This manuscript describes the NASA Ames COMA instrument, a high-altitude airborne sensor for measurement of carbon monoxide and nitrous oxide gas concentrations. The manuscript describes the customization and refitting of the core commercial sensor and laboratory and chamber testing. Data from the ACCLIP science campaign is discussed, along with carbon monoxide intercomparisons with two other sensors during that campaign. The paper is well within the scope of AMT, and presents new, novel measurement technology. The paper is clearly written with some small exceptions. I recommend publication after addressing some minor changes detailed below:

Line 93: Last sentence seems unnecessary since this describes the next section.

Figures 2 & 4: The choice of colors may be challenging for color-blind individuals. I would recommend altering the colors or adding dashed/dotted lines. At a minimum, reorganizing the legend in the same order as the color traces vertically would help.

Line 120: I would recommend citing the calibration source papers directly rather than the website.

Line 153: Which segment in figure 4 was used to perform the Allan variance calculation? Was it the entire timeseries? Seems like this would be somewhat of a worst case scenario, since most UTLS missions would have a single ascent to altitude with some profiling up high (similar to the latter half of the chamber timeseries in Fig 4). A little more information would be useful for context.

Sect. 2.2.2 & Fig 6: this section is a bit light and imprecise. Linearity is always with caveats with respect to uncertainty. How accurate are the flow controllers? Are they new with factory traceable calibrations or were they recalibrated for the experiment? I suppose the uncertainty in the standard would cancel out when just proving linearity, but the mixing errors are definitely important. Usually one can say something like “instrument is linear to within X% between MM-NN ppm”.

Line 174: “Slight degradation…was accounted for.” How was it accounted?

Line 175: maybe change “small terms due to accuracy of the standard gases” to “small contributions due to the accuracy of the standard gases”, it took me awhile to figure out what a small term was referencing

Line 178: I think there is a word missing here…maybe “equally between the residuals”?

Line 189: Is there any theory as to why the precision varies

Eq. 1&2: this might be more readable if the slopes were expressed as percents? That’s what I’m typically looking for here…just a suggestion.

Figure 7: what happened after the Aug 8th so that there are no longer 2 point NOAA gas calibrations?

Line 214: Cite data DOI?

Sect 2.1/Figure 9: Why only data from one flight? I think it is important to include all data from the campaign unless there are flights where this is not possible (e.g. missing data), along with a discussion of where they disagree and what that might mean. I also usually like either ratio or difference plots for intercomparisons rather than full scale concentration timeseries, as it highlights differences more.

Conclusion: the summary is a little slim, I would add more summation about the laboratory experiments and calibration