Biology & Philosophy (2020) 35:29
https://doi.org/10.1007/s10539-020-9742-7

Microbiome causality: further reflections (a response to our commentators)

Kate E. Lynch[1] · **Emily C. Parke**[2] · **Maureen A. O'Malley**[3]

Received: 23 November 2019 / Accepted: 28 January 2020

Introduction

Microbiome research and causal explanation are both topics of considerable attention in philosophy of biology. In our target article (Lynch et al. 2019), we brought these two topics together. Now, commentaries on our article explore this connection in a number of constructive ways. We organize them into five main themes. Each addresses in different ways how to refine and supplement interventionist accounts of explanation in microbiome and other research contexts.

A key aim of our target article was to challenge the notion that whole microbiomes are causally responsible for two human health conditions: obesity and mental illness. We undertook this project by first examining causal explanation in a well-known and non-controversial example in microbiology: *Helicobacter pylori* as the cause of ulcers. We showed that even this seemingly straightforward single-microbe case achieved low scores on two of the interventionist dimensions of causal explanation (stability and specificity). We suggested it did better on the third dimension, proportionality. When we assessed all three dimensions in causal explanations putatively connecting microbiomes to obesity and mental health, the explanatory achievements were much more limited. Scores on all three interventionist dimensions of causal explanation were low.

This reply refers to the commentaries available at https://doi.org/10.1007/s10539-020-9739-2, https://doi.org/10.1007/s10539-020-9738-3, https://doi.org/10.1007/s10539-019-9710-2, https://doi.org/10.1007/s10539-019-9711-1, https://doi.org/10.1007/s10539-019-9713-z, https://doi.org/10.1007/s10539-019-9707-x, https://doi.org/10.1007/s10539-019-9708-9, https://doi.org/10.1007/s10539-019-9728-5, https://doi.org/10.1007/s10539-020-9740-9, https://doi.org/10.1007/s10539-019-9715-x, https://doi.org/10.1007/s10539-019-9709-8, https://doi.org/10.1007/s10539-019-9718-7.

✉ Kate E. Lynch
kate.lynch@sydney.edu.au

[1] Department of Philosophy and Charles Perkins Centre, University of Sydney, Sydney, Australia

[2] Department of Philosophy, University of Auckland, Auckland, New Zealand

[3] School of History and Philosophy of Science, University of Sydney, Sydney, Australia

Published online: 17 April 2020

Some of the commentaries on our target article undertake a closer examination of those dimensions in order to understand the relationships between stability, specificity, and proportionality. These reflections are the focus of "Theme 1: Refining the interventionist dimensions".

Other authors propose that lessons can be drawn from additional scientific fields, and that microbiome research, as well as interventionist approaches, can learn from such comparisons. We examine these suggestions in "Theme 2: Using comparative cases to illuminate causal explanation".

The interventionist account writ large is the target of several commentaries, some of which propose alternative or additional criteria for assessing causal explanations. We discuss these arguments in "Theme 3: Additional approaches to causal explanation".

A number of commentators argue for broader pragmatic concerns in explanation, such as considering the practicalities of communicating causal explanations, alongside the inevitable use of imperfect models and the influence of background explanatory factors on causal accounts. We cover these issues in "Theme 4: Pragmatic considerations for causal explanation".

Finally, several commentaries challenge our conceptual typology of microbiomes and their potential causality. Our target article identified four microbiome concepts that have different levels of description and causal roles. We argued that in many cases, a 'causal core' microbiome concept is the most appropriate for claims about microbiome involvement in obesity, mental health, and other host-related states. We discuss the commentaries that explore this argument in "Theme 5: Communities and causality".

Theme 1: Refining interventionist dimensions (Oftedal, Attah et al., Schneider)

Although a cluster of commentaries are broadly concerned with whether interventionist criteria are appropriate for evaluating microbiome research and other fields, in this theme we focus on suggestions that the interventionist framework needs specific conceptual and operational refinements. These proposals are made in order to capture more fully how causal explanations work in actual scientific practice.

Oftedal (2020) suggests that the utility of each dimension depends upon the context of inquiry. In some cases, one or more of the dimensions of stability, specificity or proportionality will matter less than the others, or will not matter at all. The relative importance of each dimension might also be diminished when there are large effect sizes. For example, if a microbiome causal core were to have a large effect on a trait, but only in particular background conditions (e.g., host demographics or diet), this effect size would become more important than the explanation's limited stability. Something equivalent is occurring in broadly effective HIV treatments that nevertheless score poorly on the specificity dimension (Neal 2019).

We fully agree with Oftedal that some dimensions are more important than others depending on the explanatory context. Our single-microbe example, of *Helicobacter*

pylori causing ulcers, did not perform well on specificity and stability assessments, yet still proved to be a valuable causal explanation due to its proportionality. The *H. pylori* explanation identifies the most appropriate cause on which to intervene, and this intervention results in a successful treatment. Unfortunately, most microbiome research does not achieve similarly large effect sizes and treatment success. Any putatively causal explanations thus fall short on all dimensions, and succeed primarily when a 'causal core' supplants the whole microbiome.

However, it may more generally be the case that no explanation is going to score equally well on all dimensions. This would occur if there are necessarily trade-offs between each dimension, as suggested by some of our commentators (Attah et al. 2020; Oftedal 2020). Although we hint at this important issue near the end of our target article (Lynch et al. 2019: 19), it is a topic that needs considerable development within the interventionist literature, in which each dimension has mostly been addressed independently.

