This manuscript uses a framework developed for examining the development of small groups, that has been further developed in the literature to explain the development of sports teams, to explore the development of collaborative PhD research groups. This is an interesting application of an existing framework that offers considerable potential to better understand how, why and when successful collaborative research can be achieved. The data set is centred on case study analysis of an interdisciplinary doctoral education pilot project (the Majakka project) using the sports teams framework, data from “journey mapping” by members of the doctoral project and the wider programme in which it is embedded, as well as interviews and workshops with the participants during the project.

The potential strength of this paper is the unique data set provided by the case study that has considerable value for improving understanding of how small collaborative research groups develop, can be supported, and what they can achieve in terms of interdisciplinary research and education outcomes. However, additional explanations and clarifications, some further data analysis and paper restructuring are needed to maximise the explanatory potential of the data and its findings.

Firstly, in the introduction and throughout the paper, all key terms need to be clearly defined and explained: inter and transdisciplinary, collaborative learning, self-directed peer learning, social capital, transformational leadership, culture-building, group storming, active learning goal mapping etc.

Some more robust arguments as to why the sports team framework was chosen would be beneficial i.e. did this come from the authors past experiences? The background to this needs quite some reworking e.g. paragraph 3 of the introduction: "The sports team framework holds high potential for research groups and doctoral education development as well: it can improve and create new processes of peer learning and collaboration, and
clarify the need for diversity, support, and shared motivation”. The framework itself cannot improve and create new processes etc. but it can be used to analyse how these emerge. The research questions also require some refinements as in the manuscript in its current form they are not convincingly addressed.

Figure 1 should be introduced and much more fully explained, either in the introduction or in a new section 2, that specifically reviews the literature, develops the framework and places it in its new context of collaborative (and interdisciplinary) research. (Also, the references related to Figure 1 are missing from the reference list).

In the methods section, a more thorough description of the journey mapping workshops (who was there, what did they do, and when) and more specific details of the workshops, interviews and surveys should be provided along with how the data were collected (recorded, notes taken, photo of images and figures created) and managed (transcribed, categorised, sorted etc.).

The results and discussion do not seem particularly well aligned to the Framework that has been proposed in Fig 1. Paragraph 1 seems unnecessary and could be removed. It would then make most sense, if the sub-headings of each element of the framework were aligned with the key elements of each section.

Under 3.1 Forming, the subheadings of i) Boundaries, ii) Dependencies, and iii) Leadership, would be much more logical as these words appear in Tuckman’s original description. Some more detailed analysis of research boundaries (how they are set up, what constrains them and how they are defended in an interdisciplinary setting would be extremely interesting). Under Dependencies, you can talk about how roles are allocated and human resources selected to build a “team”. Here I would expect some discussion of how different disciplinary skill sets are brought together. Under Leadership you can bring together all the material scattered throughout the analysis and discussion on leadership, how it was structured and how it worked.

Under 3.2 Storming, the sub-headings of e.g. Conflicts/Polarisation, and Facilitation. Figure 2 needs much more thorough explanation here. What do the colours relate to in Fig. 2a and why do the axes cover scale and qualitative/quantitative? In 2b) what data is this based on? Could you add a quantitative element, i.e. what was the risk that the greatest number of students were concerned about? Why is maternity or paternity leave perceived as a risk? This needs to explained and contextualised (for the Finnish setting).

For 3.3 Norming, sub-headings of Cohesiveness, New standards, and New roles, would shape a very interesting discussion. The term “culture” seems vague and does not fit with the sports management framework. More details on the mechanisms and how they were operationalised to achieve the “Cohesiveness” would be needed. E.g. how was a commitment of 5% of work time for the common good implemented and monitored? How was psychological safety achieved and what is the evidence of this? How was learning
across different disciplines achieved? i.e. through joint fieldwork, shared data collection, different analysis approaches of the same data sets etc. What new standards of evaluation have emerged? How have new roles within the research groups evolved or been developed?

Under 3.4 Performing, sub-headings of e.g. Interpersonal structure, Performance, Activity. Under interpersonal structure, you have great material about the interactions between more advanced researchers and early career scientists. More details and explanations of the findings of the interactions between academic and non-academic partners would be very valuable especially in terms of the challenges to collaboration that these may present. Figure 3 needs considerably more explanation. Explain the scales – what exactly is meant by enabling, reinforcing, indivisible, constraining, counteracting, cancelling? What are the reasons why some students start the process higher on the scale than others? It is interesting that COVID affected students differently and this seems worthy of some discussion and explanation.

Provision of supporting data for any statements made is essential. E.g. page 11, “Journey mapping results highlight how doctoral students enjoyed the most those papers where they were working in an interdisciplinary team”. “At the beginning of their studies, the position of not being the corresponding author was found to increase their belongingness and psychological safety”. “Regarding the growth as a researcher, ownership of their work, and an established role in the community, the 3-to-12-month research visits were found valuable.” From where have these findings been obtained and how many students do they relate to (e.g. interview with student No. X in 202X; workshop no X).

In section 3.5, some discussion of whether the authors feel a fixed duration project in which all researchers start and finish together would generate a stronger collaboration than on ongoing, rolling project in which research start and finish at different times would be interesting.

Section 3.6 seems unnecessary and could be removed.

Additional comments

In the current form of the manuscript, the interdisciplinary aspects are not prominent. Therefore, consider changing the title, to replace “interdisciplinary” with “collaborative”.

Provide more details of the case study to give context e.g. where did funding come from and was collaborative research a specific objective of the funders, what was the background to the project, did the senior researchers have prior experience of working together etc.