Check for updates

Received: 28 November 2025 | Accepted: 22 October 2026

DOI: 10.1111/2041-210x.70449

Methods in Ecology and Evolution

RESEARCH ARTICLE

Quantifying priority effects in community assembly experiments

Benjamin M. Delory[1] | Chelsea J. Little[2,3] | Hannah Marton[2,3] | Fletcher W. Halliday[4] | Elle M. Barnes[5] | Isabelle Stiver[4] | Judith M. Sarneel[6] | Tamara L. H. van Steijn[6] | Tess N. Grainger[7] | Julia Dieskau[8] | Paul Kardol[9] | Tiffany M. Knight[8,10,11,12] | Emma Ladouceur[11,13,14,15] | Miriam N. Ojima[16] | Christiane Roscher[11,17] | James T. Stroud[18] | Vicky M. Temperton[19] | Chhaya M. Werner[20] | Tadashi Fukami[21]

Correspondence
Benjamin M. Delory
Email: b.m.m.delory@uu.nl

Funding information
Deutsches Zentrum für integrative Biodiversitätsforschung Halle-Jena-Leipzig

Handling Editor: James Alexander Orr

Abstract

1. The history of community assembly can affect biodiversity and ecosystem functioning. This historical contingency is often driven by priority effects, whereby species interactions are influenced by the order and timing of species arrival. Experimental manipulations of the order of arrival of species have been instrumental in elucidating the causes and consequences of priority effects. Nevertheless, quantifying priority effects is challenging due to the complexity of experimental designs and the lack of consensus regarding the most appropriate comparisons and metrics for estimating the direction and magnitude of priority effects.
2. This paper compares approaches that differ in (1) the reference scenario against which the performance of the response species is evaluated (synchronous arrival or reverse arrival), (2) the method used to standardise the interval between species introductions and the measurement of species performance (species-specific or assembly-specific comparison) and (3) the metric used to quantify priority effects.
3. Using a previously published in vitro experiment that manipulated the order of arrival of four human-associated *Bifidobacterium* species as an illustrative example, we show that different approaches and metrics to quantifying priority effects can yield divergent conclusions regarding the strength and, in some cases, even the direction of priority effects. This finding underscores the importance of carefully selecting and clearly reporting the approach and metric used, as each is best suited to address distinct questions related to arrival history.

For affiliations refer to page 14.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2026 The Author(s). *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

4. When designing experiments to quantify priority effects, we recommend always including a synchronous reference where all species are introduced simultaneously and grow for the same duration as the late-arriving (response) species in the target arrival sequence. It may help to include a reference scenario in which species are introduced sequentially, with the response species having the same time to establish and develop in the target and reference arrival sequences. Although we could not use this approach in our study case, it seems well suited to isolate order of arrival effects on species performance. Additionally, we advocate using standardised mean differences and log response ratios to quantify priority effects, as these metrics offer robust estimates and facilitate quantitative synthesis across taxa and ecosystems.

KEYWORDS
community assembly, effect size, experimental design, order of arrival, priority effect

1 | INTRODUCTION

Over the past two decades, there has been increasing appreciation of the century-old idea (Palmgren, 1926) that the order and timing of species arrival during community assembly influence biodiversity patterns as well as the functioning of ecological communities (Alonso-Crespo et al., 2025; Chappell et al., 2022; Chase, 2003; Fragata et al., 2022; Vannette & Fukami, 2014; Weidlich et al., 2016; Zou & Rudolf, 2023). This historical contingency is often caused by priority effects, in which arrival order and timing determine the way species affect one another in communities (Drake, 1991). Priority effects have been documented in a range of taxa, systems and spatial and temporal scales (Stroud et al., 2024). Priority effects arise because the early arrival and establishment of a species in a given location can reduce the availability of essential resources (niche pre-emption) or modify the (a)biotic environment (niche modification) in a way that affects the establishment, growth and reproduction of later-arriving species, either positively or negatively, thereby modifying species interactions (e.g., per-capita interaction strength), community composition and ecosystem processes (Debray et al., 2022; Stuble & Young, 2020; Torres et al., 2024; Wilsey, 2021; Zou & Rudolf, 2023).

Experimental studies have provided critical insights into priority effects by manipulating the order of arrival of species or groups of species (Svoboda et al., 2018). In a two-species system, such experiments typically involve the sequential introduction of an effect species, which arrives first and alters environmental conditions, followed by a response species, which arrives later and experiences the consequences of these changes (Stuble & Souza, 2016). In this paper, the cost or benefit experienced by the response species as a consequence of arriving after the effect species is what we refer to as a priority effect. Priority effects are considered negative (inhibitory) when the presence of the effect species reduces the performance (i.e. abundance and biomass) of the response species relative to an alternative arrival sequence, and positive (facilitative) when it increases the performance of the response species. A common experimental approach to quantifying priority effects is therefore to compare community outcomes across different arrival sequences, in which the identity and order of arrival of species are experimentally manipulated. Priority effects (*PE*) are then quantified by comparing the performance of a response species when it arrives later than the others (Y_{Late}) with its performance measured in a reference situation (Y_{Ref}).

In some cases, however, the focus is not on the response species but on the effect species. Here, the question shifts from quantifying the cost or benefit experienced by a species arriving late to quantifying the cost or benefit of arriving early (Delory, Weidlich, Kunz, et al., 2019; van Steijn et al., 2026; Werner et al., 2016). Although the priority effect definition used in this paper focuses on the consequences of arrival order for the response species, the same experimental logic can be applied to quantify the consequences of early arrival for the effect species. Specifically, the benefit or cost of arriving early can be assessed by comparing the performance of the effect species when introduced early with its performance when arriving synchronously or later than the response species.

Theoreticians and empiricists often define priority effects differently. In theoretical work, they are often framed in terms of long-term community outcomes, whereby the order of arrival of species leads to alternative stable states (Grainger et al., 2019; Ke & Letten, 2018; Song et al., 2021; Sutherland, 1974). Under this equilibrium-based view, differences in transient dynamics would not constitute a priority effect if all trajectories ultimately converge to the same equilibrium (Stroud et al., 2024). However, in many systems equilibrium is rarely achieved because of frequent disturbance (Fukami & Nakajima, 2011, 2013) and equilibrium dynamics are inherently difficult to identify in natural systems and to achieve in experiments (Grainger et al., 2025). Consequently, the focus of much empirical work on priority effects has been on how arrival order shapes community assembly and ecosystem

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

functioning in systems that may never reach a stable equilibrium, particularly as disturbance-prone systems where the repeated re-setting of the community allows species to arrive sequentially at an unoccupied site, making priority effects more likely. In this paper, we therefore adopted a definition of priority effects that focuses on the consequences of arrival order over a specified time frame without restricting the definition to cases where communities reach an equilibrium, while recognising that both the time lag between species introductions and the interval between species introduction and measurement of species performance can strongly influence the strength and direction of priority effects. Likewise, we focus here on experimental studies, as detecting priority effects in observational data is a related unresolved topic that is outside the scope of this paper.

So far, two factors have hindered quantitative synthesis of priority effects across studies. First, experimental design and timing can differ in various ways, including the choice of reference situation and approach used to standardise the time interval between introductions and measurements. Second, metrics used to quantify priority effects differ between studies (Delory, Weidlich, Kunz, et al., 2019; Delory, Weidlich, von Gillhaussen, & Temperton, 2019; Dieskau et al., 2024; Sarneel et al., 2016; van Steijn et al., 2025; Vannette & Fukami, 2014), and these metrics cannot be easily combined or compared. These challenges underscore the need for a common framework for quantifying priority effects. Developing such a framework is essential not only for designing experiments that enable better standardisation in quantifying priority effects, but also for comparing and synthesising results across studies. Such a framework is a key step towards understanding how common and strong priority effects are, and what ecological factors shape them across systems.

In this article, we provide an overview of methodological approaches used to quantify priority effects in community assembly experiments. We focus on how to quantify the strength of priority effects without identifying the mechanisms behind the priority effects. However, we hope that the methods discussed in this paper will facilitate future comparison of the strength of priority effects arising from different mechanisms. Specifically, we compare approaches that differ in (1) the reference scenario against which the performance of the response species is evaluated, (2) the method used to standardise the interval between species introductions and the measurement of species abundance, and (3) the metric used to quantify priority effects. For clarity, we describe these approaches in the context of simple two-species systems. We then illustrate their application using data from a previously published experiment that manipulated the order of arrival of four human-associated *Bifidobacterium* species (Ojima et al., 2022). Using these data, we first examine how estimates of priority effects vary with the choice of reference scenario and time interval standardisation for a given priority effect metric. Then, we assess the extent to which different methodological approaches yield comparable estimates of priority effects across metrics. Building on this analysis, we offer recommendations for designing experiments and quantifying priority effects. We conclude by discussing how these recommendations can be extended to more complex systems with multiple introduction events.

2 | METHODS

2.1 | Four methodological approaches for quantifying priority effects

Let us start with a simple case, where two species, species E (effect species) and R (response species), interact with each other. Here, E and R are two individual species, but they could also be two different genotypes of the same species, or two different species groups (e.g., functional groups or life strategy groups). In this scenario, species E arrives before species R (Figure 1a). The performance of species present in the community is then monitored over a period of time, which varies depending on the system studied (e.g., microbes, insects and plants). We define Δt_R as the time interval between the arrival of the late-arriving species (R, response species) and the moment when species performance is recorded to quantify the priority effect of species E on species R. Depending on the study system, Δt_R can be a few hours, several days or several years. The total duration of the experiment, which is the time interval between the arrival of the effect species and the moment when species performance is recorded, is denoted as Δt_{EXP} (Figure 1a). The time lag between the arrival of both species is therefore $\Delta t_{EXP} - \Delta t_R$.

To quantify the cost or benefit of arriving late for the response species (R), we would choose a reference scenario as a basis of comparison and standardise the time interval between arrival events and performance measurements (Table 1). Even though the same logic applies to situations where the cost or benefit of arriving early needs to be quantified, the next sections of this paper will mostly focus on quantifying the cost or benefit of arriving late (but see Figure S1 for an overview of approaches designed to quantify the cost/benefit of arriving early). Studies that compare how well a species grows in monocultures vs. in mixed cultures with other species cannot quantify priority effects as defined in this paper. These studies can quantify the effect of species on one another's growth, but not priority effects because they do not look at the effect of species arrival order or timing.

There are two alternative reference scenarios with respect to arrival order: a synchronous reference, where both species are introduced at the same time (upper panels in Figure 1b), and a reverse reference where, all other things being equal, the positions of E and R are reversed (bottom panels in Figure 1b).