We suggested in our target article that a small causal core of microbiome members might be responsible for the majority of certain host effects, in which case that 'core' group should be selected as the most proportional causal explanation. However, because this selection relegates additional variables (the 'non-core' microbes in the microbiome) to the background conditions, the explanation could then become more unstable. Increased instability would occur because of the increased potential for variation in that expanded background to perturb the focal causal relationship and moderate the effect size. Relatedly, Schneider (2020) suggests that broader (less proportionate) causes that include larger numbers of background conditions enable a more specific causal relationship. How would assessments of each dimension be evaluated with such trade-offs in mind?

We can see two potential solutions. One would be to develop a precise and commensurate quantification of each dimension, such that evaluations of competing causal factors would rely on quantified measures of stability, specificity and proportionality. Although interventionism is far from having such measures, recent work demonstrates a precise, theoretically principled way in which to measure specificity (Griffiths et al. 2015). Such achievements suggest that each dimension might be amenable to quantification. However, making those measures comparable for each dimension will add considerably to the difficulties of such a project.

The other option is that the interventionist dimensions and their general framework might be understood as heuristics (see Fagan 2019). As several commentators suggest below, the importance and value of each dimension will vary according to the research context and explanatory context. In some cases, proportionality might be central to an explanation; in others, stability will matter more, and so on. We think there is value in assessing explanations with interventionist criteria even when they are considered 'mere' heuristics. Our qualitative application of these heuristics to microbiome research demonstrates this fruitfulness by illuminating explanatory shortfalls and ways in which to realize better explanations. Especially when explanations are coarse and preliminary, applying interventionist criteria can help focus explanatory strategies in an epistemically efficient way.

Theme 2: Using comparative cases to illuminate causal explanation (Fagan, Gomez-Lavin)

Two of our commentators turn to comparative cases to illustrate the complexities of different explanatory contexts. Fagan (2019) describes parallels in the emergence of stem cell research, and Gomez-Lavin (2019) discusses neuroscience. In both fields of research, findings have been overblown, and the focal causal entities are complex and poorly understood. Fagan suggests that microbiome researchers might apply an 'integrated HPS' approach in order to draw on successful methodological outcomes in other fields. We agree this is a useful strategy, but do not think implementing it would preclude a focus on interventionism. Indeed, what Fagan calls for primarily is more attention to relevant scientific practices, and we take that as an invitation to elaborate on the microbiome-obesity material we discussed in our target article.

There, we described how microbiome research held early hopes for simple ratios of the phyla Bacteroides and Firmicutes to explain and ultimately help treat obesity. These hopes have been thoroughly dashed, and it is now acknowledged that the vast diversity in a phylum such as Bacteroides makes it unrealistic to expect a phylum-level effect (Johnson et al. 2016). Small sample sizes and background variations also mean correlations and experimental findings are simply not stable. Since we wrote our target article, obesity research has delivered new insights into the potentially causal relationship between microorganisms and obesity. In one particularly notable finding, Petersen et al. (2019) found that reintroducing the bacterial class Clostridia can prevent obesity in mice, possibly via control of lipid absorption in the hosts. Put the other way around, when Clostridia decrease in abundance and there are concomitant increases in bacteria belonging to the genus *Desulfovibrio*, these revised ratios are associated with obesity, at least in this study.

Without wanting to diminish the achievements of such research, we see the potential for disproportionality when a whole class of bacteria is suggested as the causal agent. Proposing a class as a cause is similar to positing a whole microbiome as a cause. Clostridia are very diverse (they are a sub-group within Firmicutes) and are recognized as a class on the basis of very broad characteristics (e.g., obligate anaerobic respiration and spore-forming capacities). Perhaps they all bear a specific pathway that brings about obesity protection in mice, but in that case, it would have to be shown that other groups do not have this pathway, and that all organisms classified as Clostridia are doing this particular thing to their hosts. An application of the interventionist framework could usefully guide future research on the basis of this promising but overly broad finding. This application could involve, for example, identifying the proportionate group within Clostridia responsible for the effect, tracking specific relationships between microbe abundance and the magnitude of the effect, and observing stable associations between those microbes and obesity. A more proportionate causal relationship is potentially found between the co-varying *Desulfovibrio* genus (a far smaller taxon). However, Petersen et al. (2019) focus less on this group for contextual reasons:

it is in many cases more useful for treatment to find microbes that protect against obesity (i.e., Clostridia), than to identify those that might contribute to it (i.e., *Desulfovibrio*). We discuss contextual factors further in Theme 4.

Fagan (2019) describes how early stem cell research parallels the emergence of microbiome research. An early approach to stem cells attempted to identify a 'causal core' of genes expressed by all stem cells and no other cells. However, this approach was not fruitful, because only a single gene met these criteria. Instead, researchers went on to engineer stem cells in order to identify a 'core set' of interacting molecular factors that switched on pluripotency. Fagan argues that the explanatory goal of selecting a causal core by interventionist criteria was not viable in stem cell research and is similarly unlikely to work for microbiome research. However, our reading of both accounts is that underlying assessments of dimensions such as proportionality are actually at work despite the different methodologies employed. As Fagan describes so clearly, the first finding of a single gene was ruled out on grounds similar to proportionality (too narrow a cause for the wider effects). That 'failure' was followed by the eventual identification of a 'core set' of three molecular factors of just the right scope, stability and specificity. Indeed, the means by which the identification was made are novel and interesting, but the general strategy is the same as in microbiome research. It too requires the identification of the causal 'cores' responsible for host states, and this is achieved by a range of back-and-forth methods that aim to exclude causally irrelevant (or less relevant) microbes (see, for example, the methods used by Buffie et al. 2015). Although Fagan does not conceptualize her rich case study material in terms of the interventionist framework, we are tempted to conclude that interventionist dimensions are effectively guiding methodological choices in both stem cell and microbiome research.

