Reviewer 2

Ernle et al. present an analysis that evaluates the role of ozone, and the impacts a understudied (and affordable) ozone scrubbing material, on VOC artifacts observed by PTR-MS and GC sampling. The authors first evaluate the role of ozone on producing artifacts in the absence and presence of VOCs, then evaluate the efficacy of implementing a $Na_2S_2O_3$ scrubber to remove ozone and limit inferred artifacts. The authors then characterize the scrubber determine important characteristics such as breakthrough, lifetime, and effects of humidity

Overall, the paper is well organized, easy to read, and the figures are easy to interpret. I agree with a number of the conclusions drawn by the authors regarding artifacts of aldehydes, but I have a number of concerns about the conclusions drawn from the observations of the alkenes. Specifically, I question whether these observations are an “artifact” – i.e., something artificially produced by the instrumentation and sampling setup - or a reflection of an ozonolysis experiment that is expected to occur when alkenes are mixed with ozone. My main comment (below) elaborates on this further, and I hope that the authors can dig deeper into the data to address this primary concern.

We thank the reviewer for the positive comments and for bringing up the discussion about the artifacts.

Main Comment:

My main comments pertain to the conclusions drawn from the various tests. What I haven’t fully appreciated from the discussion is whether the tests really demonstrate an artifact, or simply shows that effect of VOC ozonolysis. To help frame my questions, I’ve listed the main takeaways I drew from the discussion and conclusions.

- In the presence of high ozone (but absence of VOCs), aldehyde and isoprene artifacts are observed due to ozone surface reactions with organics bound to Teflon tubing. Aldehydes are observed to increase in both the GC and PTR, while isoprene artifacts are observed by the PTR due to fragmentation of aldehydes.

  Yes, we agree.
- When steady mixing ratios of VOC standards are sampled by both instruments, the mixing of high ozone results in a positive aldehyde artifact. Part of this artifact results from surface reactions described in (1), while the remaining artifact may be due to reactions of VOCs with ozone in the tubing, instruments, or both.

  Yes, we agree.

- When steady mixing ratios of terpenes and sesquiterpenes are sampled by the instruments, the introduction of high ozone results in monoterpene and sesquiterpene decay for both instruments.

  Only for PTR, fast GC does not measure any terpenes with the method used here.

- The presence of the scrubber material removes ozone and limits decay of monoterpenes, but not sesquiterpenes.

  It prevents the decay of mono- & limits the decay of sesquiterpenes.

I'm convinced that conclusion (1) is consistent with a sampling artifact, and I fully agree with the authors that this presents an important consideration when sampling in high ozone environments.

Conclusions (2) and (3) are drawn from an experimental setup that essentially simulates VOC ozonolysis, and point (4) simply shows that removing ozone prevents VOC oxidation. So, is this really an artifact, or just an ozonolysis experiment? What I would like to know is if there is additional chemistry in the tubing (or in the instruments) that is prevented by the presence of the scrubber? Or in other words, if I don't have a scrubber, and I were to measure terpenes in the atmosphere, would I measure a bias because of surface reactions? Right now, as described, it is not clear to me that that this is true and if a scrubber is needed for preventing these potential artifacts

Thanks for these questions. The reviewer is correct, that in the presence of the scrubber gas-phase ozonolysis reactions are prevented (= ozonolysis experiment). The scrubber leads to a reduction of the gas-phase chemistry and therefore less interference. Additionally, the scrubber prevents ozone-surface reactions inside the tubing (and the GC), which cause for example the formation of positive acetaldehyde artifacts. If you don't have a scrubber and were to measure terpenes in the atmosphere, there would still be ozone induced depletion of the terpenes (gas phase), which could be prevented by placing the scrubber at the very front end of the inlet line. Unfortunately, we cannot be sure whether or not surface reactions are additionally occurring with the terpenes. An important point for us was, that the $Na_2S_2O_3$ scrubber does not interfere the terpene measurement, which means we can keep our list of target analytes as it is when using the scrubber.

I do think that the authors have the data to demonstrate whether these effects are present and perhaps can expand this discussion. For example, Fig. 3 seems to show that the GC has a higher positive bias than the PTR for acetaldehyde as the VOC mixture is increased at high ozone (i.e., the change in acetaldehyde from 50 – 1000 ppb is much higher for the GC than for the PTR). Is it possible that this is an artifact of the GC preconcentration? In such a case, I would agree that this an instrument artifact. Are similar observations made for the monoterpenes and sesquiterpenes?