In addition, there are two options with respect to time intervals. In the first option, which we will refer to as the species-specific comparison, the time lag between the arrival of the response species (R) and the measurement of its performance (Δt_R) is kept constant in the target and reference scenarios (left panels in Figure 1b). In this case, calculations are standardised for

20412108x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

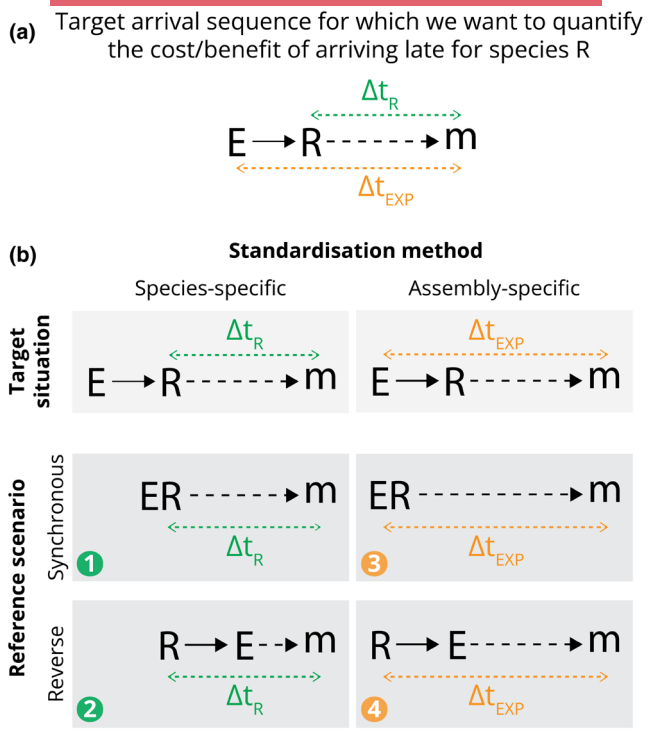


FIGURE 1 Four methodological approaches for quantifying priority effects (i.e. the cost/benefit of arriving late). See Figure S1 for an overview of approaches to quantify the cost/benefit of arriving early. Panel (a) shows the target arrival sequence for which we want to quantify a priority effect. Panel (b) illustrates four possible approaches to quantify a priority effect using data collected in a community assembly experiment. In this figure, effect (E) and response (R) represent the introduction timing of two different species, functional groups or genotypes. The solid black arrows represent the time between the arrival of effect and response species. The dashed black arrows represent the interaction time between effect and response species. The letter m marks the time at which a performance measurement (e.g., abundance of E and R) is recorded for both the target and reference arrival sequences. To quantify the priority effect of the effect species on the response species (R), two questions must be answered: (1) What reference arrival sequence should be used? and (2) How should the time between arrival and measurement be standardised? These choices define the four approaches to quantifying priority effects (numbered from 1 to 4 in (b)). Two of these approaches use a synchronous reference situation (1 and 3 in (b), corresponding to scenario iii in Table 1), while the other two use a reverse reference situation (2 and 4 in (b), corresponding to scenario iv in Table 1). In the left panels, the time from species R's arrival to measurement (Δt_R) is held constant in the target and reference sequences. This species-specific time interval ensures that the response (R) species spends the same time in the environment across arrival scenarios. In the right panels, the total time from the first species' arrival to measurement (Δt_{EXP}) is fixed. This assembly-specific standardisation ensures that the duration of community assembly is identical in the target and reference scenarios being compared.

the species of interest (i.e. the response species), which was present in the environment for the same amount of time in the two arrival orders being compared. In the second option, referred to as the assembly-specific comparison, it is the time lag between the first introduction event and the measurement (Δt_{EXP}) that is kept constant in the target and reference arrival orders (right panels in Figure 1b). In this approach, the calculations are standardised for the duration of community assembly, which is identical in the two arrival sequences compared.

Figure 1b distinguishes the four methodological approaches of quantifying priority effects, depending on whether a reverse or synchronous arrival is used as a reference and whether a species-specific or assembly-specific comparison is used to calculate the direction and magnitude of the priority effect. In all cases, priority effects can happen through different mechanisms, which include changes in per-capita interaction strength (Stroud et al., 2024; Zou & Rudolf, 2023). Hereafter, we will refer to these approaches by their numbers in Figure 1b (Approaches 1 to 4). All approaches have been used in the literature (Table 1), but inconsistently and often not simultaneously. As we will demonstrate, the four approaches do not necessarily provide the same conclusions. This variability highlights not only the importance of choosing an approach that is appropriate for the study system and research question, but also of clearly describing the method used so that the conclusions drawn are transparent, and the findings can be compared with those of other studies.

2.2 | Metrics for quantifying priority effects

Priority effects have been quantified using different metrics (Delory, Weidlich, Kunz, et al., 2019; Delory, Weidlich, von Gillhaussen, & Temperton, 2019; Sarneel et al., 2016; van Steijn et al., 2025; Vannette & Fukami, 2014). Existing commonly used metrics for priority effects are as follows: additive intensity index (Equation 1, bounded between −1 and +2), a normalised difference (Equation 2, bounded between −1 and +1), and a log response ratio (Equation 3a: uncorrected log response ratio; Equation 3c: log response ratio shown with bias correction as described in Lajeunesse, 2015). The sampling variance of the uncorrected and bias-corrected log response ratio are shown in Equation 3b and 3d, respectively. In these equations, $\overline{Y}_{Late}$ is the average performance of the response species when it arrives late, while $\overline{Y}_{Ref}$ is the average performance of the same species measured in a reference situation (either a reverse or a synchronous sequence of arrival). The parameters s_{Late} and s_{Ref} are the observed standard deviations, and n_{Late} and n_{Ref} are the number of observations in each group. Although any of these metrics enable us to assess the direction and strength of priority effects across treatments within single studies, they have different mathematical properties (Díaz-Sierra et al., 2017), which makes quantitative synthesis of priority effects challenging and leaves cross-metric comparisons difficult (Figure 2). We do not recommend using the bias-corrected log response ratio (Equations 3c and 3d) to quantify priority effects because, in some situations, the bias correction can change the direction of the priority effect, which is not a desirable mathematical property for a priority effect metric (Figure 2).

20412108x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Commonly used study designs for quantifying priority effects. The list of references is not exhaustive.

No.	Design	Part of Figure 1b where scenario is depicted	Strengths	Caveats	Recommendations	Example references
i	Manipulation of species' initial (relative) density (with all species introduced simultaneously)	Not depicted	Mimics a scenario where species must invade from rarity, facing an already- established competitor while reducing experimental complexity	Without a time lag, this design may not always detect niche modification because the early-arriving species (here, the more abundant species) may not have time to change the environment before the arrival of the later species	Ensure that design encompasses scenarios where each species has the higher density, and scenario(s) where all species have equal densities, to assess the role of species interaction at different densities	Fragata et al. (2022); Price & Morin (2009)
ii	Manipulation of ages of individuals in different species (with all species introduced simultaneously)	Not depicted	Mimics a scenario where species arrives early and has a head-start in growth compared with the later-arriving species while reducing experimental complexity	Same as above	Same as above	Moore & Franklin (2012); Rasmussen et al. (2014); Stevens & Fehmi (2011)
iii	Comparison of scenario with time lag to scenario where all species arrive simultaneously	Upper two panels of Figure 1b (Approaches 1 and 3 with synchronous reference)	Allows the researcher to isolate the effects of order of arrival from the effects of species interaction under simultaneous arrival	Does not account for how response species would perform without effect species (e.g. in monoculture or when arriving early)	We recommend this design, possibly in combination with design iv	Alonso-Crespo et al. (2023, 2025); Carpenter et al. (2021); Fragata et al. (2022); Green et al. (2024); Kardol et al. (2013); Sarneel et al. (2016); Stuble & Souza (2016); Tucker & Fukami (2014); van Steijn et al. (2025); von Gillhaussen et al. (2014); Weidlich et al. (2016, 2018); Werner et al. (2016)
iv	Comparison of scenario where species arrives early to when it arrives late	Bottom two panels of Figure 1b (Approaches 2 and 4 with reverse reference)	Allows the researcher to measure the costs or benefits of arriving both early and late	Relative density and age will vary among scenarios, so does not assess the competitive ability of any species when on even footing with the others	We recommend this design, possibly in combination with design iii	Delory, Weidlich, Kunz, et al. (2019); Fragata et al. (2022); Hoverman et al. (2013); Sarneel et al. (2016); Vannette & Fukami (2014); Werner et al. (2016)

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by iadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

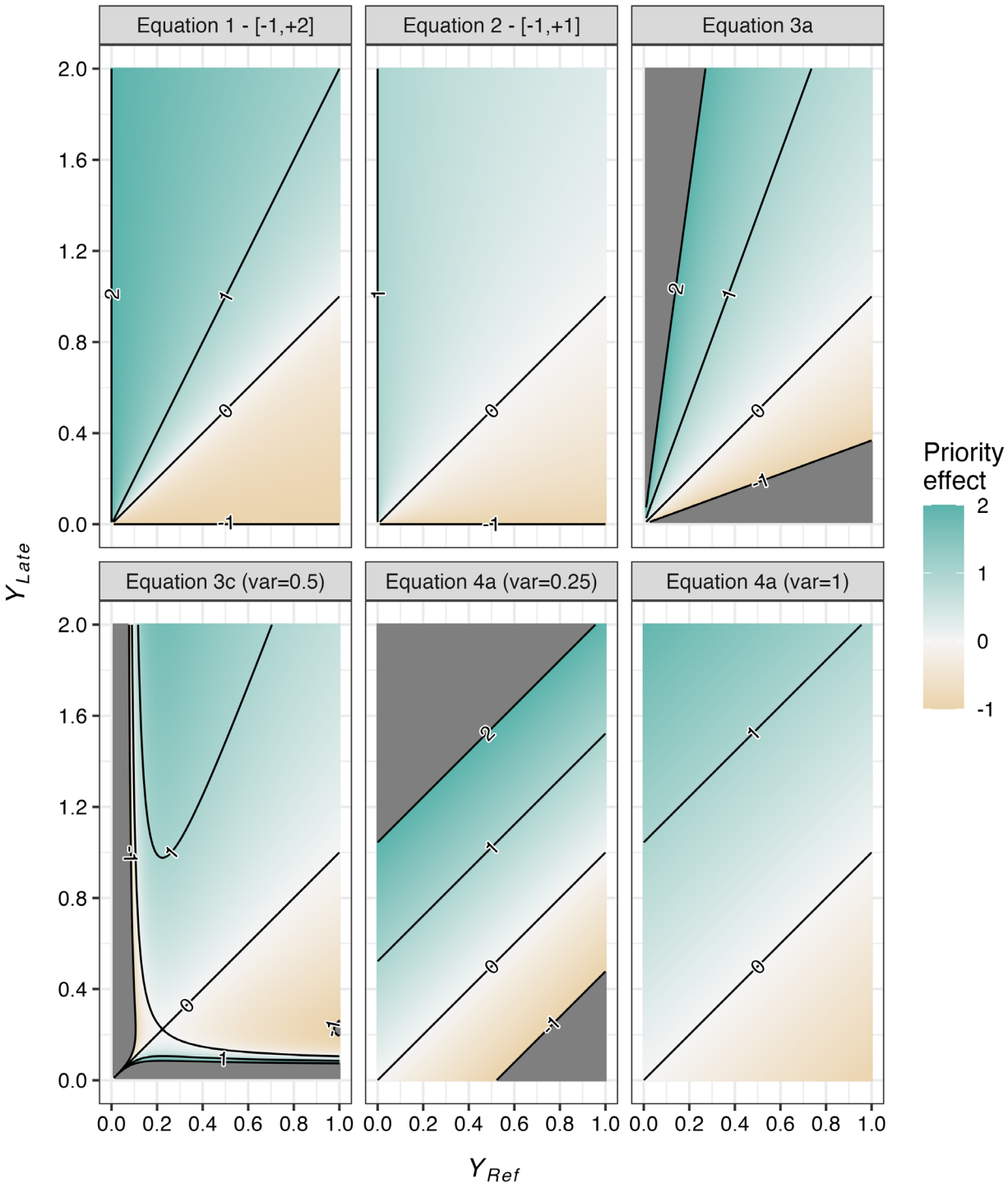


FIGURE 2 Metrics for quantifying priority effects and their relationship/properties between Y_{Late} and Y_{Ref}. The figure shows four metrics for quantifying priority effects (Equations 1 to 4a described in the main text). Y_{Late} and Y_{Ref} are the absolute abundance of the response species (R) measured in a late-arrival and a reference scenario, respectively. When a bias-corrected log response ratio (Equation 3c) or standardised mean difference (Equation 4a) is used to quantify priority effects, the outcome depends not only on the difference in mean performance of the species measured in a late-arrival or a reference scenario (Y_{Late} and Y_{Ref}), but also on the number of observations (n_{Late} and n_{Ref}) and the variance (s^2_{Late} and s^2_{Ref}) in each group. In the lower panels, the number of observations in each group was set to 10 and the calculations were done using different variances. In the calculations, we assumed that the Late and Reference groups have the same variance. Positive and negative priority effects are represented by green and yellow areas, respectively. For clarity and to allow for use of a common colour scale across all panels, only priority effect values between −1 and +2 are displayed. Grey areas indicate regions where metric values fall outside this range, with values below −1 shown on the yellow end of the scale and values above +2 on the green end.

