1. Method 26 should be included in the list of examples.

   The list of example reference methods in section 2.3 includes Method 26A, Method 320, Method 321, and ASTM D6348-12, We suggest that is also include Method 26. At least one of our members uses an HCl CEMs. They have been successful in showing that the instrument passes a RATA using Method 26 for a number of years. We see no reason why Method 26 should not be added to the list.

2. Does it make sense to require paired Method 26A trains? (11.7.4.4, p 29)

   Performance Specification (PS) 12A is the only PS CRWI knows about that requires the use of paired reference method sampling trains (see Section 8.4.2 of PS 12A) when using Method 29. Paired trains are recommended in PS 11 (see section 8.6(1)(i)), but they are not required. The reason for paired runs was to demonstrate that the results of the paired runs were consistent by having to meet a relative difference requirement. Both Methods 26 and 26A have been widely used for a number of years to develop data to set standard and to show compliance. We are not aware of any data that would indicate that the proper use of either method would result in inaccurate data. It should be noted that Method 26A has a known negative bias below 20 ppmv (Section 123.1 of Method 26A). However, this bias would show up in both trains (if a dual train was used) and would not have any impact on determining accuracy.

   However, we acknowledge that the Agency may have additional data that suggests the paired train should be required. If this is the case, we hope the Agency will include that in the docket for this rulemaking. Dual trains are more expensive and complicate the testing process. Before requiring them, we suggest that the Agency show they are warranted (i.e., the data from a single train is unreliable). One alternative is to follow the requirements in PS 11 and recommend the use of paired trains but not require them.

   While we would prefer that the Agency allow the user the option of using either single or paired trains, if they decide to require paired trains, we suggest modification of the relative difference (RD) requirements. In the current draft, the data from the paired trains must pass the relative difference criterion (≤ 10%) before it can be used. We are concerned that it would be difficult to get data to meet this criterion, especially at very low concentrations.

   In Method 5i, EPA modified the criterion based on how close you are to the method detection level. The requirements in Method 5i are as follows.

   > 12.2 b. A minimum precision criteria for Reference Method PM data is that RSD for any data pair must be less than 10% as long as the mean PM concentration is greater than 10 mg/unit volume. If the mean PM concentration is less than 10

mg/unit volume higher RSD values are acceptable. At mean PM concentration of 1 mg/unit volume acceptable RSD for paired trains is 25%. Between 1 and 10 mg/unit volume acceptable RSD criteria should be linearly scaled from 25% to 10%. Pairs of manual method data exceeding these RSD criteria should be eliminated from the data set used to develop a PM CEMS correlation or to assess RCA.

A similar sliding scale can be found in PS 12A.

8.4.6.2. The minimum performance criteria for RM Hg data is that RD for any data pair must be at least ≤ 10 percent as long as the mean Hg concentration is greater than 1.0 $\mu g/m^3$. If the mean Hg concentration is less than or equal to 1.0 $\mu g/m^3$, the RD must be ≤ 20 percent or ≤ 0.2 $\mu g/m^3$ absolute difference. Pairs of RM data exceeding these RD criteria should be eliminated from the data set used to develop a Hg CEMS correlation or to assess CEMS RA.

The use of a sliding scale is based on the idea that as you get to lower numbers, the accuracy becomes less and the relative difference gets higher – all for a number that is relatively small. We suggest using a sliding scale for the RD criterion much like what is used in Methods 5i and 12A.

3. RA requirements (11.7.4, p. 28)

As proposed, PS 18 allows for either reference Method 26A or dynamic spiking to conduct a relative accuracy test. We support allowing options. The sentence in section 2.4 (page 3) could be read to require dynamic spiking at sources with emissions near the detection limit of the CEMS. We believe the intent of this paragraph was to say that to get reliable results, a facility may need to use dynamic spiking since Method 26A may not be sufficiently reliable at certain concentration levels. We suggest that the sentence be re-worded to make its intent clear.

Should a facility use Method 26A, they are required to do at least 9 runs with a maximum of 12 (allowing you to discard up to 3). This appears to be the norm for PS where you see the results at the time of the testing. However, when using Method 26A (or 26), the facility sends the samples off-site for analysis and may not get the results back for several weeks. PS 12A allows a facility to run as many tests as you want but you are required to report all data. Since it expensive to restart a test once you have discovered that you need more data, we suggest that PS 18 be modified to reflect the data requirement of PS 12A – that is, a minimum of 9 runs with no maximum but reporting all data. This way a facility can decide how many runs they want to do while the stack testing team is on site to insure they get to the minimum required.

Appendix A – Dynamic spiking procedure

4. CRWI is concerned that the dynamic spiking procedure in Appendix A is too specific in some cases. Eli Lilly and Company developed a dynamic spiking protocol for their Ecochem HCl CEMs. This is posted on EPA's web site as OTM 25. This dynamic spiking protocol spiked at three different levels, collected 30 minutes of data at each level (about 30 data points), and used linear regression to define the precision, accuracy, and bias. It uses the O_2 concentration as an indicator of the dilution ratio (spiked gas to flue gas) since O_2 was very consistent. Thus, any drop in O_2 concentration would be due to the dilution with the calibration gas. OTM 25 recommends defining a maximum dilution ratio or range, number of data points, and range for doing the spiking). Lilly has successfully used this protocol to perform the RATAs on their instrument. They have found that three points over the calibrated range was a very good indicator of the linearity of the instrument. While this protocol may be written for a specific instrument, we see no reason why it could not and should not be used as a starting place to develop a more generic dynamic spiking protocol. CRWI suggests that the Agency re-write this protocol using OTM 25 as a starting point.

5. P 10, 11.2.1.6 and 7 and 11.2.2.4 (p. 11). Must use 6 or more dynamic spiking measurements. How does this relate to the 9 required in other places? A12 (p 12) requires results from 9 runs. Is the "6" a typo?