Perspective

Do biodiversity-ecosystem functioning experiments inform stakeholders how to simultaneously conserve biodiversity and increase ecosystem service provisioning in grasslands?

Valentin H. Klaus,⁎ Mark J. Whittingham, András Báldi, Sönke Eggers, Richard M. Francksen, Matthew Hiron, Eszter Lellei-Kovács, Caroline M. Rhymer, Nina Buchmann

Institute of Agricultural Sciences, ETH Zürich, Universitätstr. 2, 8092 Zürich, Switzerland
School of Natural and Environmental Sciences, Newcastle University, NE1 7RU Newcastle-Upon-Tyne, UK
Institute of Ecology and Botany, Centre for Ecological Research, Alkotmány u 2-4, 2163 Vácrátót, Hungary
Department of Ecology, Swedish University of Agricultural Sciences, Box 7044, SE-750 07 Uppsala, Sweden

ARTICLE INFO

Keywords:
BEF research
Biodiversity crisis
Real-world grassland management
Ecosystem services
Knowledge generation
Plant species richness

ABSTRACT

Two key stakeholders primarily important for nature conservation are farmers (and their lobby groups) and conservationists. Both have substantial inputs into environmental strategies and policies calling for biodiversity conservation aimed to directly increase ecosystem services. The scientific literature concurs that as biological diversity increases so do ecosystem functions and services in grasslands. While the evidence for this is strong, the majority comes from controlled small-scale biodiversity-ecosystem functioning (BEF) experiments. Thus, it is unclear whether the scientific basis for implementing BEF relationships into practice is sufficiently evidenced. Here we explore the applicability of findings from BEF experiments to the conservation and management of temperate grassland, a widespread and potentially highly biodiverse habitat. While we acknowledge that BEF research can reveal insights into fundamental mechanisms, the saturation of biodiversity effects at low levels and unrealistic (management) treatments widely impede the applicability of these experimental results to permanent grasslands. Additionally, the integration of BEF research results into practice is considerably hampered by experimental studies not answering stakeholders’ crucial questions, e.g. is there evidence of biodiversity conservation potentials? Thus, stakeholders do not have a strong evidence base for taking decisions for the addressed management goals, except intensive production in (species-poor) temporary grasslands. If BEF work is to inform stakeholders future research needs to overcome unrealistic management, missing stakeholder involvement and ineffective communication. A new generation of applied BEF experiments employing applied, multi-actor approaches is needed to facilitate the relevance of BEF research for nature conservation, agriculture and land management.

1. Introduction

Biodiversity is no longer only regarded as a passive feature of an ecosystem, but also as an option to positively affect the functioning of ecosystems and the services that flow from them (e.g. Cardinale et al., 2012; Isbell et al., 2011, 2015; Kleijn et al., 2019). For example, Cardinale et al. (2012) in their review study state the ‘general rule’ that reductions in the number of genes/species-functional groups of organisms reduces the efficiency by which whole communities capture biologically essential resources and convert those resources into ecosystem services (e.g. biomass). As a consequence, governments and international institutions have promoted policies and strategies, which reward practices expected to conserve and promote high levels of biodiversity to simultaneously increase ecosystem service provision (Balvanera et al., 2013; Mace, 2014). Examples are the UN CBD’s National Biodiversity Strategies and Action Plans, the global assessment of the Intergovernmental Science-policy Platform on Biodiversity and Ecosystem Services, the European Union’s Science for Environment Policy (2015), the USAID’s Biodiversity Policy or the Endangered Species Act in the USA. The scientific concept behind positive biodiversity
effects (sensu Loreau and Hector, 2001) on ecosystem service provisioning has arisen from biodiversity-ecosystem functioning (BEF) research.

