

## Report on New Evaluation Procedure Survey

### INTRODUCTION

In October, 2005, the evaluation procedure and report format were modified for the Proficiency Testing program. These changes were made in response to feedback from CAEAL members solicited during workshops and other discussions.

The major changes to the evaluation procedure were,

- Eliminate the assigned point system and calculate the PT score using the average z score.
- Allow for the inclusion of a laboratory's detection level in PT evaluations.
- Include flags for bias in the PT reports.
- Add a statement about the acceptability of a PT score to the reports.

### SURVEY SUMMARY

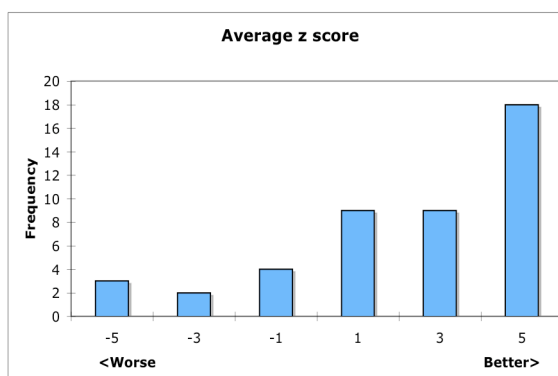
A total of forty-five laboratories submitted a completed survey.

Opinions on the new evaluation procedure and report format range from them being significantly worse than before to being a significant improvement. The following are some of the general comments received with the survey,

- *Pleased with changes.*
- *Good.*
- *Initially very confusing but after closer examination and a bit of clarification the new system is OK.*
- *No real improvement noted.*
- *This is more in accordance with other PT facilities scores. An overall improvement.*
- *Actual PT evaluation is clearer than past.*

#### *Average z Score*

Seventy percent of respondents reported that the average z score system is an improvement over the previous scheme and twenty percent reported that it is worse.



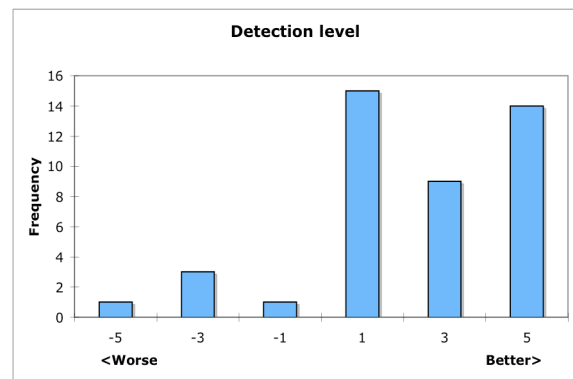
Comments include,

- *This evaluation procedure is satisfactory as it increases the accuracy of measurements.*
- *The average z score system is terrible. We failed the arsenic test because one sample has very high standard deviation and other 3 samples are acceptable. The average z score is 68%. This calculation does not represent our average performance.*
- *It is more real that you report average z score.*

### ***Detection Level Evaluation***

Fifty-six percent of respondents reported that the detection level evaluation was an improvement and twelve percent reported that the evaluation is worse. The large proportion of laboratories that reported no or very little improvement may be due to the fact that many laboratories did not avail themselves of this option.

There were no comments submitted specifically about the detection level changes.



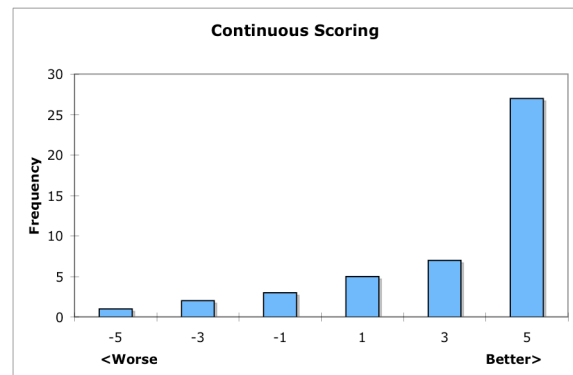
### ***Continuous Scoring Scheme***

The use of assigned points was flagged by several members as a weakness in the program because of its stepwise nature. For this reason, the average z score was introduced. This provides a continuous scoring system that better reflects a laboratory's real performance, relative to other laboratories.

Seventy-eight percent of laboratories reported that this is an improvement and thirteen percent reported that it is worse.