Gomez-Lavin (2019) introduces a different comparative case of neuroscience. Unlike Fagan, who turns to the history of stem cell research to demonstrate how progress can be made in causal investigation, Gomez-Lavin is pessimistic about the search for causal explanations in certain epistemic contexts. This pessimism rests on the notion that both brains and microbiomes are complex dynamic systems with built-in redundancy mechanisms that produce multiply realized effects. We agree that microbiomes share some of these features, which is why an entire microbiome is an inappropriate causal explanation for most health and disease states. This does not mean that successful causal explanations cannot be made in microbiome research, but that there is limited room for successful *whole* microbiome causal explanations.

While it is not very useful to say that 'the brain' is a cause of language abilities and deficiencies in humans, there *is* explanatory power in identifying particular brain regions. Wernicke's and Broca's areas, for instance, are causally important for particular language-related outcomes. Lesions in these regions result in particular dysfunctions of language across populations. These areas are differentially and consistently activated for specific stimuli, which indicates specificity and stability (Alexander et al. 1987; Embick et al. 2000). Likewise, we believe there is promise in identifying key component parts of the microbiome that make a difference to human health and disease, and thereby generate causal explanations that satisfy interventionist dimensions. We take it that Gomez-Lavin (2019) believes that these kinds of modular brain areas are the exception rather than the rule in neuroscience,

and perhaps that will also be the case for causal core explanations in microbiome research. This worry is an empirical matter to be addressed with further research. However, we are happy to go on the record and predict that microbiome research will achieve most of its future explanatory success by focusing on particular microbial groups or pathways in microbiomes (e.g., the *Clostridiodes difficile* case in our target article and further below[1]).

Even if microbiome effects on human health and disease states turn out to be mainly the result of irreducibly complex interactions between billions of microbes, some specific causal interpretations would probably still be warranted. Genetics provides another useful parallel to microbiome research. Most human traits are influenced by many genes that interact in complex ways with one another and the environment. Yet genetic causal explanations are still achieved, with methods such as heritability analysis and genome-wide association studies. Findings from such studies indicate the relative causal contribution of genetic differences across the genome. Note that the genome itself is rarely cited as the focal cause in this field. Instead, causation is attributed to combinations of 'genes'. Additionally, there are thousands of human traits in which a single or just a few genes explain traits both specifically and stably (OMIM 2019). Microbiomes may act similarly, with many traits influenced by diverse microbes interacting in numerous ways with multiple environments. But just as for genes, microbes, rather than whole microbiomes will be the appropriate causal focus for specific health and disease states.

Theme 3: Additional approaches to causal explanation (Gillies, Greslehner and Lemoine, Klassen)

Rather than interrogate the interventionist framework with different cases, Gillies (2019) argues for an altogether stronger take on explanation. He suggests that standard criteria of necessity and sufficiency lead naturally to causal laws as the desirable focus of biological explanations. Gillies argues that in medicine the identification of *avoidance actions*—those that prevent or cure diseases—is more important than the identification of *production actions*, which are those that cause the disease. Based on this schema, he calls for an amendment to Koch's postulates. The third postulate now becomes: "If the microorganisms are prevented from multiplying in the patient's body, then the patient will not have the disease".

Gillies then proposes that instead of employing interventionist dimensions to indicate the degree of causal explanatory strength, the amended postulates can be used to obtain causal laws that identify *the cause* of a disease. Optimal causal laws are found when the cause is a necessary condition for the effect, and that cause can

[1] We should note here that in our target article we referred to '*Clostridium difficile*', the spore-forming pathogen in intestines that can cause illness and death. However, in 2016, the genus this organism belongs to was renamed, and we should have said *Clostridiodes difficile* (Lawson et al. 2016). *C. difficile* throughout our response now refers accordingly, although it is commonly known as *C. diff*. Nomenclature may soon become even more complicated because *C. difficile* appears to be speciating (N. Kumar et al. 2019).

be prevented from occurring, thereby precluding the effect. Gillies' amendment means Koch's postulates are in fact fulfilled in the *H. pylori* case, because the bacterium is necessary for the effect, and ulcers can be eliminated by intervening on the bacterium. A causal law is thus obtained. The implication of Gillies' argument for microbiome research is that whole microbiomes could be identified as causes of ailments that involve *C. difficile*. Gillies identifies unhealthy microbiomes as causes of *C. difficile* because they are necessary (but not sufficient) causes of the adverse effects of the infection, and because the symptoms disappear when whole microbiomes are intervened upon via fecal microbiome transplants (FMTs).

However, as we argued in our target article, 'necessity' could be achieved only in a circular way for the ulcer case: *H. pylori* is necessary for *H. pylori*-caused ulcers. Other ulcers are caused by other factors (e.g., non-steroidal anti-inflammatory drugs), hence the need for specifying the kind of ulcer. Our target article also discussed the circularity of identifying an unhealthy or 'dysbiotic' microbiome, due to the cause being defined in reference to its effect (Lynch et al. 2019: 10). But even leaving the circularity issue aside, identifying whole microbiomes as the relevant causal agent is problematic under Gillies' account if interventions on just a few members of a microbiome (i.e., a causal core of microbes) bring about the same avoidance action for *C. difficile* infections. A microbiome subgroup is exactly what Buffie et al. (2015) found to counter *C. difficile* overgrowths, which means that many components of a whole microbiome are not, in fact, necessary to cause, nor to overcome the problematic infectious state. Gillies' own account of necessary and sufficient conditions should thus reduce whole microbiomes to the causal core that necessarily brings about the effect. We arrive at the same conclusion in our paper but via a different route: causal cores can provide a better causal explanation due to proportionality, because it narrows the focus and excludes irrelevant microbial factors.