I.e., are there relatively larger negative biases for the GC than for the PTR? Again, this would be convincing of a negative bias owing to instrument sampling and would warrant the use of a ozone scrubber to limit the sampling artifact.

Thank you for the question about the different acetaldehyde increase. We added some sentences for clarification on this point. We did not make similar observations for the terpenes with the GC, these were only measured by PTR.

l. 203: “Both, the PTR- and GC-MS measured higher acetaldehyde mixing ratios when O3 was above 150 ppb (see Figure 3). This indicates that the interference is not instrument specific but more likely a function of the common inlet tubing exposure to ozone. Note that the inlet lengths to GC and PTR were roughly the same. The higher enhancement of the GC acetaldehyde could be due to emission of oxidation products from the material of multiposition valves as described by Apel et al.. In contrast to the PTR data, the ozone induced enhancement of the GC signal increases with acetaldehyde concentration. This effect can be due to the different materials used for the tubing inside the instruments: Deming et al. showed, that in glass and metal tubing competitive adsorption occurs, which depends on humidity, concentration and functionality of the analyte, while polymer tubing shows independent absorption. Our fast GC instrument is equipped with heated silico-steel tubing, which allows competitive adsorption, while the PTR is equipped with PFA tubing. Additionally, with increasing O_3 mixing ratios ozonolysis reactions during trapping are gaining importance. It seems, that the interferences on the VOC measurements caused by high ozone exposure are an effect of both, inlet line and instrument’s surfaces.”

Other comments:

Line 64: Are there studies which show the effects of new vs. old tubing on VOC measurements?

To our knowledge there are no such studies, but Northway et al. mention that the history of the tubing is important. According to that study, the level of artifact production depends a) on the tubing (material, geometry) and b) on the tubing history. Conditioning with ozone improved the artifact formation, but was unfortunately a reversible effect.

Line 90: A lot of experiments were performed under a range of different VOC conditions. I might suggest including a table that lists out experiment conditions for clarity to the reader.

Thanks for this suggestion, we included a table with the different conditions at l. 73.

Condition	**O_3 levels / ppb**	**Calgas MR / ppb**	**RH / %**
Effect of O_3 on VOCs	0, 50, 1000	0, 0.5, 1, 2, 4	0
Effect of O_3 on tubing	0, 25, 50, 100, 150, 400, 750, 1000	0, 0.5	0
Effect of RH on VOCs/scrubber	0, 50, 150	0, 0.5, 2	0, 50, 80
Scrubber endurance	50, 150, 1000	0	0, 80

Line 114: While fragmentation is less of an issue for many of the analytes, a number of these species (and their products) do fragment (e.g. siloxanes, monoterpenes, sesquiterpenes, etc), and can impact important measurements of species such as isoprene, as demonstrated by the authors. I suggest rephrasing and referencing relevant fragmentation papers (e.g. Pagonis et al).

Thank you for this suggestion, we added some sentences to cover this point:

l. 125: “This is a soft ionization technique and therefore causes little fragmentation of the analytes during the detection process. This is the case for most analytes in this study. However, some species (e.g. terpenes, siloxanes) do fragment during ionization (Pagonis, 2019). Fragments can impact the measurement of target species such as isoprene if they have exactly the same mass.”

Line 150 and Table 1: Could the authors propose a quantitative measure for each interference? For example, the increase in signal during zero VOC injection (e.g., amount of VOC signal produced per ppb of ozone introduced) would be helpful in quantifying positive artifacts owing to ozone interactions with the walls of the tubing and/or instrumentation.

We agree, that it would be helpful if there was a quantitative measure for each interference expressed in terms of ozone and tubing length. Unfortunately, according to Northway et al., the interferences from the tubing depend on both, tubing material and history of the tubing (previous exposure of the tubing to VOCs and ozone). This means that the specific degree of interference when unscrubbed high ozone levels are introduced, is dependent on the history of the inlet. Therefore, any relationship we derive is not generally applicable and might be misleading for the reader. Nevertheless, as an example we calculated the yield (ppb of VOC produced per ppb of ozone introduced) for acetaldehyde, which is the most affected compound measured by both instruments in our study. The average acetaldehyde yield for the GC $y_{acetaldehydeGC}$ was 0.0043±0.0055 per ppb of ozone while the average yield for the PTR $y_{acetaldehydePTR}$ was 0.0003±0.0011. The GC has higher production of VOC interference signal compared to the PTR. This is in line with our observations from Figure 3, that the GC signal increases with VOC and ozone concentration.