20412108x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Another candidate metric that has been overlooked so far in priority effect studies is the standardised mean difference (Equations 4a and 4b), which is a common effect size used in meta-analyses and is also known as Hedges' *g* (Hedges, 1981; Hedges & Olkin, 1985; Koricheva et al., 2013). Equations 4a and 4b differ in how the bias correction is calculated. When sample size is small, the exact solution relying on the gamma distribution (Γ) can be used (Equation 4a), but when sample size is large, it may be preferable to use the approximate solution (Equation 4b). In comparison with the other metrics mentioned above, Hedges' *g* has a number of advantages. Unlike log response ratios, Hedges' *g* remains valid even if a species' performance metric is zero. This makes it an appropriate fit for priority effect experiments because exclusion (e.g., a species failing to establish) can be a common outcome. Hedges' *g* expresses priority effects in standard deviation units, making results comparable across studies with different measurement types, such as biomass vs. percent cover. Finally, the Hedges' *g* formula adapts to homoscedastic or heteroscedastic data. Assuming homoscedasticity of population variances, the standardised mean difference can be calculated using Equations 4a and 4b. In case of heteroscedastic population variances, the denominator in Equations 4a and 4b can be replaced by the square root of the average variance. The sampling variance of *PE* can be obtained with Equation 4c and can then be used to estimate 95% confidence intervals.

$$PE = 2\frac{\overline{Y}_{\text{Late}} - \overline{Y}_{\text{Ref}}}{\overline{Y}_{\text{Ref}} + |\overline{Y}_{\text{Late}} - \overline{Y}_{\text{Ref}}|} \tag{1}$$

(additive intensity index).

$$PE = \frac{\overline{Y}_{\text{Late}} - \overline{Y}_{\text{Ref}}}{\overline{Y}_{\text{Late}} + \overline{Y}_{\text{Ref}}} \tag{2}$$

(normalised difference).

$$PE = ln\left(\frac{\overline{Y}_{\text{Late}}}{\overline{Y}_{\text{Ref}}}\right) \tag{3a}$$

(uncorrected log response ratio).

$$s^2_{PE} = \frac{s^2_{\text{Late}}}{n_{\text{Late}}\overline{Y}^2_{\text{Late}}} + \frac{s^2_{\text{Ref}}}{n_{\text{Ref}}\overline{Y}^2_{\text{Ref}}} \tag{3b}$$

(sampling variance of uncorrected log response ratio).

$$PE = ln\left(\frac{\overline{Y}_{\text{Late}}}{\overline{Y}_{\text{Ref}}}\right) + \frac{1}{2}\left(\frac{s^2_{\text{Late}}}{n_{\text{Late}}\overline{Y}^2_{\text{Late}}} - \frac{s^2_{\text{Ref}}}{n_{\text{Ref}}\overline{Y}^2_{\text{Ref}}}\right) \tag{3c}$$

(bias-corrected log response ratio).

$$s^2_{PE} = \frac{s^2_{\text{Late}}}{n_{\text{Late}}\overline{Y}^2_{\text{Late}}} + \frac{s^2_{\text{Ref}}}{n_{\text{Ref}}\overline{Y}^2_{\text{Ref}}} + \frac{1}{2}\left(\frac{s^4_{\text{Late}}}{n^2_{\text{Late}}\overline{Y}^4_{\text{Late}}} + \frac{s^4_{\text{Ref}}}{n^2_{\text{Ref}}\overline{Y}^4_{\text{Ref}}}\right) \tag{3d}$$

(sampling variance of bias-corrected log response ratio).

$$PE = \frac{\overline{Y}_{\text{Late}} - \overline{Y}_{\text{Ref}}}{\sqrt{\frac{(n_{\text{Late}}-1)s^2_{\text{Late}} + (n_{\text{Ref}}-1)s^2_{\text{Ref}}}{n_{\text{Late}}+n_{\text{Ref}}-2}}}\left(\frac{\Gamma\left(\frac{n_{\text{Late}}+n_{\text{Ref}}-2}{2}\right)}{\sqrt{\frac{n_{\text{Late}}+n_{\text{Ref}}-2}{2}}\Gamma\left(\frac{n_{\text{Late}}+n_{\text{Ref}}-3}{2}\right)}\right) \tag{4a}$$

(Hedges' *g* with exact bias correction).

$$PE = \frac{\overline{Y}_{\text{Late}} - \overline{Y}_{\text{Ref}}}{\sqrt{\frac{(n_{\text{Late}}-1)s^2_{\text{Late}} + (n_{\text{Ref}}-1)s^2_{\text{Ref}}}{n_{\text{Late}}+n_{\text{Ref}}-2}}}\left(1 - \frac{3}{4(n_{\text{Late}}+n_{\text{Ref}}-2)-1}\right) \tag{4b}$$

(Hedges' *g* with approximate bias correction).

$$s^2_{PE} = \frac{n_{\text{Late}}+n_{\text{Ref}}}{n_{\text{Late}}n_{\text{Ref}}} + \frac{PE^2}{2(n_{\text{Late}}+n_{\text{Ref}})} \tag{4c}$$

(sampling variance of Hedges' *g*).

2.3 | Case study

Here, we use a recently published study to illustrate three of the four approaches shown in Figure 1b that are used for quantifying priority effects and demonstrate how conclusions can vary depending on which approach and metric is used. This case study uses data from Ojima et al. (2022). This study was an in vitro experiment manipulating the order of arrival of four *Bifidobacterium* species (*B. bifidum*, *B. breve*, *B. infantis* and *B. longum*) found in the human infant gut. The experiments were conducted in a medium containing human milk oligosaccharides (HMOs), a group of complex sugars found in human breastmilk, as the sole carbon source. The selected strains differ in their ability to utilise HMOs, where *B. bifidum* and *B. infantis* are avid utilisers of all types, *B. longum* shows limited utilisation, and *B. breve* can only use a limited range and prefers monosaccharide and disaccharide degradant products of HMOs. In their experiment, Ojima et al. (2022) created microbial cultures in which two *Bifidobacterium* species were introduced sequentially or synchronously. When the bacteria were introduced sequentially, the second species was introduced 12 h after the first. When species were introduced synchronously, they were introduced at time point zero (i.e. at the same time as the arrival of the first species in non-synchronous sequences). All possible pairwise combinations of the four bacterial species were tested, representing a total of 18 arrival orders, six of which were synchronous. Each arrival scenario was replicated four times. Species abundance was monitored for 24 h, every 4 h from the start of the experiment, defined as the time when the first species were introduced. In this paper, the performance of each bacterial species was calculated by multiplying the optical density (600 nm) value measured at a specific time point (20 h after the start of the experiment), used as an estimate of total bacterial abundance, by the

20412108x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

relative abundance of a species estimated from qPCR data measured at the same time point.

We calculated priority effects using four metrics (Equations 1, 2, 3a and 4a) and three of the four approaches (1, 3 and 4 in Figure 1b). We were only able to calculate sampling variances for the log response ratio (Equation 3b) and Hedges' *g* (Equation 4c). Since the second species was introduced 12 h after the first, we quantified priority effects 16 and 20 h after the first introduction, corresponding to 4 and 8 h after the second introduction, respectively. After 20 h, the cultures had reached a stable physical growth plateau, but the abundances had not yet started to decline. After 24 h, many treatments showed a noticeable decline in copy numbers, which is likely caused by bacterial death in the late-stationary phase. Data were not available to calculate priority effects using a species-specific time interval and a reverse reference scenario (Approach 2 in Figure 1b), which is a common situation as this approach has rarely been implemented.

Data analysis was performed in R version 4.5.2 (R Core Team, 2025). Bias-corrected standardised mean differences (Hedges' *g*), uncorrected log response ratios and sampling variances were calculated using the *escalc* function of the metafor R package (Viechtbauer, 2010). By default, *escalc* calculates the bias correction of Hedges' *g* using the exact solution (Equation 4a). To compare priority effect values obtained using a different combination of methodological approach (Figure 1b) and metrics (Equations 1, 2, 3a and 4a), we assessed the strength of their association using Spearman's rank correlation coefficient. In addition, we modelled the relationship between methods using standardised major axis (SMA) regression implemented in the smatr R package (Warton et al., 2012), which is an appropriate choice when both variables are subject to measurement error. A limitation of this analysis is that some performance measurements used when calculating priority effects for different plotted observations originate from the same arrival sequence, sometimes taken at different time points. Consequently, some of the plotted pairs may not be fully independent. We considered fitting a linear mixed-effects model with unordered species pair as a random effect, but each comparison included only six unordered pairs, with two observations per pair. The SMA analysis retained here does not account for the potential dependence between priority effect estimates and its inferential results should therefore be interpreted cautiously.

For each fitted model, we calculated the 95% confidence intervals (CIs) for the slope and intercept. The fitted relationship was considered to differ significantly from the 1:1 line if the 95% CI of the slope did not include 1 and/or the 95% CI of the intercept did not include 0. When comparing two methodological approaches for quantifying priority effects, we concluded that they agreed with each other if three criteria were met simultaneously: (1) the Spearman's rank correlation coefficient was significantly different from zero, (2) the 95% CI of the SMA slope included 1, and (3) the 95% CI of the intercept included 0. In addition, the following R packages were used for data exploration and visualisation: here (Müller, 2025), readxl (Wickham & Bryan, 2026), tidyverse (Wickham et al., 2019), ggpubr (Kassambara, 2023), metR (Campitelli, 2021) and viridis (Garnier, 2024).

Data and R code are available on Zenodo (B. Delory et al., 2026).