Comments include,

- *Overall, the new evaluation is an improvement over the step process.*

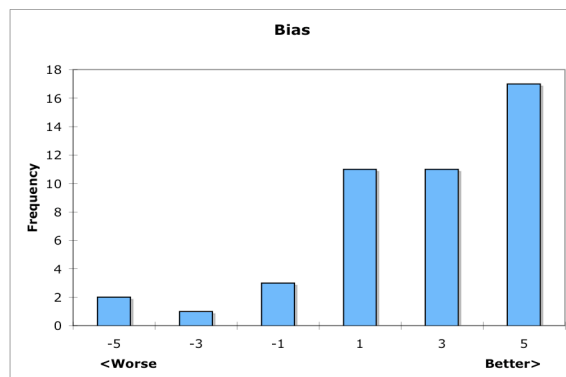


- *Need to introduce an error bar for your 70 percentile since you do not go in steps of 5 anymore.*

### ***Bias***

Bias flags were introduced to the confidential report to provide participants an additional tool in evaluation of performance. This information was always available on the generic reports posted on the CAEAL web site but it was felt that they would be of greater value on the laboratory reports.

Seventy-three percent of respondents reported that including bias flags on the reports was an improvement. Thirteen percent reported that it made the reports worse.



One of the laboratory comments detailed a problem with this procedure that was not identified during development of the new scheme. CAEAL's current scheme will assign a z score of 2 or 3 to a non-detect result. This automatically results in a high bias flag even though there may be no bias present. This will be corrected for the next study.

Comments include,

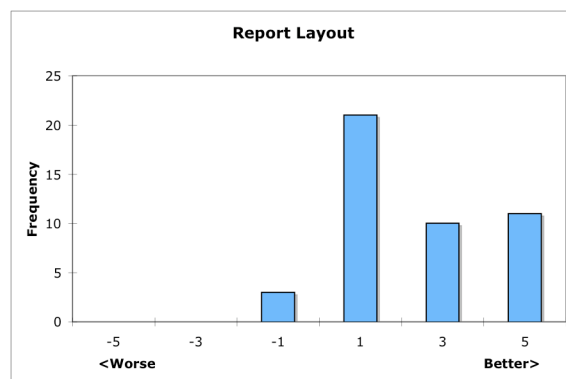
- *No relevance in bias evaluation.*
- *Evaluation of bias is not appropriate when a non-detect is reported because a z score of 2 or 3 is automatically assigned. This results in a high bias when none may exist.*

### ***Report Layout***

Only forty-eight percent of respondents reported that the report layout was an improvement over the previous layout. However, only seven percent reported that it was worse.

Comments include,

- *Sorting parameters by appendix is more relevant for lab.*
- *Would like to see a report generated that lists method, bias and PT score*





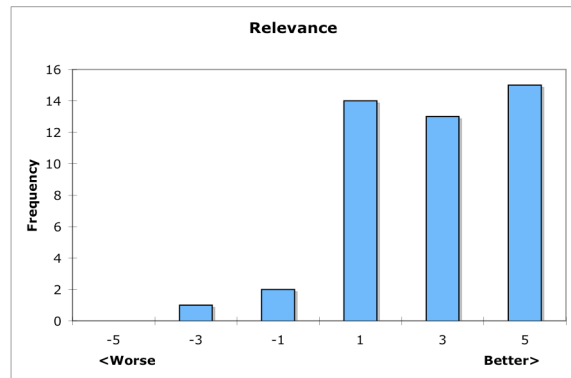
- *I like the summary column.*
- *The report format is clear and simple.*
- *In general, reports are easy to read.*

### **Relevance**

Sixty-four percent of respondents reported that the PT reports were more relevant than the previous ones. Seven percent reported that they were less relevant.

Comments include,

- *Needs to be separate evaluations for ICPOES vs MS vs hydride and should not be penalized with a 3 as a z score for < dl reporting.*
- *It is good to report s and z score rather than 2s and z/2.*



### **CONCLUSIONS**

Based on responses from the survey, it can be concluded that most laboratories perceive that the new evaluation procedure and report relevance are an improvement over those used in the past. However, it is obvious that laboratories do not believe that the layout of the report is better or worse than the previous reports.

A survey response identified a weakness in the current estimates of bias. Laboratories that report one or more non-detects are likely to receive a high or very high bias even though a real bias may not exist. This will be examined and addressed before the next study.