Table 1: As mentioned in my main comment, I'm still not quite sure if what is presented for monoterpenes and sesquiterpenes is an interference per say, and so I would be hesitant to include a down arrow for these species without further digging into the data and demonstrating that the high ozone is leading to additional biases beyond those of an ozonolysis experiment.

We agree, that the name 'artifact' might be misleading for compound depletion due to gas-phase ozonolysis reaction. Therefore, we replaced 'artifact' with 'interference' where necessary. Nevertheless, the aim of our study was to investigate the influence of O3 and humidity on VOC measurements. This includes also ozonolysis reactions, that happen during sampling. As Table 1 presents the effect of ozone on the measured species, we include ozonolysis reactions and sampling artifacts.

Figure 2: This caption feels incomplete and really doesn't describe what the authors are trying to show. It may be better to say "Effect of ozone mixing ratios (0, 50, 1000 ppb) on GC and PTR-MS measurements of chlorobenzene under a range of VOC mixing ratios (X - X ppb)"

Thanks for spotting this. We rephrased as following:

"Figure 2: Effect of Ozone (0, 50, 1000 ppb) on GC and PTR-MS measurements of chlorobenzene under different standard gas levels (0-4 ppb)."

Lines 158 – 161: These sentences feel a bit distracting and I don't think are necessary for the discussion. This statement could be removed, or simply stated – e.g. "While the instruments observe slightly different mixing ratios owing to differences in sensitivity, the relative change in chlorobenzene mixing ratios remain unchanged when ozone is present."

Thank you for this suggestion, we changed the manuscript accordingly.

Line 165 - 166: Can the authors expand here? What trends or aspects of the measurements agree with what was observed by Northway et al. and Lehumpuhl et al.?

Thank you for the suggestion, we expanded the section accordingly.

l. 196: "Based on current literature, some carbonyl compounds such as aldehydes and ketones do have the potential to show higher mixing ratios when ozone is present. Northway et al. and Apel et. al observed positive artifacts under ozone presence for acetaldehyde. Additionally, Apel et al. observed artifact formation for propanal, acetone and butanal in their fast GC-MS system, which were emitted by parts of their system when ozone was present. Lehmpuhl et al. found positive interference also for larger carbonyl compounds."

Line 169 – 173: Here, the authors discuss the observations of acetaldehyde for each instrument independently. I'm also interested in the relative changes between the instruments. For example, it seems that the GC observes a significantly higher change in acetaldehyde at high O3 than the PTR. Is this due to a longer inlet line, differences in the residence time, or possibly differences in the instruments that that are causing these effects (e.g., ozone in the GC trap)? At first glance, this looks like this could be a combination of line impacts and instrument artifact for GC, while for the PTR this looks more like an effect of aldehydes production from the surface reactions of the tubing since the changes in the presence of ozone at the various VOC mixing ratios seem to match the changes in signal when VOCs are absent (as demonstrated in figure 4b).

Thank you for this comment, we agree with this interpretation. The inlet tubing lengths are kept the same for both instruments. Ozonolysis during trapping is likely. During field operation we use an internal ozone filter scrubber for the GC-MS, but for this experiment we removed this intentionally to investigate the influence of the scrubber on the VOC measurement. With increasing O_3 mixing ratios ozonolysis reactions during trapping become more important. Now, we discuss the differences in acetaldehyde enhancement more in detail, as already mentioned in the “Main comments” section.

l. 203: “Both, the PTR- and GC-MS measured higher acetaldehyde mixing ratios when O3 was above 150 ppb (see Figure 3). This indicates that the interference is not instrument specific but more likely a function of the common inlet tubing exposure to ozone. Note that the inlet lengths to GC and PTR were roughly the same. In contrast to the PTR data, the ozone induced enhancement of the GC signal increases with acetaldehyde concentration. This effect can be due to the different materials used for the tubing inside the instruments: Deming et al. show, that in glass and metal tubing competitive adsorption occurs, which depends on humidity, concentration and functionality of the analyte, while polymer tubing shows the phenomenon of independent absorption. Our fast GC instrument is equipped with silico-steel tubing, leading to competitive adsorption, while the PTR is equipped with PFA tubing. The higher enhancement of the GC acetaldehyde could also be due to emission of oxidation products from the material of multiposition valves as described by Apel et al. Additionally, with increasing O_3 mixing ratios ozonolysis reactions during trapping are gaining importance.”