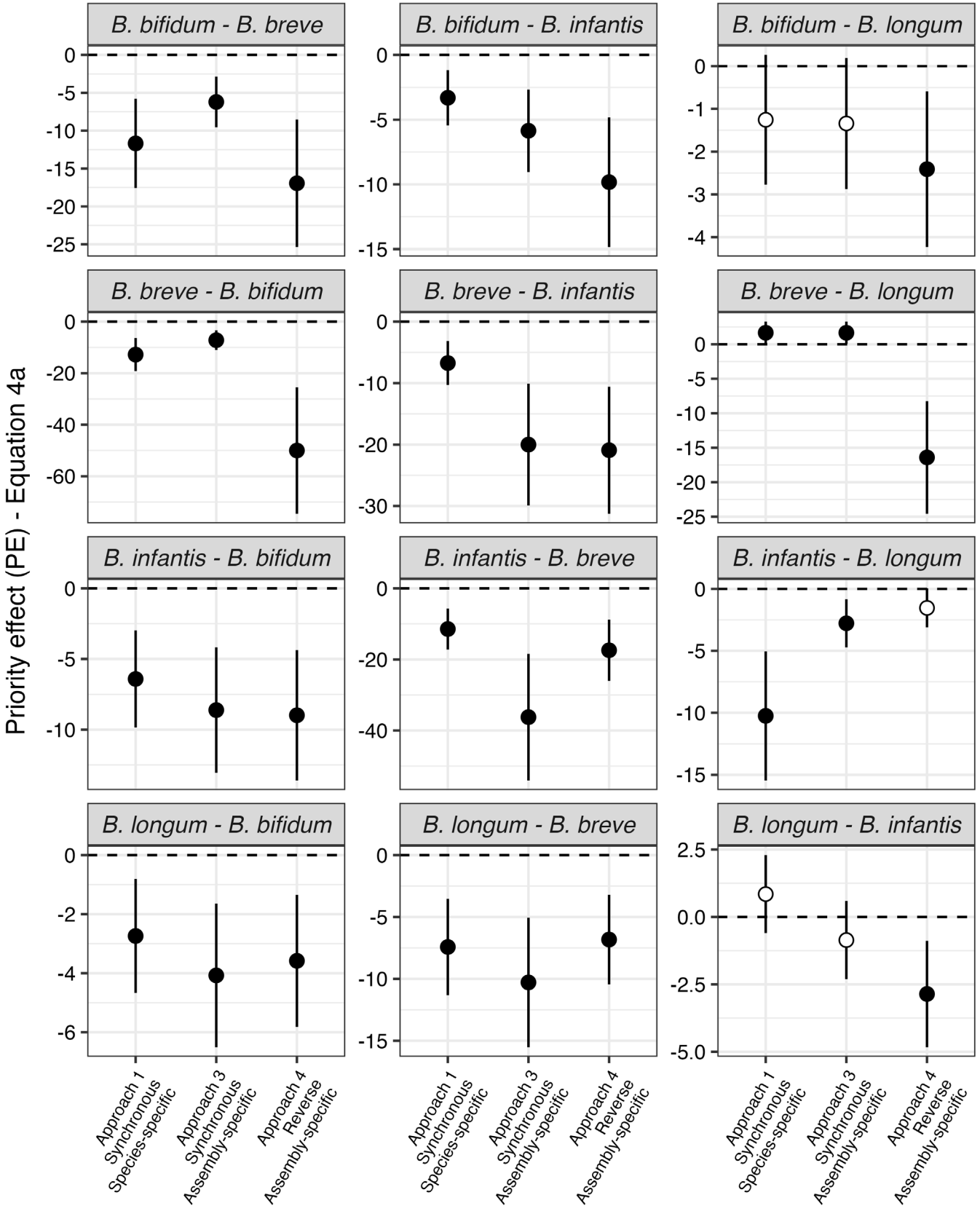
3 | RESULTS

3.1 | Different methodological approaches to quantifying priority effects can lead to contrasting conclusions about their strength and even their direction

In most cases, priority effects calculated 8 h after the introduction of the response species were negative, meaning that a *Bifidobacterium* species performed worse when it arrived late in comparison with a reference situation where it was introduced first or at the same time as the effect species in the community (Figure 3). However, in four arrival sequences including *B. longum* (Figure 3 right-hand panels), we found that approaches using a synchronous reference scenario (Approaches 1 and 3) gave different results than using a reverse reference scenario and an assembly-specific time interval (Approach 4). When *B. infantis* was introduced after *B. longum*, or when *B. longum* was introduced after *B. bifidum*, the priority effect was not significant when using a synchronous reference scenario (Approaches 1 and 3) but indicated a negative priority effect when quantified using a reverse reference scenario and an assembly-specific time interval (Approach 4). In contrast, for the sequence *B. infantis*–*B. longum*, we found an inhibitory priority effect using Approaches 1 and 3 but

FIGURE 3 Priority effects in two-species *Bifidobacterium* communities. In Approach 1, priority effects were calculated using species abundance data obtained 8 h after the introduction of the response (late-arriving) species. In Approaches 3 and 4, priority effects were calculated using the abundance of the target species measured 20 h after the start of the experiment (or 8 h after its introduction) in the target situation and the abundance of the same species measured after 20 h in the reference scenario. Each panel shows results for a specific species combination (effect species–response species). Priority effects were quantified using standardised mean differences (Equation 4a). We used two reference scenarios (Figure 1b): A reverse reference, in which the positions of the effect and response species are reversed while keeping all other factors constant (Approach 4) and a synchronous reference, in which both species are introduced simultaneously (Approaches 1 and 3). Using Approach 1, the time lag between the arrival of the response species and the measurement is kept constant in the target and reference arrival sequences (i.e. species-specific). In Approaches 3 and 4, it is the time lag between the first introduction event and the measurement that is kept constant in the target and reference arrival sequences (i.e. assembly-specific). Error bars are 95% confidence intervals (CIs). When 95% CI did not include zero, values are shown with a filled dot, indicating a statistically significant positive or negative priority effect. Values are otherwise shown with an empty dot. No data were available to calculate priority effects using Approach 2 (reverse reference scenario and species-specific time interval) described in Figure 1b.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



no priority effect using Approach 4. And in a final case (*B. breve*–*B. longum*), a facilitatory priority effect was detected using approaches with a synchronous reference scenario (Approaches 1 and 3), but an inhibitory priority effect using Approach 4. These results illustrate that the choice of reference scenario can change what outcomes are detected.

Under the synchronous reference scenario, differences between priority effects estimated with the species-specific

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(Approach 1) and assembly-specific (Approach 3) comparisons varied across species combinations, with each approach sometimes detecting stronger effects (Figure 3). This pattern shows how the choice of baseline can influence effect size estimation for the later-arriving species.

3.2 | Different combinations of metrics and methodological approaches can lead to different conclusions about the presence or absence of a priority effect

Applying different metrics (Equations 1, 2, 3a and 4a) to the same experimental data sometimes led to contrasting conclusions for specific arrival sequences (Figure 3; Figures S2–S4). As shown in Figures S2 and S3, metrics bound within finite limits (Equations 1 and 2) can produce extreme priority effect values when the response species performs substantially worse when arriving after the effect species than in the reference scenario ($Y_{Late} << Y_{Ref}$). In such cases, priority effect values approach the lower bound of these metrics (−1), reducing the ability to distinguish differences in priority effect strength among different approaches. In general, the log response ratio (Equation 3a) detected a greater number of priority effects that differed significantly from zero than the standardised mean difference (Equation 4a). For example, all three approaches identified an inhibitory priority effect of *B. bifidum* on *B. longum* when the log response ratio was used (Figure S4). However, when the standardised mean difference was used, only the approach based on a reverse reference scenario (Approach 4) detected this inhibitory effect, whereas the two approaches relying on a synchronous reference scenario (Approaches 1 and 3) did not (Figure 3). A similar pattern was observed for the effect of *B. infantis* on *B. longum*. All three approaches detected a strong inhibitory priority effect when quantified using the log response ratio (Figure S4), whereas only the approaches using a synchronous reference scenario (Approaches 1 and 3) detected a significant inhibitory effect when the standardised mean difference was used (Figure 3).

3.3 | Except for a few metrics, different approaches to quantifying priority effects are not in agreement

For each metric (Equations 1, 2, 3a and 4a), we assessed the agreement among different methodological approaches for quantifying priority effects (summarised in Figure 1b) at two successive time points (16 and 20 h after the start of the experiment).

After 20 h, and under a synchronous reference scenario, approaches based on species-specific (Approach 1) and assembly-specific (Approach 3) standardisation agreed with each other only when priority effects were quantified using an additive intensity index (Equation 1) or a normalised difference (Equation 2) (Figure 4a, Table 2). In contrast, when priority effects were quantified using a log response ratio (Equation 3a) or a standardised mean difference (Equation 4a), the strength of the association between approaches was lower for Equation 4a, and the approaches no longer agreed for Equations 3a and 4a (Figure 4a, Table 2). Under an assembly-specific standardisation, estimates of priority effects obtained using synchronous and reverse reference scenarios were not in agreement, regardless of the priority effect metric used (Figure 4b, Table 2). Our analysis also found no consistent agreement between Approaches 1 and 4 across all metrics (Table 2, Figure S5).

At 16 h, all relationships were weaker than at 20 h (Table 2, Figures S6 and S7). Under the synchronous reference scenario, approaches based on species-specific standardisation (Approach 1) and assembly-specific standardisation (Approach 3) agreed only when priority effects were quantified using either a normalised difference (Equation 2) or a log response ratio (Equation 3a) (Table 2, Figures S6 and S7).

Overall, these results indicate that the conclusions of experiments on priority effects depend on when and how priority effects are quantified, with the priority effect metric, the reference scenario (synchronous or reverse) and the standardisation method (species-specific or assembly-specific) exerting a major influence. In the following section, we present recommendations on the approaches to choose, taking into account research questions and logistical constraints.

4 | DISCUSSION

As demonstrated in our case study, different approaches and metrics used to quantify priority effects do not always yield consistent results and are best suited to addressing different types of questions about arrival history. Here, we connect approaches to research questions and offer recommendations on when and how to apply them (Figure 5).

When designing an experiment on priority effects that aims to evaluate the cost of arriving late, we recommend always including the reference scenario from Approach 1 (synchronous reference scenario, species-specific time interval) (Figure 5). In this approach, priority effects are quantified using a synchronous reference scenario in which all species are introduced at the same time and have the same time for growth as the response species in the target scenario (Alonso-Crespo et al., 2023; Delory, Weidlich, Kunz, et al., 2019; van Steijn et al., 2025). This approach is best suited for quantifying priority effects because the response species has the same amount of time to establish, develop and interact with other species in the target and reference arrival sequences, making them directly comparable. The use of a synchronous baseline is also advantageous when the experiment includes a large number of species whose order and/or timing of arrival needs to be manipulated (see Box 1), as it allows the number of arrival sequences to be kept logistically manageable, sometimes providing opportunities to test for factors that may affect the strength and direction of priority effects, such