In addition, although Gillies's view of causal laws may indeed prove useful for separating causes from non-causes, it imposes a demanding set of conditions for causation. Determining causation in terms of necessity and sufficiency in line with Koch's postulates implies maximal stability—that the cause will bring about the effect under every possible background (Postulate 1)—and a high degree of specificity—that the cause does not occur in other diseases or non-pathogenically (Postulate 2). Many causal relationships in biology simply will not meet these criteria, especially in systems with many complex interacting components such as microbiome-host relationships. Using interventionist dimensions as an alternative tool for assessing causal explanation permits a more applicable understanding of causal relationships and their relative explanatory strengths.

A second commentary also looks at ways in which to supplement interventionist approaches. Greslehner and Lemoine (2020) propose a 'dual decomposition' strategy in which the 'phenotypic' (i.e., host) effects of interest are conceptually decomposed in the same way as microbiomes. We agree that rigorous conceptualization of the effects of interest is important for scientific research and its neglect may account for some misunderstandings of causal relationships in microbiome research. Indeed, carefully specifying the effect of interest is a necessary step for identifying the most proportionate causal explanation, which we discuss in further detail in Theme 4. In our target article, Footnote 6 already elaborated on how proportional causes and

effects must be identified (or in Greslehner and Lemoine's terminology, 'decomposed'), relative to one another (see also Attah et al. 2020). In that earlier discussion, our analysis revealed problems in the conceptualization of the host phenotypic effects of mental health states. While claims are often made about the causes of human anxiety, more proportionate descriptions of the effect observed in these studies are likely to be 'anxiety-like states' or particular anxiety-indicating behaviours in rodents. It is probably common for interventionist accounts to focus more on proportionality of causes rather than proportionality of effects, so this commentary's reminder to pay attention to effect proportionality is useful.

Another supplementary approach is provided by Klassen (2020) (see also Schneider (2020)), who suggests that a better understanding of causation in microbiome research requires ecological considerations to be used alongside an interventionist approach. Klassen argues that experimental manipulations of microbes are limited in their applicability outside of experimental settings. This issue is usually diagnosed in the intervention literature as a problem of stability: causal relationships can be observed only in a limited set of background conditions, which are usually tightly controlled in laboratory experiments. The stability of causal relationships between specified microbes and their effects is an empirical matter, and there are some examples in microbiology that attest to a high degree of stability (e.g., bacterial neurotoxins, such as those produced by *Clostridium botulinum* and *Clostridium tetani*—see Humeau et al. 2000). However, when this degree of stability does not hold, examination of other background factors may be required. We discuss the potential problems that arise when considering larger communities of microorganisms as part of the causal explanation in Theme 5.

Theme 4: Pragmatic factors in assessing causal explanations (Donhauser et al., Greslehner and Lemoine, Attah et al.)

A number of commentators drew attention in various ways to the pragmatic limitations of explanatory science. Perennial issues arise from reliance on imperfect but useful models; indeed, Donhauser et al. (2019) refer repeatedly to 'mooted models' of causal relationships as synonymous with hypotheses that are as likely to be false as not. More broadly, every scientist faces the day-to-day practicalities of explanation, plus communication of those explanations (Greslehner and Lemoine 2020). We agree with these observations, though maintain that microbiome researchers still face particular conflicts in their current explanatory practices. Notably, clear acknowledgement of the senses in which models are flawed, and appropriate qualifications of the causal claims being made about microbiomes, would alleviate overselling complaints about microbiome findings.

Greslehner and Lemoine (2020) push back against our argument for precision, and express sympathy for 'place-holder' claims that communicate causal relationships pragmatically but imprecisely. Attah et al. (2020) also follow this line when they acknowledge the temporary need for black-boxing unknown explanatory elements. While strategies that accommodate uncertainty and lack of knowledge may sometimes be useful, there are extraordinary misinterpretations—both deliberate

and accidental—of microbiome findings, in social and professional media as well as on commercial platforms (Eisen 2017; Hooks et al. 2019). As we and others have argued, looseness in scientific communication can lead to broader misinterpretations that ultimately harm scientific fields, such as microbiome research, far more than they help (e.g., Perez-Muñoz et al. 2017).

Such problems indicate to us that the context of any explanatory black-boxing should be taken into account when assessing an explanation's appropriateness. In a microbiology laboratory or the broader microbiome research community, where much knowledge is shared, it might be reasonable to refer to 'the microbiome' as shorthand for the causal agent of interest. In such situations, it is possible that everyone using the shorthand is aware that there are important unknown details about the roles of different organisms and their interactions in a complex ecosystem with unidentified causal structure. However, our target article showed that there are multiple characterizations of the microbiome concept used in the literature. Given this, a black-boxing approach is likely to lead to miscommunication as researchers fill their black-boxes with different meanings. When communicating outside the immediate scientific community, it becomes even more clear how problematic it is to make claims about 'the microbiome' causing a disease or mental health state, without appropriate qualifiers (i.e., about the extent of the black-boxing involved).

