Comment on hgss-2022-7
B. Owens (Referee)

This is a very interesting manuscript that describes the *Challenger* and *Gazelle* expeditions that were carried out in 1869-1871 during a time of remarkable evolution of science, particularly oceanographic sciences. Until recently the Gazelle expedition was largely ignored and is important that the two expeditions be described and compared.

The descriptions of the vessels, the naval and scientific parties and the subsequent data analyses and reports explained the differences in the scientific impact of the two expeditions. However, it would be good if the section on how the sailing instructions were formulated was expanded. It was also difficult to determine exactly what measurements and how many of them were obtained from each expedition.

Until recently naval hydrographic services, were typically charged with providing bathymetry, tidal amplitudes, surface current charts and meteorologic data that are used for surface navigation. Preparation of these charts for navigation did not require knowledge of the distribution of temperature below the surface. "Naval Surveyors" were presumably responsible for getting the data for these products. Making these observations for the hydrographic services would have been the responsibilities of Nares, Tizard and von Schleinitz during their earlier naval careers. The advent of submarine cables also would require knowledge about ocean depths and the composition of the bottom, but not sub-surface temperature data. Maury's Brussels 1853 conference laid out a uniform procedure for reporting only meteorological data. That is, it did not set up a procedure to report sub-surface oceanographic data or provide a rational for these observations. It appears that preparation of these navigation charts and "showing the flag" were the justification for the *Gazelle*. It is clear from the first sentence of the Thompson and Murray cruise report that sub-surface measurements were a primary objective of the *Challenger* expedition. The present manuscript discusses conversations between Thompson and Carpenter about making biological measurements at the sea floor, which explains the biological sampling. The Challenger subsurface temperature measurements were revolutionary. Was the focus on deep ocean measurements due to George Richards who was the head of the naval hydrographic service and also a member of the Royal Society? Maury (1855) did describe the surface currents and temperatures of the North Atlantic and noted the differences in temperature (and density) along the path of the Gulf Stream and attempted to explain how the currents were generated. Was the scientific rationale for
these sub-surface temperature measurements purely scientific exploration or was there some underlying basis to make these observations? The present manuscript mentions that a Royal Society committee prepared a report that was used to justify and set up the Challenger cruise, but there is no entry in the reference list to this report. Does this report set out the justification for the sub-surface measurements? Someone deserves the credit for proposing these observations and it would be good to make this clear in the present manuscript.

It is difficult to discern what and how many measurements were made from each expedition. The equipment for each ship is given, but how they were used is not clear. For example, both ships had hemp rope for sub-surface measurements. I expect that both ships used the rope for bottom depth measurements, and the Challenger used it to measure sub-surface temperature. I did look through the references to find out how the temperature measurements were made. It might be interesting to discuss this in this manuscript. Did the Gazelle make other sub-surface measurements? Specific volume measurements were made on both expeditions. I was unable to find in the references how the specific volume measurements were made? It appears from the Gould and Cunningham (2021) reference that these measurements were only made from surface bucket samples which is not surprising since the capability to collect sub-surface samples had not yet been invented. From the Roemmich, Gould and Gilson (2012) paper one can get an estimate of the number of sub-surface casts. Similarly, one can get an estimate of the number of surface bucket measurements from the Gould and Cunningham (2021) paper. It would be useful to have a table that gave the list of physical measurements for each expedition and the number of these measurements. Similarly, it would be good to list the types of biological measurements that were made, but this might get complicated if one includes the terrestrial samples. Were there many deep sea benthic biological samples that addressed the suggestions from Thompson and Carpenter? That is, it would be beneficial to have a summary list of the sampling from each cruise.

If these two issues are addressed, I strongly recommend that this manuscript be published.