20412108x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

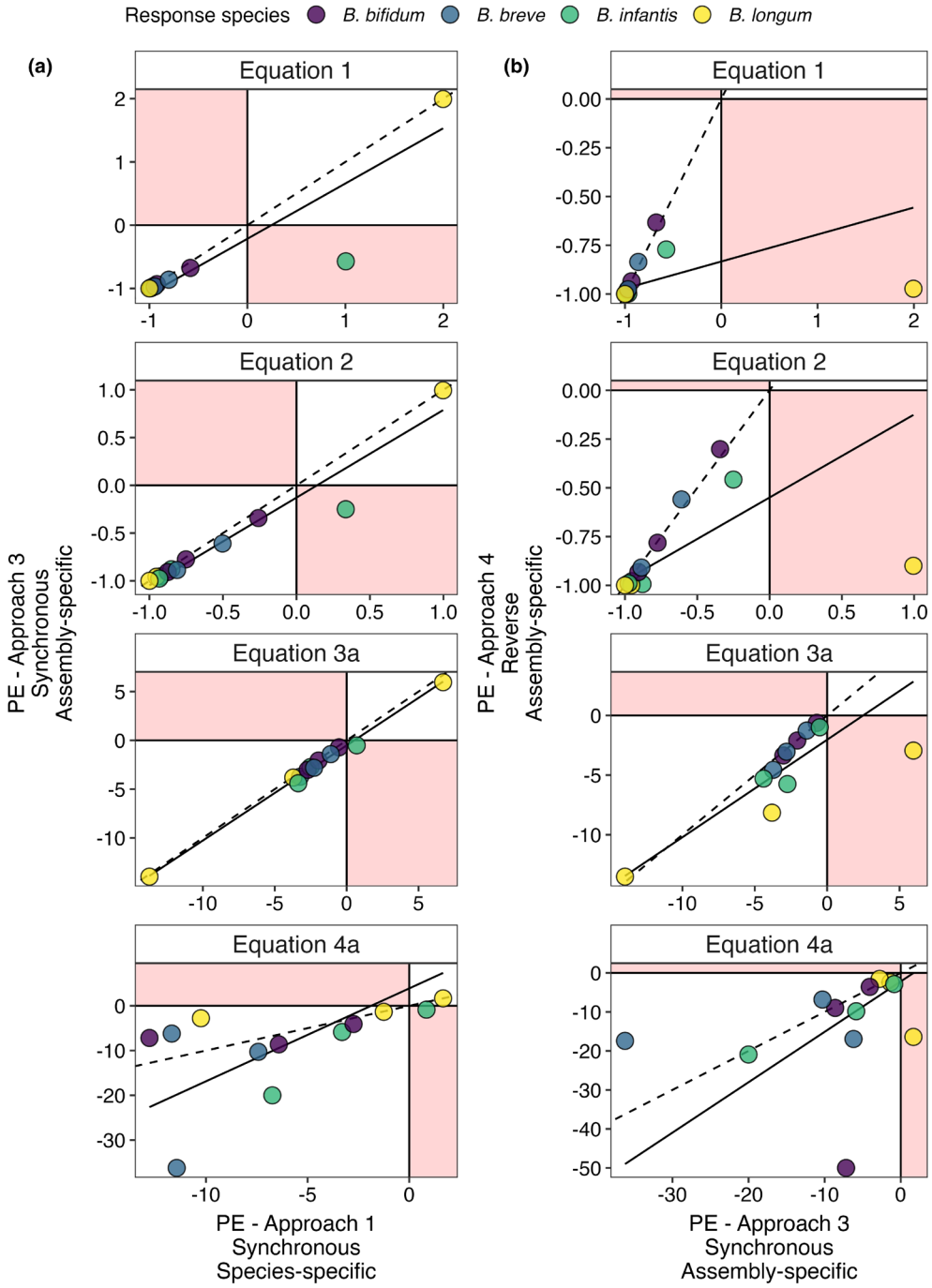


FIGURE 4 Pairwise relationships between priority effects (PE) in two-species *Bifidobacterium* communities measured using different approaches and metrics. For each response species, priority effects were calculated 20 h after the start of the experiment (i.e. 8 h after their introduction). Priority effects were quantified using four different metrics (Equations 1, 2, 3a and 4a). We used two reference scenarios (Figure 1b): A reverse reference, in which the positions of the early- and late-arriving species are reversed while keeping all other factors constant (Approach 4), and a synchronous reference, in which both species are introduced simultaneously (Approaches 1 and 3). Approaches 1, 3 and 4 further differ in the way they standardise time intervals between introduction events and performance measures. In Approach 1 (species-specific), the time lag between the arrival of the response species and the measurement is kept constant in the target and reference arrival sequences. In Approaches 3 and 4 (assembly-specific), the time lag between the first introduction event and the measurement is kept constant in the target and reference arrival sequences. Panel (a) compares species-specific Approach 1 with assembly-specific Approach 3, while panel (b) compares synchronous Approach 3 with reverse Approach 4. See also Figure S5 for a comparison between Approaches 1 and 4. In each panel, the dashed line represents the 1:1 line. The solid black lines are SMA regression models, built separately for each plot (Table 2). The red areas represent zones where two approaches do not agree on the direction of the priority effect.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Extent of agreement among different methodological approaches to quantifying priority effects.

Time since first introduction	Priority effect metric	Approaches (Figure 1b)		Correlation (spearman)	Linear regression		Agreement
		x	*y*		Slope (95% CI)	Intercept (95% CI)	
20 h (Figure 4)	Equation 1	1	3	0.99*	0.87 (0.63, 1.20)	−0.21 (−0.52, 0.09)	Yes
	Equation 2	1	3	0.99*	0.92 (0.77, 1.10)	−0.13 (−0.26, 0.01)	Yes
	Equation 3a	1	3	0.99*	0.97 (0.92, 1.03)	−0.52 (−0.78, −0.26)	No
	Equation 4a	1	3	0.68*	2.08 (1.18, 3.66)	3.89 (−6.16, 13.94)	No
	Equation 1	3	4	0.84*	0.14 (0.07, 0.27)	−0.83 (−0.96, −0.70)	No
	Equation 2	3	4	0.84*	0.43 (0.23, 0.80)	−0.55 (−0.81, −0.29)	No
	Equation 3a	3	4	0.84*	0.82 (0.54, 1.24)	−2.03 (−3.89, −0.18)	No
	Equation 4a	3	4	0.58	1.29 (0.68, 2.45)	−2.18 (−15.56, 11.21)	No
20 h (Figure S5)	Equation 1	1	4	0.88*	0.12 (0.06, 0.23)	−0.86 (−0.97, −0.76)	No
	Equation 2	1	4	0.88*	0.39 (0.22, 0.71)	−0.60 (−0.82, −0.39)	No
	Equation 3a	1	4	0.88*	0.80 (0.52, 1.21)	−2.46 (−4.22, −0.69)	No
	Equation 4a	1	4	0.48	2.68 (1.51, 4.76)	2.85 (−10.33, 16.04)	No
16 h (Figure S6)	Equation 1	1	3	0.59*	−1.05 (−2.02, −0.54)	−0.56 (−1.53, 0.41)	No
	Equation 2	1	3	0.59*	1.06 (0.55, 2.04)	−0.30 (−0.91, 0.30)	Yes
	Equation 3a	1	3	0.59*	0.98 (0.59, 1.61)	−1.09 (−3.26, 1.08)	Yes
	Equation 4a	1	3	0.27	0.87 (0.45, 1.67)	−5.18 (−10.93, 0.57)	No
	Equation 1	3	4	0.29	−0.14 (−0.26, −0.07)	−0.97 (−1.11, −0.84)	No
	Equation 2	3	4	0.29	0.43 (0.22, 0.82)	−0.59 (−0.89, −0.29)	No
	Equation 3a	3	4	0.29	1.03 (0.57, 1.87)	−2.07 (−5.21, 1.07)	No
	Equation 4a	3	4	0.37	2.76 (1.44, 5.29)	0.35 (−18.37, 19.06)	No
16 h (Figure S7)	Equation 1	1	4	0.84*	0.14 (0.10, 0.20)	−0.90 (−0.95, −0.85)	No
	Equation 2	1	4	0.84*	0.45 (0.30, 0.68)	−0.72 (−0.84, −0.60)	No
	Equation 3a	1	4	0.84*	1.01 (0.76, 1.34)	−3.20 (−4.35, −2.05)	No
	Equation 4a	1	4	0.29	−2.39 (−4.59, −1.24)	−18.30 (−33.68, −2.92)	No

Note: Spearman's rank correlation coefficients quantify the strength of the association between pairs of methods. Standardised major axis (SMA) regression coefficients are presented with their 95% confidence intervals (CIs) (lower, upper). Agreement between two approaches was inferred when (1) Spearman's rank correlation coefficient was significantly different from zero, (2) the 95% CI of the SMA slope included 1 and (3) the 95% CI of the intercept included 0, indicating no significant deviation from the 1:1 line. When at least one criterion was not met, the approaches were considered incongruent. Spearman correlation coefficients marked with an asterisk (*) differed significantly from zero (*$p < 0.05$).

as the time lag between species arrival, year effects, functional dissimilarity between interacting species, or environmental conditions (Alonso-Crespo et al., 2025; van Steijn et al., 2025; von Gillhaussen et al., 2014; Werner et al., 2020; Young et al., 2017).

To understand the effects of arrival order on species performance, it can be helpful to also take Approach 2 (reverse reference scenario, species-specific time interval) in addition to Approach 1 (Figure 5). To date, this approach has not been used much in the literature, which is perhaps surprising given that it seems the most logical scenario for quantifying a priority effect. Approach 2 differs from Approach 1 in that species are introduced sequentially in both the target and reference arrival sequences, with the response species having the same time to establish and develop in both sequences. An additional difference is that, in Approach 2, the species have less time to interact in the reference arrival sequence than in the target sequence. If the response species is facilitated by the effect species, Approach 2 is likely to yield an estimated value of positive priority effects that is higher than other approaches would, therefore overestimating the strength of positive priority effects.

In addition to quantifying the cost of arriving late for species R (i.e. what we refer to as a priority effect), researchers may also be interested in quantifying the benefit of arriving early for species E (Delory, Weidlich, Kunz, et al., 2019; van Steijn et al., 2025; Werner et al., 2016). In that case, adding Approach 3 (synchronous reference scenario, assembly-specific time interval) is a good strategy as it allows us to compare the performance of species E when it arrives early in the target sequence to its performance in a reference arrival sequence where all species are introduced simultaneously at the first introduction event (i.e. when the experiment starts). This comparison will allow us to assess the degree to which species E benefits from arriving before species R (Figure 5 but see also Figure S1).

Whether or not to also take Approach 4 (reverse reference scenario, assembly-specific time interval) depends on the research question. Given that species E and R do not have the same growth time in the target and reference arrival sequences, this approach may be the

20412108x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

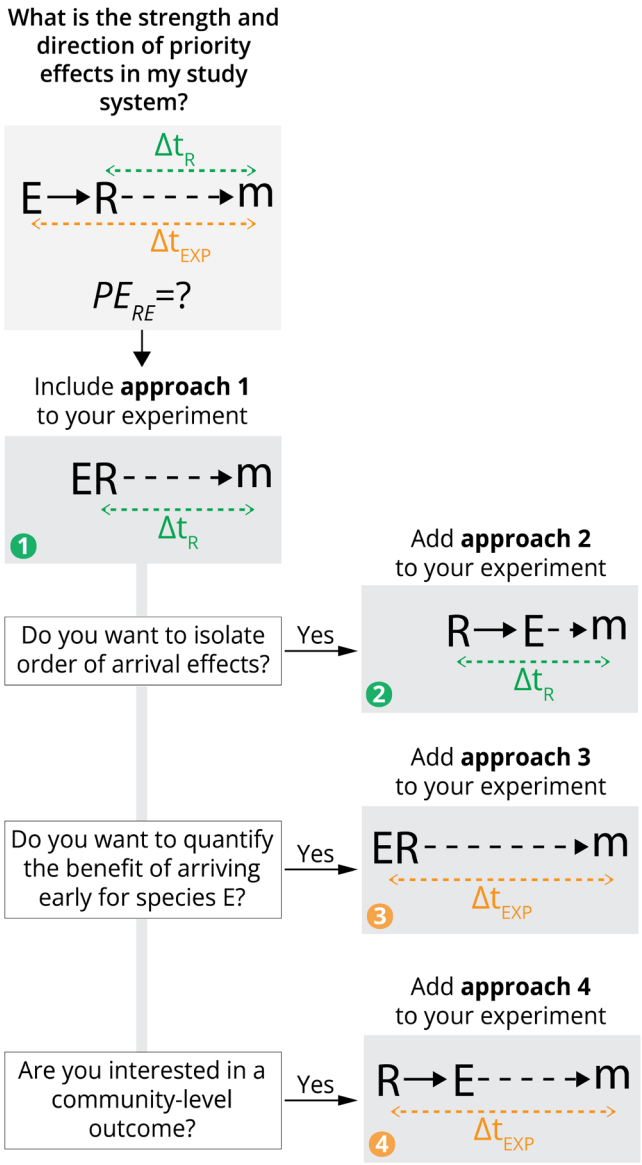


FIGURE 5 Recommendations for choosing the approach(es) for quantifying priority effects in community assembly experiments. The top grey box represents the target sequence, while grey boxes labelled 1–4 correspond to methodological approaches shown in Figure 1b. Here, E (effect) and R (response) represent two different species, functional groups or genotypes. The solid black horizontal arrows between E and R represent the time between the arrival of species E and the arrival of species R. The dashed black horizontal arrows represent the interaction time between species E and R. The letter m marks the time of performance measurement (e.g., abundance of E and R), which is recorded for both the target and reference arrival sequences. Δt_R is the time interval between the arrival of the response species (R) and a measurement. Δt_{EXP} is the time interval between the arrival of the first species (E) and a measurement. PE_{RE} is the priority effect that species E has on species R.