Line 178: Do the GC signals sum up to what is observed by the PTR?

Thanks for the question, we adjusted the text accordingly:

l. 226: “The PTR-ToF-MS in H_3O^+ mode cannot separate the aldehyde from the ketone as they have exactly the same mass, i.e. the PTR-Tof-data presented here always shows the sum of propanal and acetone (C3) and butanal and MEK (C4) respectively and should be therefore double the GC signals for the separated species."

Since this was not the case for all experiments shown, we double checked the raw data work-up for this experiment and found that there was indeed a problem for the PTR integration on this day (due to a software issue in the mass calibration). Therefore, we integrated again and requantified, taking into account the dependence on the primary ions, drift tube pressure and temperature as before. For MGC we also found a problem in the background correction that resulted in slightly negative values for zero air. However, even with these corrections we cannot explain the low mixing ratios for the C4 carbonyls measured by MGC. We double checked the data, but did not find any indication for a drift in the standard gas concentration or the existence of a leak. However, those factors would influence all compounds present in the standard gas, not only butanal and MEK. If the reason was an instrument internal problem, it should be observed for all experiments, not only on the day of the experiment presented in Figure 6b). As the low mixing ratios are measured under all O_3 conditions, a depletion through ozonolysis can also be excluded. The calibration on that day was linear. Usually, we performed a full calibration before and after the experiment. On that day however, the calibration in the morning failed, so we used only the calibration from after the experiment for the quantification. Generally, we expect a higher sensitivity during the first calibration after tuning (tuning was performed every morning) than after several hours of measurement. If this was the case and we use only the calibration data with less sensitivity i.e. lower area per ppb, we should a) measure too high concentrations during the experiment and b) see a drift towards higher concentrations with the measurement time. Neither is the case. In principal, new tuning between calibration and experiment could result in generally low mixing ratios, but we did only tune once in the morning (before the start of the experiment). Finally, we compared the response factor of the calibration from this day with the average response factor for all experiments. Unlike the C3 carbonyls, where the quantification is correct and the response factor matches the average response factor for all experiments, the response factor for C4 carbonyls indicates less sensitivity during this day's calibration than on average. Again, this does not explain the low mixing ratios as low sensitivity during calibration (and corresponding higher sensitivity during the measurement) would result in too high measured mixing ratios during the experiment. It is not satisfying that we could not find the reason for the C4 carbonyl quantification problem on that experimental day. However, we include the MGC data as the imperfect quantification does not interfere with our qualitative analysis of interferences and they fit to the qualitative results of the other experiments including the GC data. Additionally, the qualitative result for C4carbonyls is in line with the qualitative result of the measured C2-C3 carbonyls: The signals for the ketones (acetone (C3), MEK (C4)) increase with O_3 mixing ratios ≥400 ppb O_3 and the signals for the aldehydes (acetaldehyde (C2), propanal (C3) and butanal (C4)) are relatively stable with a tendency to decrease between 200 and 400 ppb O_3 and increase as well with O_3 mixing ratios ≥400 ppb.

We replaced Figures 4a) and 6 with the corrected data of both instruments:

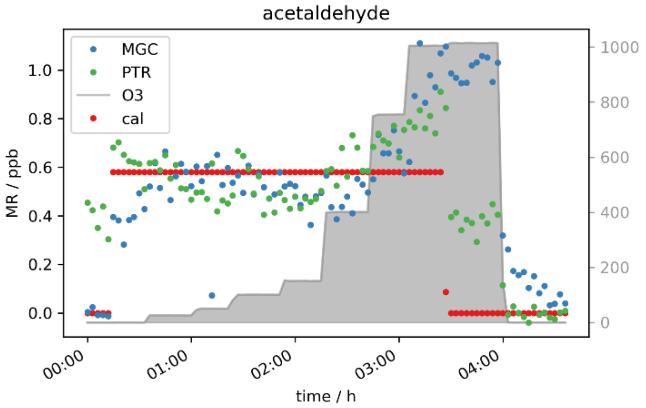

Figure 4a): Acetaldehyde mixing ratios at seven different ozone levels between 0 and 1000 ppb and 0 ppb standard gas level.