Something positive to note here is how the microbiome research community is moving toward modelling strategies that allow better assessments of complex causal relationships. Fagan (2019) mentions the potential value of in vitro experimentation ('gut organoids') for microbiome research, which we agree would allow future testing of broad hypotheses in attempts to narrow them. In vitro experiments, also known as material or experimental models (O'Malley and Parke 2018), would enable a trial-and-error approach for identifying the most proportionate causal agents of any given effect. For example, it would be a useful strategy if key members of Clostridia could be identified and certain pathways isolated in the prevention of obesity (see Petersen et al. 2019, as discussed in Theme 2). An increasingly common way of doing this involves experimental 'synthetic ecosystems', which in all the relevant respects can be considered as in vitro models (Vrancken et al. 2019; Johns et al. 2016). This involves choosing a tractable number of members of the whole community (rarely more than a dozen 'representative' members and often fewer), then exploring the interactions between members to identify relationships and their outcomes (e.g., Venturelli et al. 2018). Such systems are particularly relevant when attempting to 'restore' microbiomes in diseased hosts (Vásquez-Castellanos et al. 2019). Obviously, modelling of this sort (simplified laboratory systems), just like standard mathematical and computational modelling of microbiomes (M. Kumar et al. 2019), relies on being able to break down communities to individualized entities (pathways or organisms) before reconstituting that community and its effects via interaction dynamics. In fact, it currently seems to be the case that bottom-up interactions between pairs of microbes predict community dynamics (including stability), rather than any 'emergent' properties of the whole system (Venturelli et al. 2018; Abreu et al. 2018). Although all such models have their limits (see Donhauser et al. 2019; Klassen 2020), accepting this situation does not, very fortunately, produce epistemic paralysis in microbiome research or other fields.

In that spirit, some commentators made pragmatic suggestions that could lead to better explanatory outcomes for microbiome research. Attah et al. (2020) argue that pragmatic factors influence the proportionality of a causal description. Their account embeds causal selection via the interventionist account within strongly practical concerns. They point out that the selection of a more fine- or coarse-grained causal description is in part due to pragmatic decisions about what observers can intervene upon (see also Oftedal 2020). When there is uncertainty about the causal system, or when tools for intervention are not well developed, that context of enquiry justifies selecting a relatively coarse-grained causal description. This situation, they believe, applies to microbiome research and would justify the selection of whole microbiomes as broad causal agents.

We certainly agree that contextual and pragmatic factors can play important roles in causal explanation. The initial selection and description of effects of interest are contextually and pragmatically informed. For instance, depending on research interests, researchers could focus on the broad effect of 'depression' rather than more specific symptoms such as insomnia or hypersomnia. However, once an effect is selected and described, the most proportionate level of description for a causal explanation becomes a more tightly constrained matter. Given a targeted effect, a proportionate cause will capture the factors that bring about the effect and exclude those that are irrelevant to it. In other words, once an effect is selected (by pragmatic means), the best explanation for that effect can be determined by non-pragmatic means, which is what Woodward's interventionist account proposes. In particular, Woodward (2018) maintains that proportionality can be applied *only* with a prespecified effect in mind. It is important to note that whole microbiome explanations do not fall short on the proportionality criteria in isolation, but only in reference to particular effects (contra Donhauser et al. 2019). Pragmatic considerations thus play a limited role once a targeted effect has been selected. This entails that whole microbiomes are unlikely to be proportionate causes for specified health or disease states, even when there are uncertainties about the causal system of interest.

We suggest that rather than influencing the proportionality, stability, or specificity of causal explanations, practical issues are more likely to influence decisions about which putative cause to intervene on, particularly in medical settings. The *C. difficile* case helps illustrate this point. In our target article, we mentioned research that had found just four members of the whole faecal microbiome community to be the most efficacious in ameliorating the debilitating effects of *C. difficile*, probably via the mechanism of secondary bile acid synthesis (Buffie et al. 2015). However, most of that effect could be attributed to a single member of the four, *Clostridium scindens*. Although for treatment purposes, the maximum effect would be obtained by introducing all four bacteria, in practice it might be easier to administer and define the mechanism of effect for just one bacterium. In such a situation, the pragmatic decision is fairly straightforward. Conversely, if the same effect is achieved by FMTs (because the focal organisms are part of the transplant), and FMTs are more

easily administered than more precise interventions,[2] then pragmatic considerations mean FMTs will continue to be the intervention of choice despite their causal disproportionality.

The new obesity findings (Petersen et al. 2019) we mentioned in Theme 2 also illustrate this point. Clostridia, a broad class of bacteria, are the causal focus rather than *Desulfovibrio*, a much smaller genus. That focus occurs because intervening on Clostridia might prevent obesity, whereas intervening on *Desulfovibrio* might intervene on only one of several factors known to be implicated in producing obesity. Not only is any explanation involving *Desulfovibrio* more cumbersome, but it is less practically appealing to try and decrease a smaller target when compared to increasing the broader target of Clostridia with its potentially positive outcome.

Theme 5: Community causality (Schneider, Klassen, Gomez-Lavin, Watkins and Bocchi, Lean, Skillings)

Whether and how communities can be causal in their own right is a fifth theme in the responses, and indeed, this is a central issue addressed in our target article. Several of our commentators make cases for different views of community causality.

Schneider (2020), Klassen (2020), Gomez-Lavin (2019), Lean (2019) and Skillings (2019) make arguments about community causality that rest upon what causation means. These commentaries discuss in various ways how background conditions are important for the production of the effect. We employ a narrower difference-making view of causation, in which a focal causal entity is identified by the influence of a potential intervention. This kind of causal account requires that all background conditions are held fixed in order to determine the cause, and thus are not included in the explanation. Difference making thereby captures how experimentation gains insight into causal relationships in scientific practice (Woodward 2003). The identification of difference-making intervention points is also the basis of treatment regimes in health research.