For more than two decades, experimental BEF studies manipulated plant biodiversity primarily in grassland-like model communities to measure the responses of different ecosystem functions and services (Manning et al., 2019). BEF experiments were originally aimed to assess the consequences of species loss for ecosystems (Srivastava and Velland, 2005), and have advanced the fundamental understanding of linkages between biodiversity and ecosystem functioning in grasslands (Salvanera et al., 2006; Eisenhauer et al., 2016). However, although there is some support for positive BEF relationships within natural or semi-natural communities (van der Plas, 2019), most of the evidence has been gathered in controlled experiments (e.g. Wardle, 2016). Thus, despite the recent claims to use experimental BEF findings in real-world ecosystems (e.g. Cardinale et al., 2012; Eisenhauer et al., 2016), it remains unclear if BEF research can be used to reliably guide grassland management (Wang et al., 2019), also because it has never been designed for this purpose but to study the effects of species loss. Here, we discuss the strength of evidence to link outcomes from (mainly) experimental BEF research to the management of temperate grasslands, a widespread, potentially biodiverse and important land-use type in many countries, and the applicability of this research to inform management of non-botanical taxa and at larger than plot scales.

2. Evidence of biodiversity effects for different management goals

The vast majority of temperate grasslands rely on regular management. BEF studies were not designed specifically to link to day-to-day farm management practices but they might inform such actions. The main motivation for managing grassland is a gradient between high economic production and nature conservation depending on the desired outcomes. Recently, the idea of what agricultural land should deliver to society has shifted towards the integration of multiple aims into ecosystem service multifunctionality (Fanin et al., 2018; Manning et al., 2019), which adds a third land management goal to our considerations.

2.1. Production

To be of relevance for agricultural production, BEF studies need to mimic typical grassland management, here we explore whether this has been achieved.

First, from an agricultural point of view, the random combination of species as used in large grassland biodiversity experiments (Lepš, 2004) is problematic and leads to the overestimation of biodiversity effects when compared to actual land management. Experiments have shown a strong selection effect, which is part of the net biodiversity effect, but relies on random species assembly in the model communities (Loreau and Hector, 2001). In practice, the clearly non-random choice of plant species focusing on high-yielding and best performing species considerably diminishes this part of the biodiversity effect. Unfortunately, these agriculturally relevant species have been widely neglected when designing BEF experiments, for example the most relevant grass species of productive temperate grasslands, * Lolium perenne*, has not been included in Europe's largest grassland BEF experiment (*Jena Experiment*; Weisser et al., 2017). This makes it hard to link experimental outcomes to typical land management scenarios. Other design issues such as the annual removal of unsown species by hand to maintain the experimental biodiversity gradient have been discussed before (e.g. Huston, 1997; Lepš, 2004; Pfisterer et al., 2004; Roscher et al., 2016).

Second, BEF research rarely address stakeholders’ information needs (Binder et al., 2018). The experimental design is frequently insufficient to address concerns regarding expected outcomes in terms of agricultural and economical gains. Scientists often focus on overyielding, i.e. a mixture of species overyields when its biomass production is greater than that of the average monoculture of the species contained in the mixture (Fig. 1). In contrast, for farmers the actual difference in yield provided by a comparison of the best performing monocultures and mixtures is of relevance, i.e., transgressive overyielding (Trenbath, 1974). This concept of comparing the best or the average performance of monoculture and mixtures has been formulated for biomass production, but is basically valid for all ecosystem functions and services. Unfortunately, transgressive overyielding is not addressed in the majority of BEF studies, although it is a crucial piece of information: Web of Knowledge search (24th September 2019 across all years) identified 114 studies addressing “grassland” AND “overyielding”, while only 31 studies addressed “transgressive overyielding”. When transgressive overyielding was compared to overyielding, transgressive overyielding of the most diverse mixtures was considerably less significant (only in 12% instead of 79%) and in many cases biodiversity even appeared to have a negative effect (25% instead of 0% of studies; Wang et al., 2019; Cardinale et al., 2007).

