable policy controversies.

500   Taylor & Francis, Ltd.; Oxfordshire, UK.

501   Smith RJ, Adams WM, & Leader-Williams N (2010). Another entangled bank: making conservation

502   trade-offs more explicit. In, Trade-offs in conservation: deciding what to save (Eds., Leader-

503   Williams N, Adams WM, Smith RJ). Wiley-Blackwell; UK.

504   Sutherland WJ, & Burgman MA (2015). Use experts wisely. *Nature*, 526, 317–318.

505   Sutherland WJ, Dicks LV, Everard M, & Geneletti D (2018) Qualitative methods for ecologists and

506   conservation scientists. *Methods in Ecology & Evolution*, 9(1), 7-9.

507   Tetlock PE, Kristel OV, Elson SB, Lerner JS, & Green MC (2000). The psychology of the unthinkable:

508   taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and*

509   *Social Psychology*, 78(5), 853–870.

510   Tversky A, & Kahneman D (1974). Judgments under uncertainty: heuristics and biases. *Science*,

511   185(4157), 1124–1131.

512   Wiik E, et al. (2019). Experimental evaluation of the impact of a payment for environmental services

513   program on deforestation. *Conservation Science and Practice*, 1:e8.

514   Willis KJ, & Birks HJB (2006). What is natural? The need for a long-term perspective in biodiversity

515   conservation. *Science*, 314(5803), 1261–1265.

Figure 1: hypothetical plot of different counterfactuals (*$B_{c(A1)}$*, *$B_{c(A2)}$*) used by two actors (**Actor 1**, **Actor 2**) in an SES, compared to a baseline (*B_{base}*) at time

= 0 and the observed biodiversity outcomes through time (*B_{obs}*). **Actor 1**'s perceived impact is based on comparing *B_{obs}* against *$B_{c(A1)}$*, while Actor 2

compares *B_{obs}* against *$B_{c(A2)}$*. The difference in perceived impact is shown. Inset: real life example, of two actors judging different counterfactual deforestation