The selection of focal causes from a broader causal background has theoretical justifications beyond the interventionist account and everyday experimental practice. For any given effect, a large number of factors will count as contributing causes, despite their limited explanatory relevance (Menzies 2004). To overcome this problem of profligate causes, Dretske (1995) distinguished between 'structuring causes' and 'triggering causes', where structuring causes are defined as causal factors that enable triggering causes to produce an effect. While relevant to the effect, the enabling role of structuring causes means they can be relegated to background conditions. Triggering causes regularly give rise to the effect of interest and would be similarly identified by an interventionist account of causation. Importantly, however, an appeal to difference makers in the form of triggering causes requires some specification of the explanatory context, which other commentators have already raised as

[2] However, FMT risk calculations are likely to increase in practical importance after recent warnings of fatalities following FMTs in immunocompromised hosts (FDA 2019).

a complicating issue (Theme 4). We agree that different explanatory contexts may influence what we consider the actual difference maker to be, but this still leaves difference making at the heart of the causal explanation (for discussion, see Waters 2007).

How does all this work in the case of microbiome research? Whole-microbiome explanations include structuring causes (not just triggering causes), because many constituent microbes and their interactions play a subsidiary explanatory role in which they enable the triggering causes. Proportionality assessments offer a useful way in which to establish the key difference-making factors that are explanatorily relevant causes. A compounding problem if background conditions are part of microbiome explanations is the issue of 'normal' or baseline microbiomes. Descriptions of these entities are proving elusive because of the high inter-individual variability of microbiomes, which means microbiome community structure in relation to host health is basically unknown (Lloyd-Price et al. 2016). In these circumstances, adding extremely variable background conditions (as described by Klassen (2020) and Schneider (2020)) to any putative microbiome explanation would result in such overwhelming causal profligacy that the explanation could be only weak and unilluminating.

In efforts to deal with this problem, three commentaries suggest situations in which causality could be assigned to microbiomes or specified sub-microbiomes. These analyses identify three candidate entities at different scales of community causality: holobionts, which are all the microbes and the host (Watkins and Bocchi 2020), ecological units, which are sub-communities of whole microbiomes (Lean 2019), and core causal consortia comprising organisms containing other organisms (Skillings 2019). We address each of these in turn.

Holobiont concepts are by now well-trodden grounds of debate for researchers concerned with microbiome-host characterization (e.g., Skillings 2016; Douglas and Werren 2016; Moran and Sloan 2015; Queller and Strassmann 2016). Although the term has alternative interpretations, in its most standard sense it refers to a host and all its microbes. Watkins and Bocchi (2020) are concerned that we may not have paid this concept sufficient attention, and think that if we had, our concerns about microbiome explanations would be alleviated. In fact, we believe the opposite is true. We did not mention holobionts because we think the concept muddies the epistemic waters even more than they are currently muddied. We might even go so far as to say that large amounts of philosophical attention to putative entities in microbiome research, without due consideration of their roles in explanation, may be contributing to the reification of loose communities of animal hosts and microbes (see Douglas and Werren 2016).

Every point we made about 'whole microbiomes' applies with as much force if not more to 'holobionts'. Indeed, it becomes even harder to identify the appropriate causal locus when the host (where effects are located) is the same entity as the microbiome (where causes might be located). 'Dysbiosis' is certainly not the way in which to describe these health and illness effects, due to circularity and causal obfuscation (Olesen and Alm 2016; Hooks and O'Malley 2017). It is also far too soon to say that microbiomes and human hosts have mutualistic or even 'essential' functional relationships, as Watkins and Bocchi (2020) suggest might be the case.

Highly competitive, antagonistic interactions dominate host-microbiome states, with hosts exerting as much control as possible over long-term residents and opportunistic arrivals (Foster et al. 2017; Queller and Strassmann 2016).

We also consider it non-explanatory to say that obesity and mental health are 'emergent properties' of holobionts; our target article makes the point that such language is just a stop-gap.[3] Emergence, when asserted in a microbiome research context, usually indicates lack of current knowledge about who does what and how. Gomez-Lavin (2019) makes a case for 'global properties of a microbiome', but we think microbiome research is currently striving to understand mechanistically via decomposition how such apparent global properties come about. Although there are numerous microbiome studies of community properties that might be considered emergent, such as the stability or instability of community composition, even the best of such studies are unable to identify the causal structure between such states and the host. For instance, unstable microbiomes in hosts with particular illnesses may be the outcomes of medication for the disease, rather than causally productive of the disease itself (e.g., Halfvarson et al. 2017).

Lean (2019) focuses on how specified networks of microorganisms within microbiomes might be causal for the health status of the host. Broad health states, he suggests, are not going to be explained by single microbes. We are happy to agree. The key question for us is how the smaller networks involved in smaller scale effects will be identified, and we suggested in our target article it will be by decompositional means. Each 'function' of a microbe (i.e., its biochemical activity), and its co-dependent interactions with others in the production of that function, is probably going to have to be understood from the bottom up, rather than the top down. This recompositional work underlies some recent microbiome success stories, such as the identification of the key causal network inhibiting *C. difficile* proliferation (Buffie et al. 2015).

Many examples of recent microbiome research also follow this recompositional programme. As a case in point, Raman et al. (2019) refer to the 'emergent' function of microbiomes in producing healthy gut development, but really mean a network (an 'ecogroup') of only 15 taxa. Likewise, Schulfer et al. (2018) resist indicative findings of a single harmful bacterium having a 'crucial role' in antibiotically perturbed microbiomes. Instead, they argue that this organism is embedded in a still-to-be-fully-defined network of 'injurious' species that have displaced a 'protective' network of other species. As these authors go on to suggest, standard experimentation will be required to tease out how these sub-networks work within microbiomes and more broadly on the host. In short, we think Lean (2019) is right, but the devil is in the methodological details. At least methodologically, it is likely that recomposition will be carried out by first identifying single causal contributors by interventionist means and then their interactions.