least suitable for quantifying the cost of a late arrival or the benefit of an early arrival for a particular species. If we used this approach to quantify the priority effect of species E on species R, we would likely overestimate negative priority effects given that species R would

BOX 1 Quantifying priority effects for arrival sequences with more than two introduction events

When arrival sequences have more than two introduction events, it becomes more difficult to find a reference scenario to calculate an effect size quantifying the strength of a priority effect. This is especially the case when a reverse reference is used. To illustrate this, consider an ecological community consisting of four species (A, B, C and D) interacting with each other. In this hypothetical community, all four species were introduced sequentially, starting with A, then B, then C and finally D. In this example, how can we calculate the cost of arriving late (priority effect) for species D?

To answer this question, we first need to define a reference arrival sequence for which we can measure the abundance of species D and compare it with the target arrival sequence. Although it is straightforward to define a synchronous reference situation where all four species are introduced simultaneously (ABCD), it is more challenging to define a reverse arrival sequence for this purpose. One way to obtain a reverse reference arrival sequence is simply to invert the positions of the species for which we want to measure a priority effect (species D in our example) with that of the species that was first introduced into the community, all other things being equal. If the target arrival sequence is A-B-C-D, this means that the reverse reference situation would be D-B-C-A. However, this is not the only possible option. Another possibility would be to use the average abundance value of species D in all six possible arrival sequences in which it is introduced first (D-A-B-C, D-A-C-B, D-B-A-C, D-B-C-A, D-C-A-B and D-C-B-A), and use this value to calculate the priority effect on species D in the target arrival sequence (A-B-C-D). While this approach would work well in modelling studies, it would be difficult to implement in an experiment, as the number of arrival sequences that would need to be tested becomes difficult to manage in experiments with multiple species of interest and with a high number of introduction events. Once a reference arrival sequence has been chosen, an effect size (e.g., standardised mean difference) can then be calculated using one of the two calculation scenarios described earlier (species-specific vs. assembly-specific).

When priority effects need to be quantified in communities with multiple introduction events, we recommend using a synchronous reference situation similar to Approach 1 in Figure 1 as it is logistically simpler and would make it easier to compare results across priority effect studies. However, in interpreting the results of such a comparison, one should keep in mind the implications of a synchronous reference scenario and how the strength of priority effects detected may differ from the use of a reverse reference scenario.

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

have more time to develop and grow in the reference than in the target situation, while it would be the opposite for species E. Our case study illustrated this point as, in some arrival sequences, we found that using Approach 4 led to stronger inhibitory priority effects than other approaches, which is likely due to contrasting differences in growth time between the reference and target sequences. However, given that the interaction time is identical in the two sequences, Approach 4 may still be useful if one wishes to study how the order of arrival of species affects community- and ecosystem-level properties, such as biomass production, nutrient cycling, and matter and energy fluxes (Fukami et al., 2010; Weidlich et al., 2016).

Ultimately, the choice of approach for quantifying priority effects should be guided by considerations specific to the study system. For the *B. longum*–*B. infantis* and *B. bifidum*–*B. longum* sequences, the approach using a reverse reference (Approach 4) yielded stronger inhibitory priority effects than those using a synchronous reference (Approaches 1 and 3). These differences are consistent with bacterial substrate preferences, resource utilisation dynamics and competitive interactions. *B. bifidum*, *B. infantis* and *B. longum* compete when intact oligosaccharides are the main resource, but *B. bifidum* and *B. infantis* utilise them more efficiently, giving them a competitive advantage. In the *B. bifidum*–*B. longum* sequence, *B. longum* performs poorly when it arrives after *B. bifidum* and when both species arrive simultaneously. Consequently, the synchronous reference provides little contrast with the target sequence, resulting in weak or no priority effects. In the reverse reference, however, *B. longum* arrives first and has time to use resources before *B. bifidum* arrives. Its improved performance in this scenario increases the contrast with the target sequence, resulting in a stronger estimated inhibitory priority effect. A similar influence of arrival order and resource use may explain the results for the *B. longum*–*B. infantis* sequence.

Priority effect experiments do not necessarily need to include all the approaches described in this paper. The choice of approach will depend on the research question(s) (Figure 5), but also on logistical constraints, particularly when the species pool is large and/or when the experiment involves more than two introduction events (but see Box 1 for a discussion on how approaches described in this paper can be extended to situations with multiple introduction events).

The time interval between species arrivals is a key determinant of the strength and direction of priority effects because it affects the scope for the effect species to establish, grow and modify environmental conditions before the response species arrives (Kardol et al., 2013). When designing priority effect experiments, time intervals should ideally be selected based on a good understanding of the biology of the organisms involved and the variation in arrival histories observed in the field. Ideally, priority effect experiments should incorporate and test multiple ecologically relevant arrival intervals and determine how they affect the strength and direction of priority effects.

Where possible, it also helps to measure the performance of species at several time points during a priority effect experiment, as this repeated measurement would allow use of multiple approaches (including Approach 2) and would also allow empiricists to assess whether or not their study system is at equilibrium (see Box 2 in Grainger et al., 2025). Repeated measurements can be difficult when they take destructive sampling, which requires high replication, but their benefit can still outweigh the cost in some cases. Temporal data can be used to analyse when priority effects are weakest or strongest, as well as how long they persist in a study system, which has not been studied much to date. Investigating how priority effects change over time may provide insights into the ecological mechanisms underlying them.

When choosing a metric for quantifying priority effects, we do not recommend the additive intensity index (Equation 1) or the normalised difference (Equation 2). When Y_{Late} is much smaller than Y_{Ref}, both metrics rapidly approach their lower bound of −1, compressing differences in effect magnitude. In addition, for the formulations considered here, we do not provide generally applicable sampling-variance estimators or small-sample bias corrections for these indices, which limits their usefulness for quantitative synthesis. We also do not recommend the bias-corrected log response ratio because the correction can reverse the direction of the estimated priority effect in some cases (see Equation 3c in Figure 2). We recommend Hedges' *g* (Equation 4a) when the aim is to express differences between groups in units of pooled standard deviation, particularly when comparing outcomes measured on different scales. However, Hedges' *g* is sensitive to within-group variability and can produce large values when the variability is low. It should therefore be interpreted alongside the underlying group-level data and its sampling variance. When the relative difference between group means is of primary interest, the uncorrected log response ratio (Equation 3a) is an alternative, provided both means are strictly positive.

AUTHOR CONTRIBUTIONS

Benjamin M. Delory and Tadashi Fukami led the iDiv working group that resulted in this work, with support from Vicky M. Temperton. Benjamin M. Delory led data analysis and manuscript writing, with Chelsea J. Little, Hannah Marton, Fletcher W. Halliday, Elle M. Barnes and Tadashi Fukami. All authors discussed ideas and contributed to manuscript writing and editing.

AFFILIATIONS

[1]Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, The Netherlands; [2]School of Environmental Science, Simon Fraser University, Vancouver, British Columbia, Canada; [3]Department of Biological Sciences, Simon Fraser University, Vancouver, British Columbia, Canada; [4]Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon, USA; [5]Thomas H. Gosnell School of Life Sciences, Rochester Institute of Technology, Rochester, New York, USA; [6]Department of Ecology, Environment and Geoscience, Umeå University, Umeå, Sweden; [7]Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada; [8]Institute of Biology, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany; [9]Department of Forest Mycology and Plant Pathology, Swedish University of Agricultural

20412108x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Sciences, Uppsala, Sweden; [10]Department of Species Interaction Ecology, Helmholtz Centre for Environmental Research-UFZ, Leipzig, Germany; [11]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany; [12]Department of Science and Conservation, National Tropical Botanical Garden, Kalāheo, Hawaii, USA; [13]Department of Biology, University of Prince Edward Island, Charlottetown, Prince Edward Island, Canada; [14]Canadian Centre for Climate Change and Adaptation, University of Prince Edward Island, St. Peter's Bay, Prince Edward Island, Canada; [15]School of Climate Change and Adaptation, University of Prince Edward Island, Charlottetown, Prince Edward Island, Canada; [16]Institute for the Evolutionary Origins of Human Behavior, Kyoto University, Kyoto, Japan; [17]Department of Physiological Diversity, Helmholtz Centre for Environmental Research-UFZ, Leipzig, Germany; [18]School of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA; [19]Institute of Ecology, School of Sustainability, Leuphana University Lüneburg, Lüneburg, Germany; [20]Department of Environmental Science, Policy & Sustainability, Southern Oregon University, Ashland, Oregon, USA and [21]Departments of Biology and of Earth System Science, Stanford University, Stanford, California, USA

ACKNOWLEDGEMENTS

This paper is a joint effort of the working group sPriority supported in part by sDiv, the Synthesis Centre of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, funded by the German Research Foundation (FZT 118, 202548816). Judith M. Sarneel acknowledges support from the Swedish Research Council Vetenskapsrådet (grant number 2019-05099). Many thanks to Ton Markus (Utrecht University) for his help in designing some of the figures. The authors thank three anonymous reviewers for their constructive comments and suggestions, as well as Elmer for his unfailing support, playfulness, relaxed attitude and occasional encouragement during some of the discussions that led to the writing of this paper.

CONFLICT OF INTEREST STATEMENT

Employment of Miriam N. Ojima at Kyoto University was supported by Morinaga Milk Industry Co. Ltd. (April 2021 to September 2025).

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210x.70449.

DATA AVAILABILITY STATEMENT

Data and R code are available via https://doi.org/10.5281/zenodo.23801867 (Delory et al., 2026).