rates in Australian states (from Maron et al., 2015), abbreviations for each state given in Table 1.
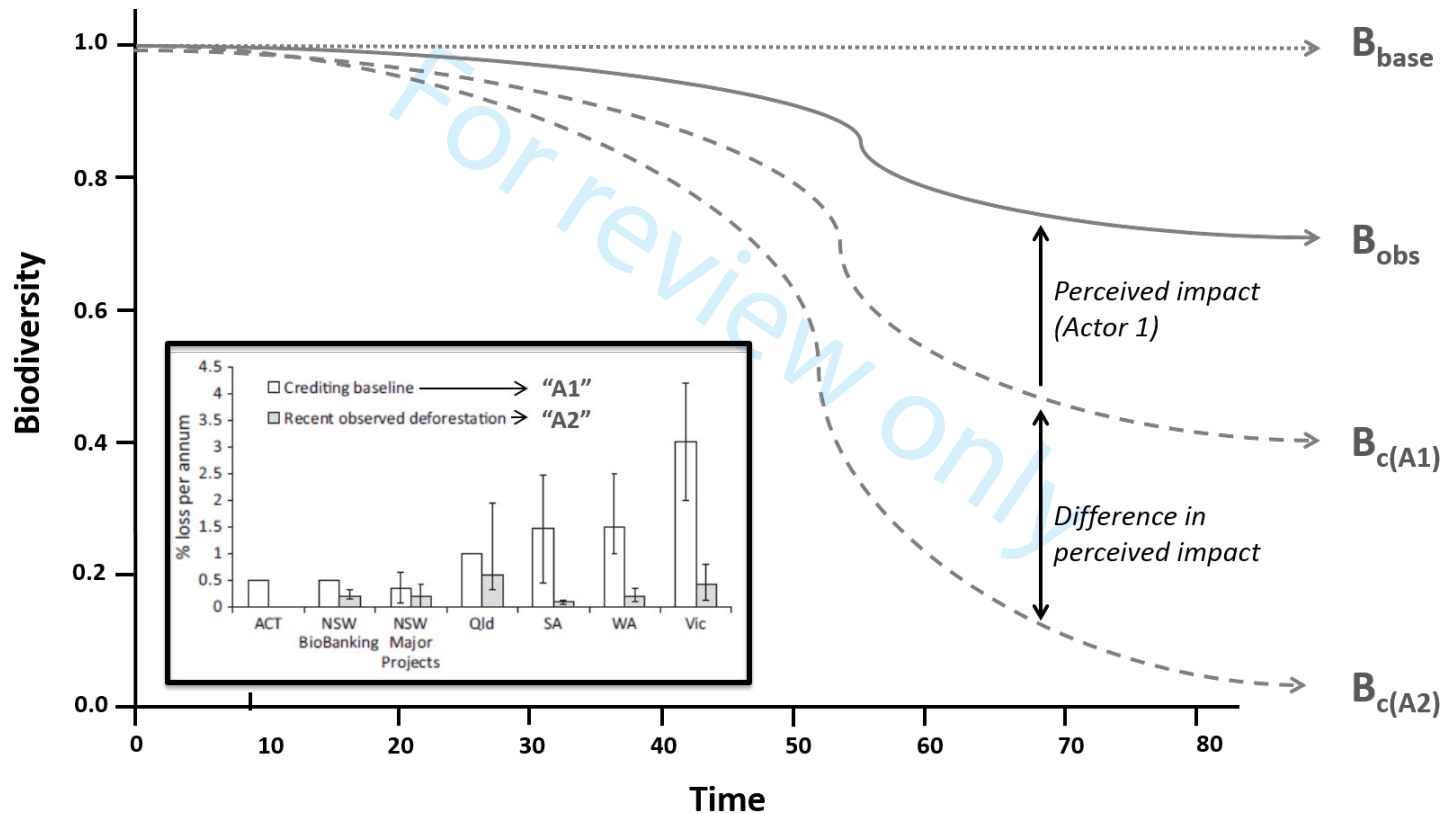
Figure 2: plot of global marine protected area coverage (%) against time (years). Blue = pre-2010 trend (counterfactual for A1); orange = post-2010 trend; black = exponential trend-line fit to entire time series (counterfactual for A2); grey = Aichi Target 11 (10%). A perceived difference in the impact of setting Aichi Target 11 arises between two actors who use different counterfactuals (the blue and black lines on the graph). Data from the WDPA (UNEP-WCMC).
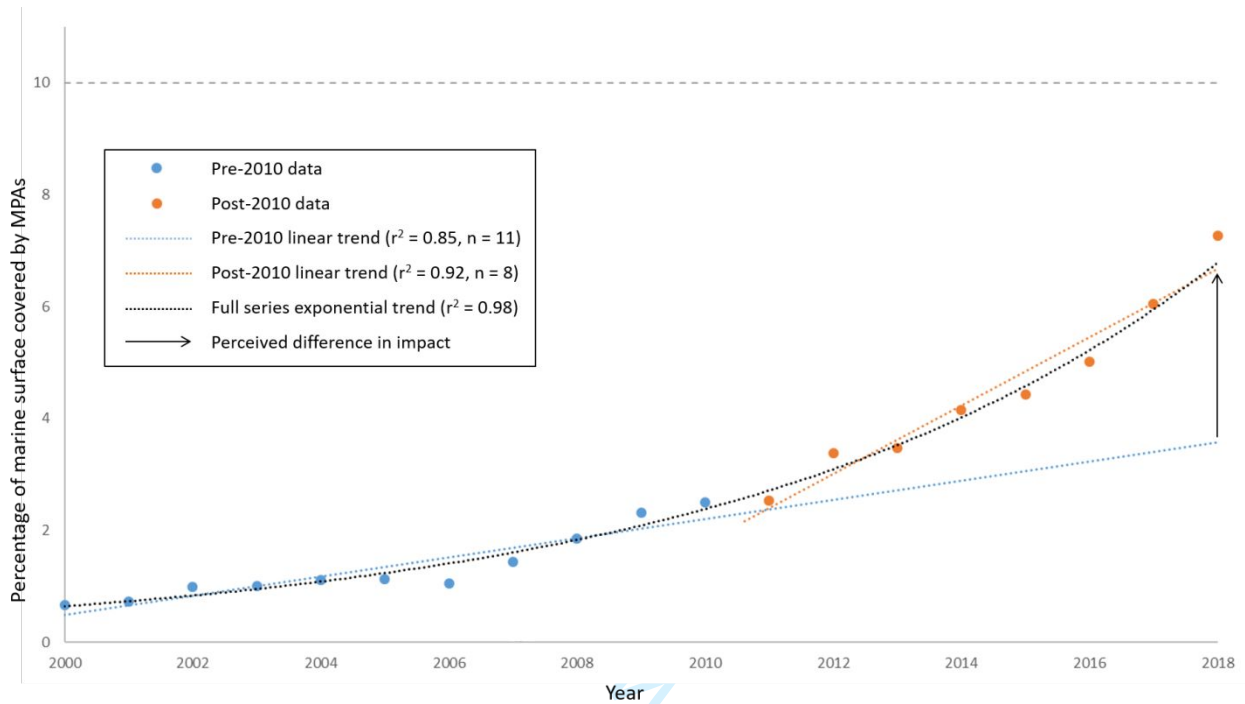
Figure 3: *evaluating the outcomes of species protection in northwest Uzbekistan, 2006 – 2018.*