[3] Here, we dispute Attah et al.'s (2020) characterization of our position as dichotomous: either individual microbes are causal, or emergent properties are causal. We do not think emergent causation is at work in microbiome-host relationships.

Skillings (2019) proposes that 'core causal consortia' of multiple organisms can meet proportionality requirements in certain explanatory contexts. He examines interactions that involve extra tiers of intermediate hosts. In some cases, such consortia may form metabolic individuals, and in even fewer cases, these individuals may have evolutionary unity. But the key point for our analysis is that Skillings suggests certain members of such consortia may not be the main causal agents, but are nevertheless essential to the production of the effect (e.g., the trematode that hosts the bacterium that causes the disease in the larger animal host).

Although we find Skillings' main example very illuminating, the key to any reconstruction of host health state is identification of the disease cause. If it requires delivery to the host by another organism, we would see this vector as a structural rather than a triggering cause. An appeal to relevant counterfactuals, such as other natural or experimental systems, picks out the bacterium as a triggering cause, and not the vector. Put more concretely, no amount of trematodes will deliver the disease if the bacteria are not there, whereas bacteria will bring about the disease even when delivered by some other means. That said, obligatory delivery systems for causal agents may be highly pertinent to the pragmatic side of the story. In the scenario described by Skillings (2019), structural causal consortia members may act as useful intervention points for treatments, because they provide reliable proxies for the core causal triggers. Identifying the delivery vector as the intervention target for the sake of treatment (but not explanation) brings us back to Gillies' (2019) emphasis on avoidance actions for treatment, and Attah et al.'s (2020) point about the importance of pragmatic considerations for identifying useful interventions.

Concluding reflections

We appreciate the opportunity given by these commentaries to reflect further on the key message of our target article and its interventionist framing. Several of the commentaries add to the growing toolkit of methods for better evaluating causal explanations. Even if the ultimate achievement of the interventionist framework is that it operates as a heuristic, it can guide researchers in causal selection and explanatory communication. Topics of particular interest for developing this framework (perhaps beyond its heuristic base) are how to accommodate trade-offs between dimensions, whether the dimensions can be further formalized and quantified, and whether interventionist accounts can be made more rigorous with contributions from other views of causal explanation.

The main message of our target article, that whole microbiomes are a poor basis for causal explanations of host effects, is the other key topic of discussion in these commentaries. Our conclusion is that even when opposing views in the commentaries are taken into account, whole microbiomes end up reinforced as inappropriately broad for causal explanation. This conclusion suggests that a lexical shift might be timely for microbiome research, away from whole-microbiome causal shorthand, and back to a more traditional focus on individual microbes and their interactions, both among themselves and with the host. Although broader communication is only part of the scientific picture, casual reliance on loose and broad terms with causal

implications may have run its course in the development of microbiome research. We will be delighted if future research proves us wrong—indeed, philosophers have been attracted to microbiome research for the very idea there might be whole-community causality and individuality—but predict that it won't.

References

Abreu C, Lopez AO, Gore J (2018) Pairing off: a bottom-up approach to the human gut microbiome. Mol Syst Biol 14:e8425

Alexander MP, Naeser MA, Palumbo CL (1987) Correlations of subcortical lesion sites and aphasia profiles. Brain 110(4):961–988

Attah NO, DiMarco M, Plutynski A (2020) Microbiomes: proportional causes in context. Biol Philos 35:22

Buffie CG, Bucci V, Stein RR et al (2015) Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. Nature 517:205–208

Donhauser J, Worley S, Bradie M, Bouzat JL (2019) The problem of mooted models for analyses of microbiome causality. Biol Philos 34:57

Douglas AE, Werren JH (2016) Holes in the hologenome: why host-microbe symbioses are not holobionts. mBio. https://doi.org/10.1128/mbio.02099-15

Dretske F (1995) Mental events as structuring causes of behaviour. In: Heil J, Mele A (eds) Mental causation. Clarendon Press, Oxford, pp 121–136

Eisen JA (2017) Microbiomania and 'overselling the microbiome'. The tree of life. https://phylogenomics.blogspot.com/p/blog-page.html

Embick D, Marantz A, Miyashita Y, O'Neil W, Sakai KL (2000) A syntactic specialization for Broca's area. Proc Natl Acad Sci USA 97(11):6150–6154

Fagan MB (2019) Commentary on 'How causal are microbiomes?'. Biol Philos 34:58

FDA (2019) Important safety alert regarding use of fecal microbiota for transplantion and risk of serious adverse reactions due to transmission of multi-drug resistant organisms. https://www.fda.gov/vaccines-blood-biologics/safety-availability-biologics/important-safety-alert-regarding-use-fecal-microbiota-transplantation-and-risk-serious-adverse

Foster KR, Schluter J, Coyte KZ, Rakoff-Nahoum S (2017) The evolution of the host microbiome as an ecosystem on a leash. Nature 548:43–51

Gillies D (2019) Commentary on Kate E. Lynch, Emily C. Parke, and Maureen A. O'Malley: 'How causal are Microbiomes? A comparison with the *helicobacter pylori* explanation of ulcers'. Biol Philos 24:56.