Third, management intensity has hardly been addressed in BEF experiments, although essential for economic production and relevant to any biodiversity effects in grasslands. For example, in an additional study to the Jena BEF Experiment, all model communities of regionally adapted species also underwent different management treatments (Weigelt et al., 2009). The effect of biodiversity on yield was larger than...
the effect of maximal intensification, meaning increasing from one to 60 sown plant species resulted in higher yields than increasing fertilization rates from 0 to 200 kg N per ha and doubling the number of cuts (from 2 to 4 per year). However, plant species typical for more intensively used grasslands were absent in this assessment and highest management intensity did not per se result in highest yields (Weigelt et al., 2009), pointing to a still suboptimal and unrealistic management scenario. Neither this experiment nor synthesis studies assessing the persistence of the BEF relationships under increased nitrogen availability (Craven et al., 2016; Tilman et al., 2012) reported transgressive overyielding or any other direct agronomic benefit. Thus, the proof of applicability is missing.

Fourth, important information for farmers is lacking. Except productivity being the most frequently assessed ecosystem service in BEF experiments, few studies considered yield quality, a crucial agronomical aspect (Balvanera et al., 2013; but see Bullock et al., 2007; Isbell and Wilsey, 2011; Schaub et al., 2020). Additionally, only limited information is currently available about resulting economic benefits (but see Finger and Buchmann, 2015). Although tools to extract economic information of direct relevance for farmers have been applied (Binder et al., 2018), the usefulness of their results is questionable given the technical issues mentioned above.

Only those BEF experiments conducted in temporary grasslands have been setup to mimic the real-world management situation. Several studies found that using four-species mixtures under different management intensities had strong positive effects on yield, weed suppression and other, mostly production-related, services when compared to monocultures (e.g. Finn et al., 2013; Nymeyer et al., 2009). This enables a direct transfer of BEF findings into practice (Manning et al., 2019), but only for intensively used, non-permanent grasslands with up to four–in some cases also six or more (Brophy et al., 2019; Grange et al., 2019)–different plant species.

Besides the debated strongly controlled BEF experiments, Bullock et al. (2001, 2007) performed a close-to-practice field trial to test the agricultural performance of species-poor mixtures (7–21 species, almost only grasses) compared to species-rich mixtures (25–41 species of grasses, legumes, and forbs) in newly sown grasslands on ex-arable land. The authors found considerably higher yields of similar or even better feed quality when using their species-rich compared to species-poor mixtures for grassland planting or restoration, pointing to a win-win-situation for agriculture and biodiversity. This study, despite its simple design with two levels of (functional) diversity, provides important information and assists stakeholders with the restoration of extensively managed grassland on ex-arable land. We suggest more studies are designed in this way, building up on established management practices and species combinations to introduce novel mixtures and approaches. Such applied experiments will be different from previous BEF studies in that not all plant species are necessarily contained in each diversity level while some cultivars should also be included in specific mixtures. Additionally, monocultures that are not realistic and require intensive maintenance (weeding) should be replaced by those of agricultural relevance.

2.2. Nature conservation

For nature conservationists, BEF research is relevant if it effectively contributes to the conservation of biodiversity across a range of taxa which occur at different spatial scales. Here, both the number of species but also whether these are rare or endangered is of relevance to conservation. To our knowledge, no BEF experiment ever included or measured particularly rare or endangered species, such as from national or international red lists such as from the IUCN Red List of Threatened Species. Studies considering less common or locally rare species as functionally relevant in real-world grasslands (e.g. Lyons et al., 2005; Soliveres et al., 2016a,b) showed positive but to some extent also some negative effects of these; and such studies have clearly not been designed to inform stakeholders about rare species’ practical relevance. As the current functional role of a species in a grassland is linked to its absolute abundance, single rare species are not functionally relevant (Schwartz et al., 2000) or even dilute positive effects of other species of particularly high value for a specific ecosystem service (Binder et al., 2018; Kleijn et al., 2015). Only if the abundance of a rare species increases due to environmental change, rare species might become important for ecosystem functioning (Jain et al., 2014).

