We thank the reviewer for their valuable feedback. Below you find our detailed response (in blue italic).

General comment:
This study uses bioassay experiments in the Weddell Sea and in the Amundsen Sea Polynya to study the effects of increasing Fe and temperature conditions on natural phytoplankton communities. Given the current predictions about upcoming changes in the Southern Ocean, it is critical to understand what effects these changes may have on natural communities. The manuscript is very well written, and this study is greatly strengthened by the extent of trace metal results reported (including both dissolved and particulate data).
However, several points could be improved. First, the differences in initial conditions could be discussed in more details (coastal vs offshore, difference in macronutrients, etc.).
We will include a comparison of the initial conditions for the different bioassays to the Results (including information from Material and Methods and Table 1), as well as include depth profiles for the stations in the Supplements to provide a general overview of the initial differences between bioassays. Lines 460 – 461 will be moved to the regional differences section in the Results.

"3.1 Sample site characteristics
The in-situ temperature was below zero for all bioassays, with lowest values for Amundsen Sea bioassay A2 and Weddell Sea bioassay W2 (-1.6 °C and -1.4 °C, respectively, compared to -0.6 °C and -0.3 °C for A1 and W1). The daily average irradiance at sampling depth on day of sampling was lowest for A1 and A2, i.e., < 6 µmol quanta m^{-2} s^{-1}, compared to 18 and 98 µmol quanta m^{-2} s^{-1} for W1 and W2. Dissolved inorganic macronutrient concentrations were relatively comparable between bioassays, except the silicate concentration in W1 being ~20 µM lower than for the other bioassays (but still far from limiting). Initial dFe concentrations in the Weddell Sea were lower compared to the Amundsen Sea (Table 1), as were dMn concentrations (Fig. 2). Bioassay A1 had the highest Chl a concentrations (sampled within the ASP), followed by W1. Both bioassays also had the highest share of >20 µm Chl a. The Chl a concentration of A2 was almost exclusively made up of < 20 µm sized phytoplankton (98% of total Chl a, Table 1). Flow cytometry derived phytoplankton abundances were highest for the Amundsen bioassays. The photosynthetic efficiency Fv/Fm at the start of the incubations was 2-fold lower for the Weddell Sea bioassays compared to the Amundsen Sea bioassays (i.e., 0.3 vs 0.6 r.u., respectively). The station for bioassay W2 was closest to the coast, followed by A1, A2 and W1, however distance to land did not seem to have a major impact on either phytoplankton community composition, or nutrient concentrations."

Further, the photophysiological results (Fv/Fm and more?) should be referred to more often to strengthen some statements.
We will refer to the F_v/F_m results more consistently, e.g. in lines 594 and 658 of the original manuscript. See also our reply to the comment on the topic by Reviewer 1.

You could also add more information on the biological results that were briefly mentioned (e.g., Chl:C ratios) to confirm your hypothesis.
We thank the Reviewer for their suggestion. We will add the following at 666 of the original manuscript: "The relatively low Chl:POC ratios in the Weddell Sea bioassays (average over all treatments 0.003 $\pm$0.003 vs 0.008 $\pm$0.002 for the Amundsen Sea bioassays) may indicate stronger Fe limitation, since Fe limited cells are known to have a lower Chl:POC ratio compared to non-limited cells (Moore et al. 2007, Deep-Sea Research Part II: Topical Studies in Oceanography, Vol 54 (5-7))."

The figures could also be improved with bigger panels and the statistically significant differences displayed to help result visualisation and interpretation.
We will adapt the figures as suggested (e.g. increase panel size and change the colour for the TF treatment). Adding significance indicators to the figures will in our opinion give the impression that we used pairwise comparisons. However, the test used is not based on pairwise comparisons, hence

we cannot add meaningful indicators for statistical significance to the figures. We will add a table to the supplements providing the statistical results for the test used.

Finally, some typos need fixing throughout the manuscript (use of abbreviations and then not, some references to fix, some commas near period).
We carefully checked the manuscript and tried to make sure the revised manuscript is thoroughly corrected.

Below, I have provided line comments which I hope will help improve the manuscript.
Line comments
-L68 This first long intro paragraph could be split for easier reading.
Thank you for bringing this to our attention. We will add paragraph breaks in line 65 and line 78 to facilitate reading.

