Dear Prof Otto and kind reviewer,

Many thanks for this opportunity to further clarify and improve our work. With this submission, we feel we have made another step forward.

Please see our responses to the comments and queries below.

We look forward to your feedback.

With warm regards,
The Authors

REVIEWER COMMENTS & OUR RESPONSES

The revised version is a completely re-written manuscript rather than a reply to the reviewer's comments. Track change mode was only used in parts of the manuscript which made it difficult to recognize changes.

We apologise for this error in incompletely tracking the changes. This time around we clearly indicate all changes in content and terminology: with new and added text in blue and a strikethrough for deleted text. (Some very small edits related to spelling and grammar we don't always indicate to keep the revision document as simple as possible). We believe that addressing some of the comments in the previous review round warranted restructuring the paper - and that this has resulted in a more compact and clearer manuscript. Many apologies if we did not convey this sufficiently clearly in the previous submission in our letter and response.

We group the following comments:

- Comments were only partially addressed and important questions such as the suitability of reuse bottle systems as 'positive tipping element' remain open.
- The tipping framework used in this manuscript differs from the cited tipping point framework by Lenton et al as (i) Lenton's framework always refers to 'positive' in the sense of more sustainable while the manuscript partly describes a negative social tipping point,

We agree that the second tipping episode is a step backwards with regards to having a positive impact and that a name change to 'negative social tipping point' is more appropriate. We have made this change in the document and in the figures.

We have grouped the following comments due to a common cause or concern expressed by the reviewer for clarity and brevity:

- (ii) terms such as balancing feedback loops are used differently across the two frameworks. This deviation needs to be made very clear in the text.
- RE: Figure 6: The authors state that 'The balancing feedback loops refer to the dynamics that stabilise the new system.' Lenton et al (2022) or the Global Tipping Point Report, use a different definition: balancing feedback loops stabilise the old system. From Lenton et al. (2022): 'negative feedbacks maintaining the initial state', 'Existing regimes, whether social, technological or ecological, are stabilised by damping feedbacks that resist change and restore the status quo' or 'balancing negative feedbacks maintaining the initial state'. A 'balancing feedback loop' as a dynamic that stabilizes a new steady state can be introduced, but then it needs to explain that this is a new addition to the existing PTP framework and that it differs to the existing understanding of 'balancing feedback loops'.
- From Lenton et al. (2022): 'To bring a system to a tipping point typically requires some forcing – that is, a change in boundary conditions – in a direction that weakens balancing negative feedbacks maintaining the initial state and/or strengthens reinforcing positive feedbacks that amplify change (Meadows, 1999, 2008).

Balancing (or mathematically 'negative') feedback loops simply act to maintain the current state (attractor) of a system regardless of whether it is a desirable or undesirable state of the system or whether it is the old state or the new state - i.e. balancing feedback loops stabilise both the old state and the new state of a system, but they can be different balancing feedback loops in each case. Previous work emphasised the role of balancing/negative feedback loops in stabilising the initial state (and of amplifying/positive feedback in propelling change) but did not elaborate the role of balancing/negative

feedback in stabilising the new state. Here we elaborate on that (and how the balance of feedback shifts through different phases) because it is essential to understanding the overall chronology and multiple episodes of tipping in our case study. This is not a change in the well established terminology and widely accepted meaning of damping/balancing/negative feedback in cybernetics and systems thinking. But it is an elaboration of the overall positive tipping points framework that we now make explicit. It necessitates clearly identifying and differentiating balancing feedbacks that stabilise the old state and the new state which we now do more clearly in our figures and narrative.

-> Balancing feedbacks are exclusively seen as something undesirable that prevents 'positive'/desirable change.
This is an incorrect statement and a misinterpretation of the previous work on PTPs (which may have been insufficiently clear). It is important to stick to the widely accepted and more general (mathematical) meaning of balancing/damping/negative feedback loops, which refers to a general dynamic that keeps things in place/ stable. We have sought to clarify that in our revisions.