Concerning the level of biodiversity needed to achieve a positive BEF effect, experiments with intensively managed temporary grasslands are rarely relevant for conserving plant diversity as they usually contain around four grasses and legumes species (Finn et al., 2013). Concerning BEF experiments with extensively managed permanent grassland model communities, which usually contain 32 to 60 species depending on the respective trial, a recent study by Buchmann et al. (2018) showed that the effect of plant diversity on yield saturates already at eight out of 60 sown plant species. This is considerably below the species richness of comparable low-input semi-natural grasslands within the same region (Buchmann et al., 2018; Klaus et al., 2011). Thus, BEF experiments cannot prove the need for high grassland biodiversity, i.e., > 30–40 species, but in contrast help to understand the mechanism underlying the benefits of using species mixtures instead of monocultures (e.g. Binder et al., 2018; Weisser et al., 2017).

Despite the restricted relevance of many BEF experiments for nature conservation, the findings from the field trial of Bullock et al. (2007) might still be of interest, as their species-rich mixtures appeared to be of similar richness as comparable semi-natural communities in the region. However, as the functional relevance of individual species in these mixtures were not tested, “free rider species” could not be differentiated from functionally important species such as legumes.

Moreover, grasslands support a range of threatened biodiversity including birds, mammals and invertebrates but the evidence of increased biodiversity impacts on ecosystem function and services from non-botanical taxa is considerably less prevalent in the literature. This coupled with less information beyond the plot scale from BEF experiments is a substantial evidence-gap to support actions on grasslands.

2.3. Multifunctionality

Positive effects of plant diversity on single ecosystem services saturate at relatively low plant species diversity, often at 4 to 8 species, as only a small proportion of a community is usually relevant to maintain a large share of one service (Fanin et al., 2018; Schwartz et al., 2000; Meyer et al., 2018). This has also been shown for services provided by birds (Hiron et al., 2018) and pollinators (Kleijn et al., 2015). As different species might be relevant for different services, managing for plot-scale multifunctionality could potentially strengthen the relevance of biodiversity effects for grassland management and might result in higher levels of conserved biodiversity (e.g. Isbell et al., 2011; Dee et al., 2017; Meyer et al., 2018). However, recent analyses have presented contradicting results, namely that targeting at multiple functions or services does not necessarily require higher levels of biodiversity and thus strongly question the usefulness of current approaches of plot-scale multifunctionality for decision-making (Gamfeldt and Roger, 2017; Slade et al., 2019).

3. Relevance for stakeholders

Farmers focused on production are likely to use species-poor mixtures of around four plant species for intensively-used temporary grasslands. Advantages of these species-poor mixtures over monocultures are of practical relevance across a wide climatic gradient (Finn et al., 2013), with low- or even unfertilized mixtures being able to achieve similar yield levels as highly fertilized grass monocultures (Sanderson et al., 2004; Nymeyer et al., 2009). However, the BEF literature provides little evidence to support management of low-input
permanent grassland given the unrealistic management treatments and experimental designs as well as missing information of agricultural and economic benefits.

**Nature conservation agencies** aimed primarily at biodiversity conservation are unlikely to find much of practical value from the BEF experiments with their focus on small-scale experiments. As long as biodiversity effects do not require truly high (> 30–40 species) and/or endangered or rare species (many of which operate at larger spatial scales than the plot), BEF research cannot provide relevant evidence for stakeholders wishing to link biodiversity conservation, (at least that focussed on rarer species) with ecosystem service benefits.

**Multifunctionality** has no applicable management goal on the plot-level as its relevance is still debated and its practical usefulness doubtful (Slade et al., 2019). Likewise, multifunctionality indices do not overcome trade-offs between services such as production and rare species conservation (Binder et al., 2018). Additionally, targeting multifunctionality is not well supported by the structure of agricultural compensation schemes (Vrebos et al., 2017).