-L71 fix ref
We will double check all references and fix them, where necessary.

-L105-107 This info could be removed from the introduction as you describe it well in your method.
We will shorten this sentence in the introduction to: "Our bioassay treatments comprised a full factorial combination of Fe and temperature increases."

-L55-120 Considering how many times you speak of Mn in your results and discussion, I wonder if you should talk about it in your introduction and describe its essential roles in phytoplankton.
We will add to the introduction line 77 onwards: "
Besides Fe and temperature, there are also other factors, e.g., other bio-essential metals (Mn, Co, Ni, Cu and Zn), where notably Mn has been shown to be (co-)limiting in the Southern Ocean (Wu et al; 2019, Browning et al., 2021, Balaguer et al.; 2022, Hawco et al.; 2022). Mn is essential for phytoplankton photosystems (Raven et al., 1990) and a co-factor for enzymes dealing with oxidative stress (Wolfe-Simon et al., 2005, Journal of Phycology, Vol 41 (3)). Moreover, light is another major limiting factor for phytoplankton growth in Southern Ocean (e.g. van Oijen et al.; 2004, Journal of Plankton Research, Vol 26 (8), Strzepek et al.; 2019, Proceedings of the National Academy of Sciences, Vol 116 (10), Vives et al.; 2022, Journal of Marine Systems, Vol 234, Latour et al.; 2023, Limnology and Oceanography Letters, Vol 9 (2))."

-Figure 1: I think you are missing in your result and discussion a general description of your different sites. Both regions are separated but within them you present results from coastal and (almost?) offshore data? This may greatly impact your results too.
We are hesitant to put too much emphasis on regional differences, given that we only performed two experiments in each region, and given that light conditions differed. We will incorporate a section on initial characteristics of the locations as a first section in the Results, please see our response to your first general comment.

-L137 what is 'PVDF'?
PVDF is a type of plastic (Polyvinylidene fluoride). We will spell it the first time we use it.

-L143 ASP previously defined
Thank you for noting. We will change accordingly.

L146 Why this difference in duration?
The difference in duration between the Amundsen Sea and Weddell Sea bioassays was due to logistical constraints (we will add this to the Materials and Methods section). Amundsen Sea bioassays were conducted before Weddell Sea bioassays and unfortunately, time was very limited.

Given the small effects observed in the Amundsen Sea bioassays, we decided to adapt light conditions and extend the duration of the Weddell Sea bioassays to allow for a more comprehensive assessment of the treatment effects.

-Table 1: could the depth profiles of each station be presented in SI so we can better visualize the different initial conditions?
We will add depth profiles for each station to the Supplements.

-L183 typo
Thank you for pointing this out, we will change the text accordingly.

-L191 How were these light levels achieved in your experiment?
Light levels were adapted using neutral density screens, we will clarify this in the Material and Methods following line 193.

-L245 Can you clarify if you expect your pTM measurements to include lithogenic material as well?
Yes, our pTM measurements include lithogenic material. We will add the following text to the Material and Methods section 2.5, following line 262: "The lithogenic fraction and concentration of pFe and other particulate metals discussed was determined by assessing the ratio between the particulate metal of interest and particulate aluminium (pAl), assuming all pAl originates from crustal material using the approach described in more detail in van Manen et al (2022). For example, we are using the observed pFe/pAl ratio in the samples and the known crustal ratio of 0.21 mol mol^{-1} (Taylor and McLennan, 1985) to calculate the lithogenic pFe fraction and concentration, see supplement xx for more details."
And: "The EDTA oxalic acid wash used on particulate samples prior to filtration should effectively remove surface-bound metals, also minimizing the authigenic Fe fraction" following line 246.

In the Results (following line 435), we will add that lithogenic particles provided a consistent background that did not affect observed changes between the treatments, and we will add the respective data to the Supplements.

-L248 at the end of the experiment?
Yes, all filters collected (stored at -20°C) were taken back to the NIOZ at the end of the cruise to be processed and analysed.