ORCID

Benjamin M. Delory https://orcid.org/0000-0002-1190-8060
Chelsea J. Little https://orcid.org/0000-0003-2803-7465
Judith M. Sarneel https://orcid.org/0000-0001-6187-499X
Julia Dieskau https://orcid.org/0000-0002-2410-415X
Paul Kardol https://orcid.org/0000-0001-7065-3435
Emma Ladouceur https://orcid.org/0000-0002-4943-4358
Christiane Roscher https://orcid.org/0000-0001-9301-7909
James T. Stroud https://orcid.org/0000-0003-0734-6795
Vicky M. Temperton https://orcid.org/0000-0003-0543-4521

REFERENCES

Alonso-Crespo, I. M., Temperton, V. M., Fichtner, A., Niemeyer, T., Schloter, M., & Delory, B. M. (2025). Exploring priority and year effects on plant diversity, productivity and vertical root distribution: First insights from a grassland field experiment. *Journal of Vegetation Science: Official Organ of the International Association for Vegetation Science*, *36*(2), e70026.

Alonso-Crespo, I. M., Weidlich, E. W. A., Temperton, V. M., & Delory, B. M. (2023). Assembly history modulates vertical root distribution in a grassland experiment. *Oikos*, *1*, e08886.

Campitelli, E. (2021). metR: tools for easier analysis of meteorological fields (Version 0.18.3). https://eliocamp.github.io/metR/

Carpenter, S. A., Vannatta, J. T., & Minchella, D. J. (2021). Host exposure history and priority effects impact the development and reproduction of a dominant parasite. *International Journal for Parasitology*, *51*(11), 935–943.

Chappell, C. R., Dhami, M. K., Bitter, M. C., Czech, L., Herrera Paredes, S., Barrie, F. B., Calderón, Y., Eritano, K., Golden, L.-A., Hekmat-Scafe, D., Hsu, V., Kieschnick, C., Malladi, S., Rush, N., & Fukami, T. (2022). Wide-ranging consequences of priority effects governed by an overarching factor. *eLife*, *11*, e79647. https://doi.org/10.7554/eLife.79647

Chase, J. M. (2003). Community assembly: When should history matter? *Oecologia*, *136*(4), 489–498.

Debray, R., Herbert, R. A., Jaffe, A. L., Crits-Christoph, A., Power, M. E., & Koskella, B. (2022). Priority effects in microbiome assembly. *Nature Reviews Microbiology*, *20*(2), 109–121.

Delory, B., Little, C. J., Marton, H., Halliday, F., Barnes, E., Isabelle, S., Sarneel, J., Van Steijn, T., Grainger, T., Dieskau, J., Kardol, P., Knight, T., Ladouceur, E., Ojima, M., Roscher, C., Stroud, J., Temperton, V., Werner, C., & Fukami, T. (2026). Data and R code used in Delory et al - Quantifying priority effects in community assembly experiments. *Zenodo*. https://doi.org/10.5281/ZENODO.23801867

Delory, B. M., Weidlich, E. W. A., Kunz, M., Neitzel, J., & Temperton, V. M. (2019). The exotic species *Senecio inaequidens* pays the price for arriving late in temperate European grassland communities. *Oecologia*, *191*(3), 657–671.

Delory, B. M., Weidlich, E. W. A., von Gillhaussen, P., & Temperton, V. M. (2019). When history matters: The overlooked role of priority effects in grassland overyielding. *Functional Ecology*, *33*(12), 2369–2380.

Díaz-Sierra, R., Verwijmeren, M., Rietkerk, M., de Dios, V. R., & Baudena, M. (2017). A new family of standardized and symmetric indices for measuring the intensity and importance of plant neighbour effects. *Methods in Ecology and Evolution*, *8*, 580–591.

Dieskau, J., Hensen, I., Eisenhauer, N., Gaberle, I., Durka, W., Lachmuth, S., & Auge, H. (2024). Phylogenetic relationships and plant life stage but not biogeographic history mediate priority effects of European grassland plants. *Journal of Ecology*, *112*(9), 2007–2017.

Drake, J. A. (1991). Community-assembly mechanics and the structure of an experimental species ensemble. *The American Naturalist*, *137*(1), 1–26.

Fragata, I., Costa-Pereira, R., Kozak, M., Majer, A., Godoy, O., & Magalhães, S. (2022). Specific sequence of arrival promotes coexistence via spatial niche pre-emption by the weak competitor. *Ecology Letters*, *25*(7), 1629–1639.

Fukami, T., Dickie, I. A., Wilkie, J. P., Paulus, B. C., Park, D., Roberts, A., Buchanan, P. K., & Allen, R. B. (2010). Assembly history dictates

20412108x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ecosystem functioning: Evidence from wood decomposer communities: Carbon dynamics and fungal community assembly. *Ecology Letters*, *13*(6), 675–684.

Fukami, T., & Nakajima, M. (2011). Community assembly: Alternative stable states or alternative transient states? *Ecology Letters*, *14*(10), 973–984.

Fukami, T., & Nakajima, M. (2013). Complex plant-soil interactions enhance plant species diversity by delaying community convergence. *Journal of Ecology*, *101*(2), 316–324.

Garnier, S. (2024). viridis(Lite) – Colorblind-Friendly Color Maps for R. R package version 0.6.5. https://cran.r-project.org/package=viridis

Grainger, T. N., Letten, A. D., Gilbert, B., & Fukami, T. (2019). Applying modern coexistence theory to priority effects. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(13), 6205–6210.

Grainger, T. N., Stark, K., Song, C., Barbour, M. A., & Germain, R. M. (2025). The equilibrium conundrum. *Ecology Letters*, *28*(11), e70232.

Green, E. T., Grunberg, R. L., & Mitchell, C. E. (2024). Order of arrival and nutrient supply alter outcomes of co-infection with two fungal pathogens. *Proceedings. Biological sciences*, *291*(2029), 20240915.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Elsevier S & T. http://www.sciencedirect.com:5070/book/monograph/9780080570655/statistical-methods-for-meta-analysis

Hoverman, J. T., Hoye, B. J., & Johnson, P. T. J. (2013). Does timing matter? How priority effects influence the outcome of parasite interactions within hosts. *Oecologia*, *173*(4), 1471–1480.

Kardol, P., Souza, L., & Classen, A. T. (2013). Resource availability mediates the importance of priority effects in plant community assembly and ecosystem function. *Oikos*, *122*(1), 84–94.

Kassambara, A. (2023). ggpubr: "ggplot2" based publication ready plots. R package version 0.6.0. https://cran.r-project.org/package=ggpubr

Ke, P.-J., & Letten, A. D. (2018). Coexistence theory and the frequency-dependence of priority effects. *Nature Ecology & Evolution*, *2*, 1691–1695.

Koricheva, J., Gurevitch, J., & Mengersen, K. (Eds.). (2013). *Handbook of meta-analysis in ecology and evolution*. Princeton University Press. https://doi.org/10.1515/9781400846184

Lajeunesse, M. J. (2015). Bias and correction for the log response ratio in ecological meta-analysis. *Ecology*, *96*(8), 2056–2063.

Moore, J. E., & Franklin, S. B. (2012). Water stress interacts with early arrival to influence interspecific and intraspecific priority competition: A test using a greenhouse study. *Journal of Vegetation Science: Official Organ of the International Association for Vegetation Science*, *23*(4), 647–656.

Müller, K. (2025). here: a simpler way to find your files (Version 1.0.2). https://CRAN.R-project.org/package=here

Ojima, M. N., Jiang, L., Arzamasov, A. A., Yoshida, K., Odamaki, T., Xiao, J., Nakajima, A., Kitaoka, M., Hirose, J., Urashima, T., Katoh, T., Gotoh, A., van Sinderen, D., Rodionov, D. A., Osterman, A. L., Sakanaka, M., & Katayama, T. (2022). Priority effects shape the structure of infant-type *Bifidobacterium* communities on human milk oligosaccharides. *The ISME Journal*, *16*(9), 2265–2279.

Palmgren, A. (1926). Chance as an element in plant geography. In *Proceedings of the International Congress of Plant Sciences*, Ithaca, New York (Vol. 1, pp. 591–602).

Price, J. E., & Morin, P. J. (2009). Community convergence in a simple microbial food web. *Ecological Research*, *24*(3), 587–595.

R Core Team. (2025). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.r-project.org/

Rasmussen, N. L., Van Allen, B. G., & Rudolf, V. H. W. (2014). Linking phenological shifts to species interactions through size-mediated priority effects. *The Journal of Animal Ecology*, *83*(5), 1206–1215.

Sarneel, J. M., Kardol, P., Nilsson, C., & Bartha, S. (2016). The importance of priority effects for riparian plant community dynamics. *Journal of Vegetation Science: Official Organ of the International Association for Vegetation Science*, *27*(4), 658–667.

Song, C., Fukami, T., & Saavedra, S. (2021). Untangling the complexity of priority effects in multispecies communities. *Ecology Letters*, *24*(11), 2301–2313.

Stevens, J. M., & Fehmi, J. S. (2011). Early establishment of a native grass reduces the competitive effect of a non-native grass. *Restoration Ecology*, *19*(3), 399–406.

Stroud, J. T., Delory, B. M., Barnes, E. M., Chase, J. M., De Meester, L., Dieskau, J., Grainger, T. N., Halliday, F. W., Kardol, P., Knight, T. M., Ladouceur, E., Little, C. J., Roscher, C., Sarneel, J. M., Temperton, V. M., van Steijn, T. L. H., Werner, C. M., Wood, C. W., & Fukami, T. (2024). Priority effects transcend scales and disciplines in biology. *Trends in Ecology & Evolution*, *39*(7), 677–688.

Stuble, K. L., & Souza, L. (2016). Priority effects: Natives, but not exotics, pay to arrive late. *The Journal of Ecology*, *104*, 987–993.

Stuble, K. L., & Young, T. P. (2020). Priority treatment leaves grassland restoration vulnerable to invasion. *Diversity*, *12*(2), 71.

Sutherland, J. P. (1974). Multiple stable points in natural communities. *The American Naturalist*, *108*(964), 859–873.

Svoboda, P., Lindström, E. S., Ahmed Osman, O., & Langenheder, S. (2018). Dispersal timing determines the importance of priority effects in bacterial communities. *The ISME Journal*, *12*(2), 644–646.

Torres, A., Morán-López, T., Rodriguez-Cabal, M. A., & Núñez, M. A. (2024). Inverse priority effects: The order and timing of removal of invasive species influence community reassembly. *The Journal of Applied Ecology*, *61*(1), 51–62.

Tucker, C. M., & Fukami, T. (2014). Environmental variability counteracts priority effects to facilitate species coexistence: Evidence from nectar microbes. *Proceedings of the Royal Society B*, *281*(1778), 20132637.

van Steijn, T. L. H., Kardol, P., Jansson, R., & Sarneel, J. M. (2026). Germination speed modulates priority effects: Evidence from a large-scale field study. *Ecology*, *107*(1), e70291.

van Steijn, T. L. H., Kardol, P., Jansson, R., Tjäder, J., & Sarneel, J. M. (2025). Priority effects can be explained by competitive traits. *Ecology*, *106*(1), e4528.

Vannette, R. L., & Fukami, T. (2014). Historical contingency in species interactions: Towards niche-based predictions. *Ecology Letters*, *17*(1), 115–124.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

von Gillhaussen, P., Rascher, U., Jablonowski, N. D., Plückers, C., Beierkuhnlein, C., & Temperton, V. M. (2014). Priority effects of time of arrival of plant functional groups override sowing interval or density effects: A grassland experiment. *PLoS One*, *9*(1), e86906.