Primary y axis: dashed grey lines = five possible counterfactual trends assumed by different actors

'A1' – 'A5' for the regional saiga antelope Saiga tatarica *population; solid line = observed population*

trend 'Ob' (ACBK, unpublished data). All six lines labelled, for clarity, at years 2013 and 2017.

Secondary y axis: block dots = annual normalised sum of perceived differences.
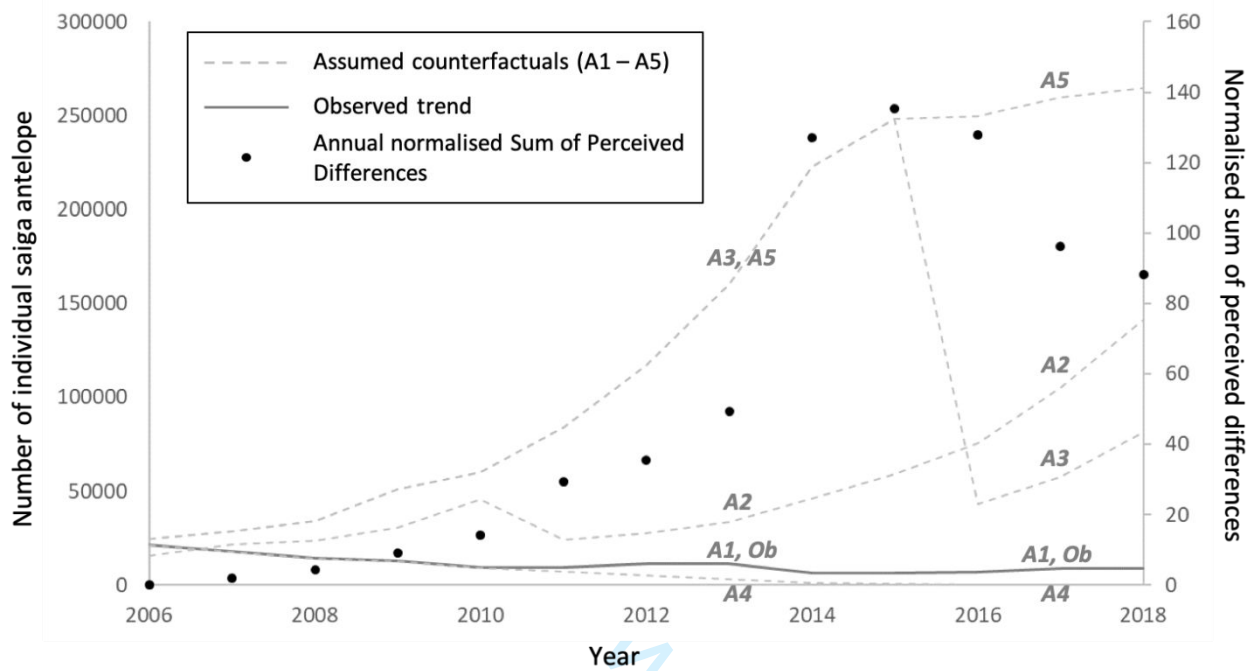
Table 1: table of the perceived difference in outcomes between the regulator and researchers (A1 and A2, respectively) in the case of six different Australian state biodiversity offset policies, using the formulation of the 'sum of perceived differences' ($\sum \Delta_i$) in Appendix 1. Normalised sum of perceived differences also given. Data from Maron et al. (2015).