-L298 could you add info on the type of measurements (flash sequence etc…)
The measuring light frequency used was set to level 5 (25 Hz) with an intensity of 8, the SAT-pulse width was set to 0.8 seconds and the far-red pulse width was set to 10 seconds, with intensities of 10 and 6, each. We will add this information to the Material and Methods, section 2.8.

-L303 Did you also rinse the cuvette with the (filtered) sample itself?
No, the cuvette was not rinsed with the sample before measuring, but UP water was removed from the cuvette by shaking and placing the cuvette upside down on lint-free paper towels to remove last droplets. We also measured technical replicates per biological replicate, between which the cuvette was not cleaned with UP water. There was no significant difference in Fv/Fm between the different replicates (non-parametric Kruskal-Wallis ANOVA, p = 0.95), indicating that rinsing with UP water did not impact reported values. We will add this to the Materials and Methods section.

-L307 Can measurements of the functional absorption cross section of PSII be derived from this instrument too?
Unfortunately, the functional absorption cross section of PSII cannot be determined with the instrument used.

-Figure 3: it would be helpful to have the significant differences displayed on the figure.
Adding significance indicators to the figures will in our opinion give the impression that we used pairwise comparisons. However, the test used is not based on pairwise comparisons, hence we cannot add meaningful indicators for statistical significance to the figures. We will add a table to the Supplements providing the statistical results for the test used.

-Figure 4: please fix the x-axis label.
We will change the x-axis label accordingly.

-L594-596 You could also refer to Fv/Fm to support this statement.
The reviewer is correct, we will add Fv/Fm in this sentence (here underlined for clarity): "However, since phytoplankton abundances, F_v/F_m and Chl a concentrations were not higher in T treatments compared to the control, and since phytoplankton requires less Fe at higher temperatures (Jabre & Bertrand, 2020), this is less likely."

-L596 It is hard to visualize this trend because of the different scales in Figure S2. It would be good to refer to the lower Si initial conditions at W1 compared to A1 and A2. Also, the t0 of Figure S2 does not match your value in Table 1 for A1 and W2? Same for NOx of A1.
W1 showing the strongest decline in Si concentrations should still be visible in Fig. S2, given that the tick marks have the same distance despite the different scales. We will add to the caption that tick marks represent the same concentration intervals aiding comparison.
We are sorry for the typos in Table 1. The correct values are the ones displayed in figure S2, we corrected the Table.

Bioassay	Station	Lat. (°S)	Long. (°W)	Temp. (°C)	Salinity (psu)	Irradiance (µmol quanta m^2 sec^{-1})	Si (µM)
A1	31	73.50	116.50	-0.6	33.99	5.0	77.9
A2	52	72.00	118.42	-1.6	33.89	3.1	77.5
W1	17	65.00	000.00	-0.3	33.90	17.7	58.3
W2	36	70.08	011.08	-1.4	33.82	97.6	71.8

Bioassay	PO_4 (µM)	NOx (µM)	Fe (nM)	total Chl *a* ($\mu g\ L^{-1}$)	< 20 µm Chl *a* (%)	Phyto ($x10^3\ mL^{-1}$)	F_v/F_m r.u.
A1	1.8	24.3	0.28	3.0	42	8.4	0.6
A2	2.0	28.2	0.10	0.4	98	7.1	0.6
W1	1.6	24.0	0.05	1.5	24	5.6	0.3
W2	1.9	27.9	0.03	0.6	65	4.4	0.3

-L610 How did you calculate the growth rates?
Phytoplankton net growth rates are based on an exponential trend line per replicate for each treatment per bioassay and were originally only determined for total phytoplankton abundances. We have now included phytoplankton group specific growth rates. We will include how phytoplankton net growth rates were calculated in the Material and Methods: "Phytoplankton net growth rates were calculated using exponential trendlines. For total abundances, the full incubation period was taken into account (i.e., day 1 - 6 for Amundsen Sea and day 2 - 8 for Weddell Sea bioassays). Starting abundances were taken prior to filling of the cubitainers and hence not taken into account. For the phytoplankton group specific rates only those time points (>3 but most often 4-5 time points) with a consecutive increase in abundances were selected."