Warton, D. I., Duursma, R. A., Falster, D. S., & Taskinen, S. (2012). smatr 3– an R package for estimation and inference about allometric lines. *Methods in Ecology and Evolution*, *3*(2), 257–259. https://doi.org/10.1111/j.2041-210x.2011.00153.x

Weidlich, E. W. A., von Gillhaussen, P., Delory, B. M., Blossfeld, S., Poorter, H., & Temperton, V. M. (2016). The importance of being first: Exploring priority and diversity effects in a grassland field experiment. *Frontiers in Plant Science*, *7*(2008), 2008.

Weidlich, E. W. A., von Gillhaussen, P., Max, J. F. J., Delory, B. M., Jablonowski, N. D., Rascher, U., & Temperton, V. M. (2018). Priority effects caused by plant order of arrival affect below-ground productivity. *The Journal of Ecology*, *106*(2), 774–780.

Werner, C. M., Stuble, K. L., Groves, A. M., & Young, T. P. (2020). Year effects: Interannual variation as a driver of community assembly dynamics. *Ecology*, *101*(9), e03104.

Werner, C. M., Vaughn, K. J., Stuble, K. L., Wolf, K., & Young, T. P. (2016). Persistent asymmetrical priority effects in a California grassland

20412108x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

restoration experiment. *Ecological Applications: A Publication of the Ecological Society of America*, *26*(6), 1624–1632.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *The Journal of Open Source Software*, *4*(43), 1686.

Wickham, H., & Bryan, J. (2026). readxl: Read excel files (Version 1.5.0). https://CRAN.R-project.org/package=readxl

Wilsey, B. (2021). Restoration in the face of changing climate: Importance of persistence, priority effects, and species diversity. *Restoration Ecology*, *29*, e13132.

Young, T. P., Stuble, K. L., Balachowski, J. A., & Werner, C. M. (2017). Using priority effects to manipulate competitive relationships in restoration. *Restoration Ecology*, *25*, S114–S123.

Zou, H.-X., & Rudolf, V. H. W. (2023). Bridging theory and experiments of priority effects. *Trends in Ecology & Evolution*, *38*(12), 1203–1216.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Approaches for quantifying the cost/benefit of arriving early. Panel (a) shows the target arrival sequence for which we want to quantify the cost/benefit of arriving early for species E. Panel (b) illustrates possible approaches to quantify the cost/benefit of arriving early using data collected in a community assembly experiment. In this figure, E and R represent the introduction timing of two different species (effect and response species), functional groups or genotypes. The solid black horizontal arrows represent the time between the arrival of the effect species (E) and the arrival of the response species (R). The dashed black horizontal arrows represent the interaction time between effect and response species. The letter m marks the time of performance measurement (e.g., abundance of E and R), which is recorded for both the target and reference arrival sequences. To quantify the cost/benefit of arriving early for species E, two questions must be answered: (1) What reference arrival sequence should be used? and (2) How should the time between arrival and measurement be standardised? These choices define three approaches (numbered from 1 to 3 in the figure above). One of these approaches use a synchronous reference situation, while the other two use a reverse reference situation. There is only one synchronous reference situation because, in the example shown in the figure, the time from species E's arrival to measurement (ΔtE) is equal to the duration of the experiment (referred to as Δt_{EXP} in Figure 1), which results in identical reference situations for the species-specific and assembly-specific approaches relying on a synchronous reference. In the lower left panel, the time from species E's arrival to measurement (ΔtE) is held constant in the target and reference sequences. This species-specific time interval ensures that the effect species spends the same time in the environment across arrival scenarios. In the lower right panel, the total time from the first species' arrival to measurement is fixed. This assembly-specific standardisation ensures that the duration of community assembly is identical in the target and reference scenarios being compared.

Figure S2. Priority effects (PE) in two-species *Bifidobacterium* communities. In Approach 1, priority effects were calculated using species abundance data obtained 8 h after the introduction of the response (late-arriving) species. In Approaches 3 and 4 (Figure 1b), priority effects were calculated using the abundance of the target species measured 20 h after the start of the experiment (or 8 h after its introduction) in the target situation and the abundance of the same species measured after 20 h in the reference scenario. Each panel shows results for a specific species combination (effect species –response species). Priority effects were quantified using an additive intensity index (Equation 1 in the main text). We used two different methods for standardising arrival sequences (Figure 1): a reverse reference, in which the positions of the effect and response species are reversed while keeping all other factors constant (Approach 4), and a synchronous reference, in which both species are introduced simultaneously (Approaches 1 and 3). We also used two different approaches to standardise time intervals between introduction events and performance measures. In the first approach (species-specific, Approach 1), the time lag between the arrival of the second species and the measurement is kept constant in the target and reference arrival sequences. In the second approach (assembly-specific, Approaches 3 and 4), it is the time lag between the first introduction event and the measurement that is kept constant in the target and reference arrival sequences. No data were available to calculate priority effects using Approach 2 (reverse reference scenario and species-specific time interval) described in Figure 1.

Figure S3. Priority effects (PE) in two-species *Bifidobacterium* communities. In Approach 1, priority effects were calculated using species abundance data obtained 8 h after the introduction of the response (late-arriving) species. In Approaches 3 and 4 (Figure 1b), priority effects were calculated using the abundance of the target species measured 20 h after the start of the experiment (or 8 h after its introduction) in the target situation and the abundance of the same species measured after 20 h in the reference scenario. Each panel shows results for a specific species combination (effect species–response species). Priority effects were quantified using a normalised difference (Equation 2 in the main text). We used two different methods for standardising arrival sequences (Figure 1): a reverse reference, in which the positions of the effect and response species are reversed while keeping all other factors constant (Approach 4), and a synchronous reference, in which both species are introduced simultaneously (Approaches 1 and 3). We also used two different approaches to standardise time intervals between introduction events and performance measures. In the first approach (species-specific, Approach 1), the time lag between the arrival of the second species and the measurement is kept constant in the target and reference arrival sequences. In the second approach (assembly-specific, Approaches 3 and 4), it is the time lag between the first introduction event and the measurement that is kept constant in the target and reference arrival sequences. No data were available to calculate priority effects using Approach 2 (reverse reference scenario and species-specific time interval) described in Figure 1.

Figure S4. Priority effects (PE) in two-species *Bifidobacterium* communities. In Approach 1, priority effects were calculated using

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

species abundance data obtained 8 h after the introduction of the response (late-arriving) species. In Approaches 3 and 4 (Figure 1b), priority effects were calculated using the abundance of the target species measured 20 h after the start of the experiment (or 8 h after its introduction) in the target situation and the abundance of the same species measured after 20 h in the reference scenario. Each panel shows results for a specific species combination (effect species–response species). Priority effects were quantified using a log response ratio (Equation 3a in the main text). We used two different methods for standardising arrival sequences (Figure 1): a reverse reference, in which the positions of the effect and response species are reversed while keeping all other factors constant (Approach 4), and a synchronous reference, in which both species are introduced simultaneously (Approaches 1 and 3). We also used two different approaches to standardise time intervals between introduction events and performance measures. In the first approach (species-specific, Approach 1), the time lag between the arrival of the second species and the measurement is kept constant in the target and reference arrival sequences. In the second approach (assembly-specific, Approaches 3 and 4), it is the time lag between the first introduction event and the measurement that is kept constant in the target and reference arrival sequences. Error bars are 95% confidence intervals. When 95% CI did not include zero, values are shown with a filled dot. Values are otherwise shown with an empty dot. No data were available to calculate priority effects using Approach 2 (reverse reference scenario and species-specific time interval) described in Figure 1.

Figure S5. Pairwise relationships between priority effects (PE) in two-species *Bifidobacterium* communities measured using different approaches and metrics. For each response species, priority effects were calculated 20 h after the start of the experiment (i.e. 8 h after their introduction). Priority effects were quantified using four different metrics (Equations 1, 2, 3a and 4a). We used two reference scenarios (Figure 1b): a reverse reference, in which the positions of the early- and late-arriving species are reversed while keeping all other factors constant (Approach 4), and a synchronous reference, in which both species are introduced simultaneously (Approach 1). Approaches 1 and 4 further differ in the way they standardise time intervals between introduction events and performance measures. In Approach 1 (species-specific), the time lag between the arrival of the response species and the measurement is kept constant in the target and reference arrival sequences. In Approach 4 (assembly-specific), the time lag between the first introduction event and the measurement is kept constant in the target and reference arrival sequences. In each panel, the dashed line represents the 1:1 line. The solid black lines are SMA regression models, built separately for each plot (Table 2). The red areas represent zones where two approaches do not agree on the direction of the priority effect.

Figure S6. Pairwise relationships between priority effects (PE) in two-species *Bifidobacterium* communities measured using different approaches and metrics. For each response species, priority effects were calculated 16 h after the start of the experiment (i.e. 4 h after their introduction). Priority effects were quantified using four different metrics (Equations 1, 2, 3a and 4a). We used two reference scenarios (Figure 1b): a reverse reference, in which the positions of the early- and late-arriving species are reversed while keeping all other factors constant (Approach 4), and a synchronous reference, in which both species are introduced simultaneously (Approaches 1 and 3). Approaches 1, 3 and 4 further differ in the way they standardise time intervals between introduction events and performance measures. In Approach 1 (species-specific), the time lag between the arrival of the response species and the measurement is kept constant in the target and reference arrival sequences. In Approaches 3 and 4 (assembly-specific), the time lag between the first introduction event and the measurement is kept constant in the target and reference arrival sequences. Panel (a) compares species-specific Approach 1 with assembly-specific Approach 3, while panel (b) compares synchronous Approach 3 with reverse Approach 4. See also Figure S7 for a comparison between Approaches 1 and 4. In each panel, the dashed line represents the 1:1 line. The solid black lines are SMA regression models, built separately for each plot (Table 2). The red areas represent zones where two approaches do not agree on the direction of the priority effect.

Figure S7. Pairwise relationships between priority effects (PE) in two-species *Bifidobacterium* communities measured using different approaches and metrics. For each response species, priority effects were calculated 16 h after the start of the experiment (i.e. 4 h after their introduction). Priority effects were quantified using four different metrics (Equations 1, 2, 3a and 4a). We used two reference scenarios (Figure 1b): a reverse reference, in which the positions of the early- and late-arriving species are reversed while keeping all other factors constant (Approach 4), and a synchronous reference, in which both species are introduced simultaneously (Approach 1). Approaches 1 and 4 further differ in the way they standardise time intervals between introduction events and performance measures. In Approach 1 (species-specific), the time lag between the arrival of the response species and the measurement is kept constant in the target and reference arrival sequences. In Approach 4 (assembly-specific), the time lag between the first introduction event and the measurement is kept constant in the target and reference arrival sequences. In each panel, the dashed line represents the 1:1 line. The solid black lines are SMA regression models, built separately for each plot (Table 2). The red areas represent zones where two approaches do not agree on the direction of the priority effect.

How to cite this article: Delory, B. M., Little, C. J., Marton, H., Halliday, F. W., Barnes, E. M., Stiver, I., Sarneel, J. M., van Steijn, T. L. H., Grainger, T. N., Dieskau, J., Kardol, P., Knight, T. M., Ladouceur, E., Ojima, M. N., Roscher, C., Stroud, J. T., Temperton, V. M., Werner, C. M., & Fukami, T. (2026). Quantifying priority effects in community assembly experiments. *Methods in Ecology and Evolution*, *00*, 1–18. https://doi.org/10.1111/2041-210x.70449

2041210x, 0, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70449 by tadashi fukami - Stanford University , Wiley Online Library on [26/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License