If the authors use 'balancing' as something so fix the new steady- state (e.g. something desirable as pool reuse systems), they need to make it very clear in the text that they deviate from Lenton's and the PTP community's definition of 'balancing'. If the authors followed Lenton et al (2022), an example of balancing feedbacks loops would be on p. 17, l. 420: lack of leadership and difficulties to align prospective partners as balancing feedback loop that prevents a tipping towards a more sustainable reuse pool system.
We are not deviating from our previous work or widely accepted definitions in systems thinking, although we are elaborating the positive tipping points framework somewhat. Given the confusion the reviewer is expressing we have chosen to carefully elaborate on the role of balancing feedbacks on both sides of tipping (pre- and post-tipping). That is: already in section 2.3 we indicate clearly that we regard destabilisation to be caused by the weakening of feedback loops that keep the current system in place, but that balancing loops also have a role post-tipping to keep the new systems state in place, and we have brought the remainder of the document in-line with this.

From the abstract: "Building on current research on positive tipping points, our case study demonstrates opportunities to create an environment for change, the role of reinforcing feedback loops in accelerating sustainable transitions, and successful interventions. However, the case also demonstrates the threat of destabilisation of newly created systems as a result of the emergence of competing technologies, in this case single-use plastic bottles. Unsuccessful efforts to stop this, included the introduction of a reusable plastic bottle and a failed policy intervention that rushed into a solution that instead accelerated the change it was designed to prevent."
-> The authors frame this as two separate tipping processes. An alternative interpretation of the German bottling system could be that recyclable single-use plastic bottles are a "shallow outcome", following the update of Lenton et al' (2022) framework in the Global Tipping Point Report (2023), section 4, p. 14.
We find it insightful to highlight two separate tipping episodes. Especially given that Episode 2 is not an improvement when it comes to environmental impact, but a step back. With the remainder of the changes (see also below), we trust we have sufficiently addressed this.

We have grouped the following comments:

- Regarding re-use versus recycling systems: The authors refer to the 'positive' tipping points framework by Lenton et al. Positive tipping in this framework refers to a transformation towards a sustainable future. 'Positive tipping points offer hope for accelerating change to avert climate and ecological emergency', state Lenton et al (2022). The Global Tipping Point Report (2023) defines on page 8, part 4: 'What is considered normatively 'positive' or beneficial, and by whom, is highly debatable. In principle, tipping points may be considered positive either: a) where they reduce the drivers of 'negative' Earth system impacts such as greenhouse gas emissions or deforestation, for example in a rapid shift to renewable energy or alternative food proteins; or b) where they improve the social foundations of sustainability.
- As such, the authors' statement that 'the quantitative assessment of which of these states is the 'most sustainable' we consider out of scope for this work' is worrying given that they use the 'Positive Tipping Framework' by Lenton as basis of the study. This question needs to be addressed in a short discussion of what the study defines as 'positive' and why the re-use bottle system is a valid example for a positive tipping element.

- In the authors' answer to my comment on Fig 5, I take that the 'desirable' state of 'better' is the level of reuse. That makes sense, but again it needs a short discussion why pool reuse of glass is 'better'/more desirable than recyclable PET.
- RE: The rapid establishment… > Again, the authors have not addressed my comment which is a link to the Stefanini paper and the question whether reuse systems are a useful example for positive tipping processes.

Taking the comments above together: we acknowledge that our response in the letter was somewhat different from what was in the final manuscript text. That is: in the manuscript we do not refer to "quantitative assessment" explicitly. However, the essence of our argument, as explained in the third paragraph of section 3.1, is that the number of reuse cycles and the distance traveled (40-50 uses and 260 km), far exceeds the break-even point (3-10 uses and 500 km) where reuse is thought to be more sustainable than recycling, according to Coelho et al., 2020; DUH, 2014b, EMF, 2023 and UBA, 2016 - as referenced in the document. It can therefore be expected that this system has less environmental impact than single-use alternatives. We therefore believe it is a reasonable assumption to frame the first tipping episode as a Positive Tipping Episode, especially in relation to the later introduction of plastic single-use and recycling. We do, however, regard doing a (retrospective) LCA or similar assessment as outside the scope of this study. To avoid confusion, we have removed the footnote, and added a sentence to the paragraph to make explicit that the case falls well within the break-even limits in the third paragraph of section 3.1. We have also amended the titles of the images: where Fig.4 is now clearly labelled as a "positive tipping point: and Fig. 5 as a "negative social tipping point."