-L642 are you referring to the 3 experiments? In your results I think you mentioned differences to the C and not the T.
We calculated the difference in pCu:POP ratios for the Fe treatments compared to the control to avoid too many comparisons. However, statistical results indicated that Fe addition treatments showed differences in the pCu:POP ratios compared to both treatments without Fe addition (i.e. control and T treatment), and both the control and T treatment showed higher pCu:POP ratios. We are referring to all 3 experiments, however, the effect was most visible for bioassay W2 and will change the text to make this clearer (changes underlined for clarity): "For example, cellular Cu requirements increase under Fe limitation (Schoffman et al., 2016), which could explain the higher pCu:POP ratios in the C and T treatments compared to the Fe addition treatments in all bioassays analysed (Fig. 3)."

-L659 maybe I missed it, but you should also discuss that some bioassays have an increase in Chl in the control compared to t0 (A1, A2, W2) while W1 did not and why is that.
We will add discussion on the topic when we relate the results of W1 with the bioassay S54-65 from the Viljoen et al. (2018) study: "The location the seawater for bioassay W1 was taken has similar coordinates as bioassay S54-65 in a study by Viljoen et al. (2018). These authors sampled 3 weeks later (different year) and at a comparable depth (30 m vs 20 m in our study) and found largely similar responses by the phytoplankton to dFe addition, i.e., total Chl a increased by ~2 µg Chl a L^{-1} and diatoms dominated the phytoplankton community. In contrast to W1 but comparable to our other bioassays, total Chl a concentration in bioassay S54-65 (Viljoen et al. 2018) increased in the control over the duration of the bioassay. The lack of increase in Chl a in the control (and T) treatment of W1 might be explained by a lower in-situ dFe for W1, indicating a stronger limitation of dFe. At the same time, POC (and < 20 µm Chl a) concentrations did show an increase over time in the control (and T) treatment of bioassay W1. Moreover, bioassay W2, with even lover starting concentrations of dFe, showed an increase in Chl a over time for the control. Given the lowest dMn concentrations in W1, it might be that dMn and not (only) dFe was limiting the production of reaction centres (Raven et al.; 1990), resulting in Chl a concentrations to not increase. Given the increased requirement for Mn under low Fe (Peers & Price; 2004), Fe addition may have relieved Mn limitation in the Fe addition treatments slightly, resulting in the observed increase of Chl a in those treatments."

-L661 you could also refer to your Fv/Fm results here.
The Reviewer is correct, we will add Fv/Fm to this part of the Discussion:
"The Weddell Sea bioassays exhibited stronger Chl a accumulation, a stronger increase in F_v/F_m and increased phytoplankton abundances in response to Fe addition than the Amundsen Sea bioassays, which is likely due to the lower dFe concentrations (and hence higher degree of Fe limitation for the phytoplankton typical for the Weddell Sea) at the start of the incubations."

-L664 Could these results (Chl:POC) be presented in SI? So we can see the difference between treatments. Results of fluorescence per Chl may also be interesting in this context?
We will add the Chl:POC ratios as a Supplement, as well as F0 values from our F_v/F_m measurements.

-L666-669 Do you mean that in-situ phytoplankton were light-limited? It is expected which such low levels.
Indeed, we expect in-situ phytoplankton to be light-limited, as we indicated in line 670.

-L671 Figure 2 still shows a great depletion of the added Fe by day 3. Do you think it is due to phytoplankton uptake or could be sorption? POC does not show any change but Chl does although with n=1. Do you think another parameter could have become limiting after day 3?
The observed depletion of added Fe by day 3 could be attributed to several factors, but the most likely is a combination of precipitation and wall adsorption, and perhaps luxury uptake as explained in section 2.3. Even in UP water, the Fe concentrations decreased albeit less rapidly than in natural seawater (with more (particle) surfaces available), where the difference between Fe amended and

control treatments often became undisguisable based on concentrations alone (see section 3.1). This implies sorption of Fe due to its high affinity for surfaces plays a role, where we envision an equilibrium concentration (balance between sorption and dissolution) develops after Fe addition that is driven by the type and amount of (particle) surface area available. If there is also significant biological uptake, the equilibrium will re-establish, leading to desorption maintaining the values near equilibrium despite uptake (till all available Fe is used). As such, not much can be concluded from the dissolved concentration alone and hence, we switched to the practice of using $d^{57}Fe$ for the Weddell Sea experiments to better track the Fe addition as described in this manuscript.
The decline in dFe is what we would expect, so we should be careful not to overinterpret this feature. Regarding whether another parameter became limiting in A2 after day 3: It is possible that another parameter became limiting, but given that we measured all macro- and micronutrients, and none was drawn down to limiting concentrations, it is unlikely it was nutrient limitation. Light conditions were likely limiting, potentially causing POC levels not to increase (please also see our response to the next comment).

