

**Guidelines for Proper Connecticut Mastery Test Data Analysis
Handout for RESC Workshops
November 2007**

Appropriate Analyses:

- Within a generation and grade, Mathematics, Reading and Writing results may be compared across years. (e.g., fourth-grade mathematics performance levels in 2006 can be compared to fourth-grade mathematics performance levels in 2007.) *Note: As a reminder, the Generation Three CMT was administered from 2000-2004 and the Generation Four CMT has been administered in 2006 and 2007.*
- Within a generation and grade, Mathematics, Reading and Writing results may be compared across years at the subgroup level (i.e., English language learner status, special education status, gender status, free/reduced meal status or ethnic background status). For example, the performance of 6th-grade girls in writing can be compared to the performance of 6th-grade boys in writing.
- Within a generation and grade, comparisons may be made on the basis of scale scores and achievement levels for all groups of students. In Reading, DRP unit scores may also be compared for all groups of students.
- Within a generation, grade and year, strand level performance may be compared to other strands within the same content area. For example, looking at the percent of students achieving mastery in one strand in grade five and comparing this performance to the percent of mastery in other strands will show the areas where students are excelling or need improvement. Within a generation and across years, comparisons can also be made to relative changes at the strand level. For example, looking at the change in performance in one strand in grade four at the school level can be compared to changes at the state or district level in the same strand.

Inappropriate Analyses:

- Direct comparisons across generations are inappropriate. (e.g., it is not appropriate to compare the performance of fourth-grade mathematics students in the fall of 2004 to the performance of fourth-grade mathematics students in the spring of 2006.)
- Comparisons of performance levels across grades (either within the same year or across years) within a content area cannot yet be made. For example, one cannot legitimately compare seventh-grade performance in the goal range in 2007 to eighth-grade performance in the goal range in 2007. Nor can one legitimately compare seventh-grade performance in the goal range in 2006 to eighth-grade performance in the goal range in 2007.

Note: The CSDE is currently developing a vertical scale, which will provide a way to track progress across grades.

- Averaging numbers across subjects within a grade is not appropriate. (e.g. the percentage of students at the goal level in reading across grades cannot be averaged; neither can the percentage of students at the goal level be averaged across all the content areas within a grade).

- Caution should be used when looking at strand level data. Because of test design characteristics, strand level data is not comparable across grades, nor is it very reliable to look at strand level performance across years within a content level. For example, it is not appropriate to look at the percent of students achieving mastery in strand 15 in grade four and compare it to the percent of students achieving mastery in strand 15 in grade five. Comparing strand level performance across years within the same grade should be done with caution and with certain caveats in mind since the strand level difficulty across years is pre-equated but not post equated. This means that variations in the scores may not necessarily represent changes in student performance.