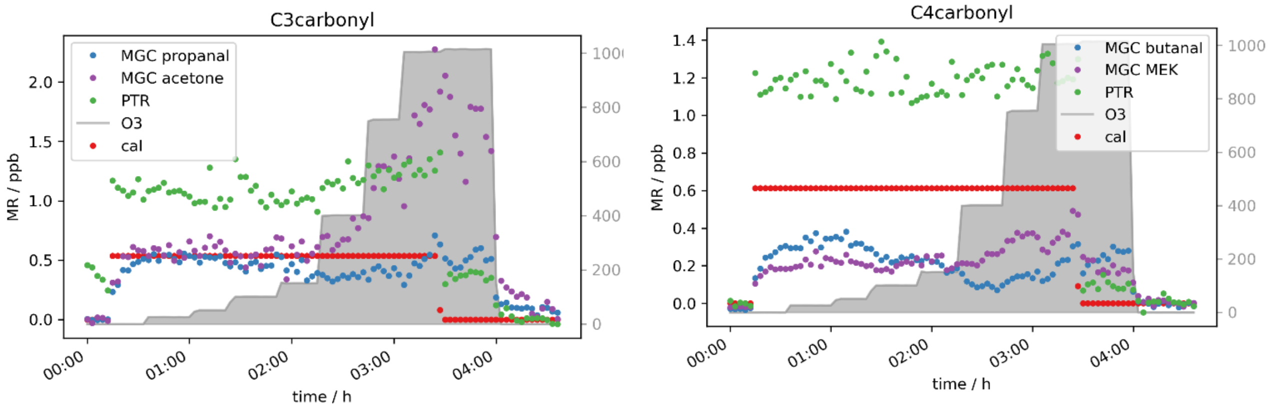

Figure 6: C3- and C4 carbonyl mixing ratios at approximately 0.5 ppb per VOC at different ozone levels.

Additionally, we adjusted the text accordingly:

l. 231: „Interestingly, the aldehyde mixing ratios are relatively stable with a tendency to decrease with ozone when the standard gas was added. Figure 6 shows this phenomenon. Propanal and butanal mixing ratios do not show a substantial increase under the same O_3 conditions where they increase in the zero air measurement, while the sum of C3~carbonyls (PTR signal) and GC acetone again increase (as in the zero air measurement). As propanal slightly decreases and acetone strongly increases with ozone, the PTR measurements show a positive net ozone effect for the C3 carbonyls. For C4 carbonyls, the GC quantification during this experiment was compromised (too low mixing ratio) for unknown reasons. However, the qualitative results match the rest of our observations: butanal decreases slightly, while MEK increases slightly, leading to a stable signal for the sum of butanal and MEK, which is shown by the PTR data presented in Figure 6b. Additionally, the qualitative results of butanal and MEK are in line with the qualitative results of C2-C3 carbonyls: The signals for the ketones (acetone (C3), MEK (C4)) increase with O_3 mixing ratios ≥400 ppb and the signals for the aldehydes (acetaldehyde (C2), propanal (C3) and butanal (C4)) are relatively stable with a tendency to decrease between 200 and 400 ppb O_3 and increase as well with O_3 mixing ratios ≥400 ppb. Between 3.5 and 4 h after start of the experiment (cf. Figure 6), not all signals dropped to background levels. They finally drop once ozone was switched off. This is consistent with the results from the zero air measurement (Figure 5) and the acetaldehyde data (Figure 4b)). It shows that exposure of the inlet tubing to high ozone does not rapidly clean the lines of the interfering compounds."

Figure 10: Is this measurement by PTR or GC? Please specify in the caption.

We agree this would be helpful and adjusted the caption accordingly:

"Figure 10. Terpene mixing ratios measured by PTR-MS with and without scrubber at 50 and 170~ppb O_3."

Line 330: I presume humidity didn't have a strong effect on the oxygenates or other species that had a positive artifact during ozonolysis?