-L674 these low light levels are comparable to this study: https://doi.org/10.1002/lol2.10366.
Thank you for the reference. We will change lines 671 and following: "Although earlier studies reported positive responses of phytoplankton to Fe addition also under low light conditions (Viljoen et al., 2018; Alderkamp et al., 2019), the light intensities used for the low light treatment in those studies were still relatively high (i.e., 15 and 30 µmol quanta m^{-2} s^{-1}) and well above those in A1 and A2 (average of 3.4 and 1.5 µmol quanta m^{-2} s^{-1}). In addition to higher light levels, the lower initial dFe concentrations in the Ross Sea study (Alderkamp et al. 2019) compared to our study indicate a stronger Fe limitation and subsequently a stronger response to dFe addition. A recent study on Southern Ocean deep chlorophyll maximum phytoplankton responses to Fe addition (Latour et al., 2023) reported a Chl a increase at low light intensities (similar to our Amundsen Sea bioassays) and no change of POC (similar to bioassay A2) until light levels increased. This supports our suggestion that the low light condition in A2 was a determining factor for the lack of a response to dFe addition."

-L676 Do you think you could have seen a response if the experiment lasted longer? With the initial combination of very low temperature, light and biomass at that station.
Given that the light levels stayed low, we do not think extending the incubation would have made much diference. Both F_v/F_m and total phytoplankton abundances over time already showed levelling off at the day the experiment was ended. Only in case weather conditions had improved, leading to higher light levels, may we have seen a response.

-L677 Or maybe phytoplankton were just not Fe-limited at the A2 station, as supported by the high Fv/Fm. It is likely the cells were content with the slow growth rates and would only need more Fe if light levels were increased (refer to the paper mentioned in one of the previous comments + https://doi.org/10.1073/pnas.1810886116)
We agree that maybe phytoplankton in bioassay A2 were not limited by dFe concentrations as indicated in line 670 – 671 of our original manuscript. We will clarify better and add the following: "The high initial F_v/F_m value suggests that the phytoplankton may not have been limited by dFe (under these low light conditions) and would only require more dFe once light intensities increased again (Strzepek et al.; 2019, Vives et al.; 2022, Latour et al.; 2023). The small increase in F_v/F_m in the Fe addition treatments may suggest growth became dFe limited during the incubation (Fe-addition did show a significant effect on F_v/F_m at the last day of the incubations), despite the light conditions remaining low."

-L686 Do you mean for A1 only? It is hard to visualize this treatment effect in Figure 4.
The effect was stronger in bioassay A1, however also bioassay A2 showed a significant positive effect of Fe addition on Fv/Fm values. We will adapt the symbol-sizes in the figure to make differences easier to see.

-L696 Which figure are you referring to? I cannot see this in Figure 6.
We agree that for some phytoplankton groups, differences are hard to see in this figure and will increase the panel sizes of the figures to make it easier to see such differences. The effect of temperature on phytoplankton groups 19 and 22 can also be seen in the Supplemental data where all phytoplankton abundances for each day of the bioassays are given. Phyto 19 showed higher abundances in the T and TF treatments in bioassays A1 and A2 (not present in Weddell Sea bioassays) and Phyt 22 showed an increased abundance in the T and TF treatments for both A2 and W2. We will also include Phyto 3 here, since it showed increased net growth with temperature for bioassay A1.

-L700 change in what?
L700 in the original manuscript refers to changes in phytoplankton abundances. We will clarify this.

-L764 but POC results were not extensively described in this study.
L764 refers to how a change in phytoplankton community composition and size structure may affect organic matter fluxes in the Southern Ocean. This part of the Conclusions is now moved to the Discussion based on a suggestion by Reviewer 3.