	A1: assumed counterfactual (% habitat loss per annum)	A2: assumed counterfactual (proxy, % deforestation per annum)	Perceived difference ($\sum \Delta_i$)	Normalised sum of perceived differences
New South Wales (NSW) 'BioBanking'	0.55	0.21 (min 0.14, max 0.32)	0.34	1.6
New South Wales (NSW) 'Major Projects'	0.35 (min 0.07, max 0.65)	0.22 (min 0.00, max 0.42)	0.13	0.6
Queensland (Qld)	1.00	0.60 (min 0.32, max 1.96)	0.4	0.7
South Australia (SA)	1.48 (min 0.47, max 2.49)	0.09 (min 0.07, max 0.14)	1.39	15.4
Western Australia (WA)	1.50 (min 1.00, max 1.50)	0.19 (min 0.09, max 0.34)	1.31	6.9
Victoria (Vic)	3.12 (min 2.01, max 4.20)	0.45 (min 0.12, max 0.80)	2.67	5.9

22

Table 2: *some possible causes of divergence between different reference frames, for actors observing conservation interventions*

Component of frame (F)	Possible areas of divergence	Examples of divergence potentially leading to conflict	Relevant references
Physical parameter space (E^n)	Number and type of parameters	Which parameters to incorporate (e.g. objectives and targets of interventions)	Burgman et al., 2011; Maxwell et al., 2015
		Decision whether to include certain physical parameters (e.g. temperature), or otherwise	Poiani et al, 2011
		Treating the intervention as dynamic (that is, incorporating 'time' as a parameter) or otherwise	Corlett, 2016
Observer's personal reference frame (F_{ob})	Personal values	Degree to which relationship exists between biodiversity and personal well-being	Woodhouse et al., 2015
		Assumption that high biodiversity is 'better' than low biodiversity (e.g. species numbers)	Bull & Maron, 2016
		Different intrinsic incentives to engage in conservation interventions	Reddy et al., 2017
	Social values	Preference for focus on conservation of certain components of biodiversity over others	Marris, 2013
	Cultural values	Attribution of sacred values vs. secular values to components of biodiversity	Daw et al., 2015; Biggs et al., 2017
Coordinate system (C)	Choice of indicators	Measurement of outcomes in relation to biological diversity, or to functionality (and consequent ecosystem service provision)	Naidoo et al., 2008
	Choice of scale	Evaluating interventions on the spatial scale of individual projects vs. a landscape scale	Bull et al., 2014; Maron et al., 2018
		Choice of temporal scale over which interventions are to be evaluated for efficacy	Bull et al., 2014
		Decision to include historical context, or otherwise	Willis & Birks, 2006;