Your assumption is correct. We mention in line 389-391, that we did not observe any strong effects during this study due to drying/using a humid calibration:

l. 389: "For the measurements performed within this study, humidity did not have any influence on the GC- and PTR-ToF-MS instrument measurement capability as these dried the air before detection (GC-MS) or used humid calibrations (PTR-ToF-MS)."

Line 360: Citation for the stratospheric observations? Also, It would be helpful to compare these interferences to those observed in the stratospheric work to put into perspective the real-world implications of these interferences.

To address the reference request, we give the Apel et al. reference as the impetus to test their system for ozone was the stratospheric interference. We too have seen acetaldehyde anomalies in earlier PTR-MS data from the stratosphere, but as it was clearly erroneous it was not published.

As mentioned in the answer concerning your comment about line 150 and Table1, in our experiments the average acetaldehyde yield per ppb ozone for the GC $y_{\text{acetaldehydeGC}}$ was 0.0043±0.0055 while the average yield for the PTR $y_{\text{acetaldehydePTR}}$ was 0.0003±0.0011. The GC has higher production of VOC interference signal compared to the PTR. This is in line with our observations from Figure 3, that the GC signal increases with VOC and ozone concentration. The most obvious indicator of an ozone interference in stratospheric air was with acetaldehyde. For example, in one such incursion we saw similar dependencies: Our PTR measured 0.88 ppb acetaldehyde at an altitude of ~13000 m, when ozone was 465 ppb. Unfortunately, the fast GC-MS did not measure acetaldehyde during this campaign so a direct comparison cannot be made from this dataset.

We cover these points with the following added text.

l. 416: "Signals for acetaldehyde, propanal, acetone and butanal increased with ozone levels above 150 ppb in background measurements with GC as well as PTR.

Thus, it can be concluded, that there are positive artifacts generated in the experimental setup i.e. in the tubing, inside the ozone generator or within both of the mass spectrometers. As we observed the same during stratospheric measurements before, the ozone generator cannot be the only source. Apel et al. originally conducted ozone sensitivity tests on their airborne GC-MS system due to anomalous acetaldehyde observations in the stratosphere. Our experiments also show acetaldehyde to be the species most affected by ozone interference. When our PTR encountered a stratospheric intrusion in flight as on 2nd June 2020 we found 0.88 ppb acetaldehyde (altitude 13000 m, O_3 465 ppb), extremely suspect for such a short-lived molecule under otherwise clean conditions. Unfortunately, the fast GC-MS did not measure acetaldehyde during this flight campaign. When 0.5 ppb of VOC standard gas was measured, the signals for propanal and butanal decreased due to the reaction with the OH radical, generated from ozonolysis of terpenes."

Technical Comments:

Line 27: Suggest "measured" in place of "covered"

We replaced the word:

l. 26: "With these measurement techniques a wide range of volatile organic compounds can be measured including aliphatic and aromatic hydrocarbons,[…]"

Line 31: Suggest "mixing ratios" in place of "values"

We replaced the word:

l. 33: "Northway et al. (2004) and Apel et al. (2003) reported for example increased mixing ratios for acetaldehyde in their systems for measurements in the lower stratosphere where ozone levels are high and humidity is low."

Line 62: It would be more helpful to point to Table 1 as opposed to Section 3 for the list of VOCs.

Thanks for the suggestion, we now refer to the table.

l. 61: "A list with the measured species is shown in Table 2."

Line 64: Suggest "instrument" or "technique" in place of "machine"

Thanks for the suggestion. We removed the word "machine".

l. 64: "Some species could be detected with both instruments simultaneously while other species could only be measured by one."

Line 95: Suggest “instruments” rather than “mass spectrometers” since the GC is not solely a mass spectrometer

l. 104: “For those experiments the flow through the scrubber was higher to provide enough air for both of the instruments.”

Line 186: “Netto” should be “net”

Thanks for finding this typo. We corrected the word accordingly.

l. 235: “As propanal decreases and acetone increases with ozone, the PTR measurements shows the net ozone effect on C3 carbonyl.”

Line 190: The wording is confusing - perhaps "compounds with positive artifacts"

Thanks. We adjusted the text as following:

l. 244: “It shows that exposure of the inlet tubing to high ozone does not rapidly clean the lines of the interfering compounds.”

Line 261: Suggest “junction” in place of “T-piece”

Thanks for this suggestion. We replaced the word.