4. Designing novel BEF studies and improving communication

We see a significant risk in deducing conclusions for decision-making from strongly controlled BEF experiments, as this can put credibility of BEF research at risk. As one example, publications often do not inform about realized diversity of the experiments but refer to sown diversity, a characteristic of the experiment but irrelevant information for nature conservationists or farmers that need to know the actual number of plant species involved. Effective communication between stakeholders and researchers is a major goal if future BEF research is to inform practical land management (Hulme, 2014). Thus, interacting with stakeholders and appropriately communicating findings of BEF studies is of greatest importance, and should be improved as follows:

4.1. Specific applied outcomes are needed

Land managers should not be left alone with the vague notion that biodiversity (e.g. plant diversity) somehow positively affects production, its stability, and other ecosystem services. Future studies should clearly address and quantify agronomic and ecological benefits from biodiversity and set these off against actual costs (e.g. for seeds) as investments in higher biodiversity (Manning et al., 2019) and assess under which (environmental) circumstances these effects can be achieved (Hiddink et al., 2009). This also includes adopting statistical methods to evaluate BEF research to provide information relevant to stakeholders, not only to scientists.

4.2. Careful communication is crucial

Dissemination of experimental results needs to clearly present experimental design, define the type and level of biodiversity considered and separate fundamental from applied BEF studies. Separating high or full biodiversity, which for conservationists usually refers to the upper end of a species richness gradient in an ecosystem, from functional effects of rather low biodiversity, i.e. mixture effects, is an important step to avoid confusion and misinterpretations. Thus, results also need to be robust to fulfill expectations that are raised, which also needs testing of BEF effects under different environmental conditions and with different management practices. Many studies performed under more natural field conditions report a negative relationship between plant diversity and productivity, due to e.g. species pool effects, natural invasion at low sown diversity and confounding effects of site conditions and management (Grace et al., 2007; Rychtěcká et al., 2014; Sandau et al., 2019). As such, it is essential that experimental findings of exactly the opposite (positive) relationship are confirmed outside of highly controlled environments before reaching out to stakeholders.

4.3. Collaboration is crucial

Stakeholders and researchers need to work together and maintain a dialogue regarding their shared objectives in biodiversity conservation and ecosystem service provision. Involving stakeholders already in designing and even conducting BEF experiments (Manning et al., 2019) could significantly increase representativeness for real-world systems, and provide ample opportunities for direct knowledge transfer.

Current BEF approaches are widely limited to newly established plant communities. Thus, it appears to be of high practical relevance for future BEF research to assess whether a species enrichment of established low-diversity grasslands, e.g. by seeding additional species in gaps (Klaus et al., 2017; Kiss et al., 2020), can significantly increase production and/or other ecosystem services; and whether the money spent for such restoration measures can be economically compensated e.g. by higher yields or lower costs for sward maintenance.

5. Conclusions

Despite 25 years of BEF research, we still see a considerable gap between (experimental) BEF research and its applicability for management of permeant grasslands aiming to reconcile biodiversity and production conflicts. BEF research has increased awareness for biodiversity and identified fundamental ecological mechanisms. However, there is still little evidence on which to base land management recommendations, neither for nature conservation with ambitious biodiversity targets nor for economical viable grassland production. Thus, BEF research now has to move on from fundamental research to applied trials and multi-actor approaches to tackle all relevant practical aspects and information needs. Beside strong stakeholder involvement, this calls for the inclusion of agriculturally relevant species in BEF experiments, for field level assessments in different environmental conditions, for more realistic management scenarios, for testing truly high biodiversity (> 30–40 species) to provide evidence for nature conservation benefits and for options to maintain BEF effects under more intensive management.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was stimulated and (partly) funded by the project SUPER-G (https://www.super-g.eu) that received funding by the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement N. 774124.

Author contributions

VHK had the idea; NB, MJW, and VHK developed the concept and all other authors contributed to intensive discussions and the final manuscript.

References