Gomez-Lavin J (2019) Why expect causation at all? A pessimistic parallel with neuroscience. Biol Philos 34:61

Greslehner GP, Lemoine M (2020) A dual decomposition strategy of both microbial and phenotypic components for a better understanding of causal claims. Biol Philos 35:1

Griffiths PE, Pocheville A, Calcott B, Stotz K, Kim H, Knight R (2015) Measuring causal specificity. Philos Sci 82:529–555

Halfvarson J, Brislawn CJ, Lamendella R et al (2017) Dynamics of the human gut microbiome in inflammatory bowel disease. Nat Microbiol 2:17004

Hooks KB, O'Malley MA (2017) Dysbiosis and its discontents. mBio. https://doi.org/10.1128/mBio.01492-17

Hooks KB, Konsman JP, O'Malley MA (2019) Microbiota-gut-brain research: a critical analysis. Behav Brain Sci 42:e60

Humeau Y, Doussau F, Grant NJ, Poulain B (2000) How botulinum and tetanus neurotoxins block neurotransmitter release. Biochimie 82:427–446

Johns NI, Blazejewski T, Gomes ALC, Wang HH (2016) Principles for designing synthetic microbial communities. Curr Opin Microbiol 31:146–153

Johnson EL, Heaver SL, Walters WA, Ley RE (2016) Microbiome and metabolic disease: revisiting the bacterial phylum Bacteroidetes. J Mol Med. https://doi.org/10.1007/s00109016-1492-2

Klassen JL (2020) Ecology helps bound causal explanations in microbiology. Biol Philos 35:3
Kumar N, Browne HP, Vciani E et al (2019) Adaptation of host transmission cycle during *Clostridium difficile* speciation. Nat Genet 51:1315–1320
Kumar M, Ji B, Zengler K, Nielsen J (2019) Modelling approaches for studying the microbiome. Nat Microbiol 4:1253–1267
Lawson PA, Citron DM, Tyrrell KL, Finegold SM (2016) Reclassification of *Clostridium difficile* as *Clostridiodes difficile* (Hall and O'Toole 1935) Prévot 1938. Anaerobe 40:95–99
Lean CH (2019) Can communities cause? Biol Philos 34:59
Lloyd-Price J, Abu-Ali G, Huttenhower C (2016) The healthy human microbiome. Genome Med 8:51
Lynch KE, Parke EC, O'Malley MA (2019) How causal are microbiomes? A comparison with the *Helicobacter pylori* explanation of ulcers. Biol Philos 34:62
Menzies P (2004) Difference making in context. In: Collins JD, Hall N, Paul LA (eds) Causation and counterfactuals. Cambridge, pp 139-180
Moran NA, Sloan DB (2015) The hologenome concept: helpful or hollow? PLoS Biol. https://doi.org/10.1371/journal.pbio.1002311
Neal JP (2019) When causal specificity does not matter (much): insights from HIV treatment. Philos Sci 86:836–846
O'Malley MA, Parke EC (2018) Microbes, mathematics, and models. Stud Hist Philos Biol Biomed Sci 72:1–10
Oftedal G (2020) Problems with using stability, specificity, and proportionality as criteria for evaluating strength of scientific causal explanations: commentary on Lynch et al. Biol Philos 35:26
Olesen SW, Alm EJ (2016) Dysbiosis is not an answer. Nat Microbiol 1:16228
Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD), 17th October 2019. https://omim.org/
Perez-Muñoz ME, Arrieta M-C, Ramer-Tait AE, Walter J (2017) A critical assessment of the 'sterile womb' and 'in utero colonization' hypotheses: implications for research on the pioneer infant microbiome. Microbiome 5:48
Petersen C, Bell R, Klag KA et al (2019) T cell-mediated regulation of the microbiota protects against obesity. Science 365:eaat9351
Queller DC, Strassmann JE (2016) Problems of multi-species organisms: endosymbionts to holobionts. Biol Philos 31:855–873
Raman AS, Gehrig JL, Venkatesh S et al (2019) A sparse covarying unit that describes healthy and impaired human gut microbiota development. Science 365:eaau4735
Schneider T (2020) Microbial activities are dependent on background conditions. Biol Philos 35:23
Schulfer AF, Battaglia T, Alvarez Y et al (2018) Intergenerational transfer of antibiotic-perturbed microbiota enhances colitis in susceptible mice. Nat Microbiol 3:234–242
Skillings D (2016) Holobionts and the ecology of organisms: multi-species communities or integrated individuals? Biol Philos 31:875–892
Skillings D (2019) Trojan horses and Black Queens: 'causal core' explanations in microbiome research. Biol Philos 34:60
Vásquez-Castellanos JF, Biclot A, Vrancken G, Huys GRB, Raes J (2019) Design of synthetic microbial consortia for gut microbiota modulation. Curr Opin Pharmacol 49:52–59
Venturelli OS, Carr AV, Fisher G et al (2018) Deciphering microbial interactions in synthetic human gut microbiome communities. Mol Syst Biol 14:e8157
Vrancken G, Gregory AC, Huys GRB, Faust K, Raes J (2019) Synthetic ecology of the human gut microbiota. Nat Rev Microbiol 17:754–763
Waters CK (2007) Causes that make a difference. J Philos 104:551–579
Watkins A, Bocchi F (2020) Pathogen versus microbiome causation in the holobiont. Biol Philos 35:7
Woodward J (2003) Making things happen: a theory of causal explanation. Oxford University Press, Oxford
Woodward J (2018) Explanatory autonomy: the role of proportionality, stability, and conditional irrelevance. Synthese. https://doi.org/10.1007/s11229-018-01998-6