RE: A reinforcing…
Again, there is a fundamental difference between Lenton et al.'s (2022)/the Global Tipping Point Report's (2023) and the authors' interpretation of 'positive tipping points'. This needs to be made clear and well described at some point in the paper. In their reply, the authors state that 'a reinforcing feedback loop can operate to amplify change in either direction.' This is not the case in how Lenton et al define 'reinforcing'. There, 'reinforcing' always refers to a positive (e.g. more sustainable) change.
There is not a fundamental difference, although our previous work may have been insufficiently clear. While the examples of reinforcing feedback given in previous studies may have always been in the normatively 'positive' direction, reinforcing feedback is simply an alternative label for mathematically 'positive' feedback and mechanistically it can always operate in either direction. (For example, just as there are 'increasing returns' to increasing adoption, there are decreasing returns to abandonment. To be more specific: the more people who abandon petrol/diesel cars and the more petrol/diesel stations close and mechanics retrain, the more inconvenient it is to (still) own a petrol/diesel car - and thus the stronger the incentive to abandon the technology (a reinforcing feedback).) Thus reinforcing feedback cannot be conflated with normatively 'positive' change, it has to be retained as a more general mechanistic concept. Normative 'positive' and 'negative' judgements can be assigned to the different directions in which a particular reinforcing feedback is operating, and that is what we now try to do in a clear and consistent way in the manuscript.

Research papers should start with an introduction into a topic. The re-written version does not provide a proper introduction but jumps straight into the research question without providing any contexts.
We apologise for that oversight and have now added a new first paragraph to the paper that clearly introduces the topic of circular economy and its relevance.

p.3, l. 80. How does today's situation regarding soap bars compare to the historic re-use bottle system case study?
We have opted to drop this example as it was causing confusion - it was intended to indicate that the solution space is not singular: e.g reuse is not the only option for introducing more sustainable options. In some cases, the product itself can also be redesigned or reformulated, realising further environmental benefits. But both reusable bottles and soap bars were once replaced by other more resource intensive alternatives.

p.10, l.251 The paragraph is confusing and was clearer in its initial form. 'Thus far, the framework has been applied to the food and land use system' sounds as if FOLU was the only application of the PTP framework so far which is incorrect. And how have food and land use systems 'provided initial insights into the adoption of renewable energy and electric vehicles'?
Apologies, this was a mistake. We have now altered the sentence and included reference to Meldrum et al. (2023) on the energy system and to the Global Tipping Points Report 2023.

Fig. 4. Again, it should be mentioned that R1 and R2 follow a different logic than Lenton et al. (2022). Following Lenton et al, R1 would be a weaking of balancing feedbacks (balancing as keeping the old bottle system and weakening would refer to the loss of these bottles in WW II). The same applies to R2. R5 belongs to the next graphic as it has nothing to do with the tipping from company bottles to pool reuse.

Thanks for these excellent suggestions. We agree that R1 and R2 were wrongly described. The figure has been updated accordingly. However, we have preserved R5 as these developments already take place early in the timeline: it indicates that already early on the possibility of tipping in a different direction existed, and it emphasises the overlap of the two tipping episodes when it comes to certain developments. Also, R5 and B4 interact as described in the narrative (post-tipping for Episode 1) - so it is appropriate to mention it here. And: it also shows that the risk of single-use already loomed early - and that such signs can be taken as early warning signs.

Fig 5. The entire example would not be defined as 'positive tipping point' following Lenton et al (2022) or Otto (2020) but rather be a 'negative social tipping point' as (according to the authors, see comment on sustainability above) PET reuse systems are less sustainable than glass reuse bottles.

We agree (and never thought of this second tipping episode as 'positive') - it was a negative tipping point from reusable glass to single use recycled plastic bottles and now clearly labelled as such.

p. 32, l. 758. Again, there is a reference to PTPs that assume a shift to a more sustainable state following Lenton which, according to the authors, is not the case in Fig. 5

This was a mistake that has now been corrected.