APPENDIX C

Comparisons Among States Based on Average Proficiency

Figure C.1 is provided as a visual representation of the distribution of proficiency results for each participating jurisdiction. The darkest box at the midpoint of each distribution shows the 95 percent confidence interval around the average proficiency. The lighter shaded boxes indicate the locations of selected percentiles of each jurisdiction distribution. The intervals take into account the sampling and measurement error associated with the estimates of average proficiency. Jurisdictions are listed by overall average reading proficiency — beginning with the state of Maine whose average reading proficiency for fourth-grade public school students is 229 with a standard error of 1.3 points.

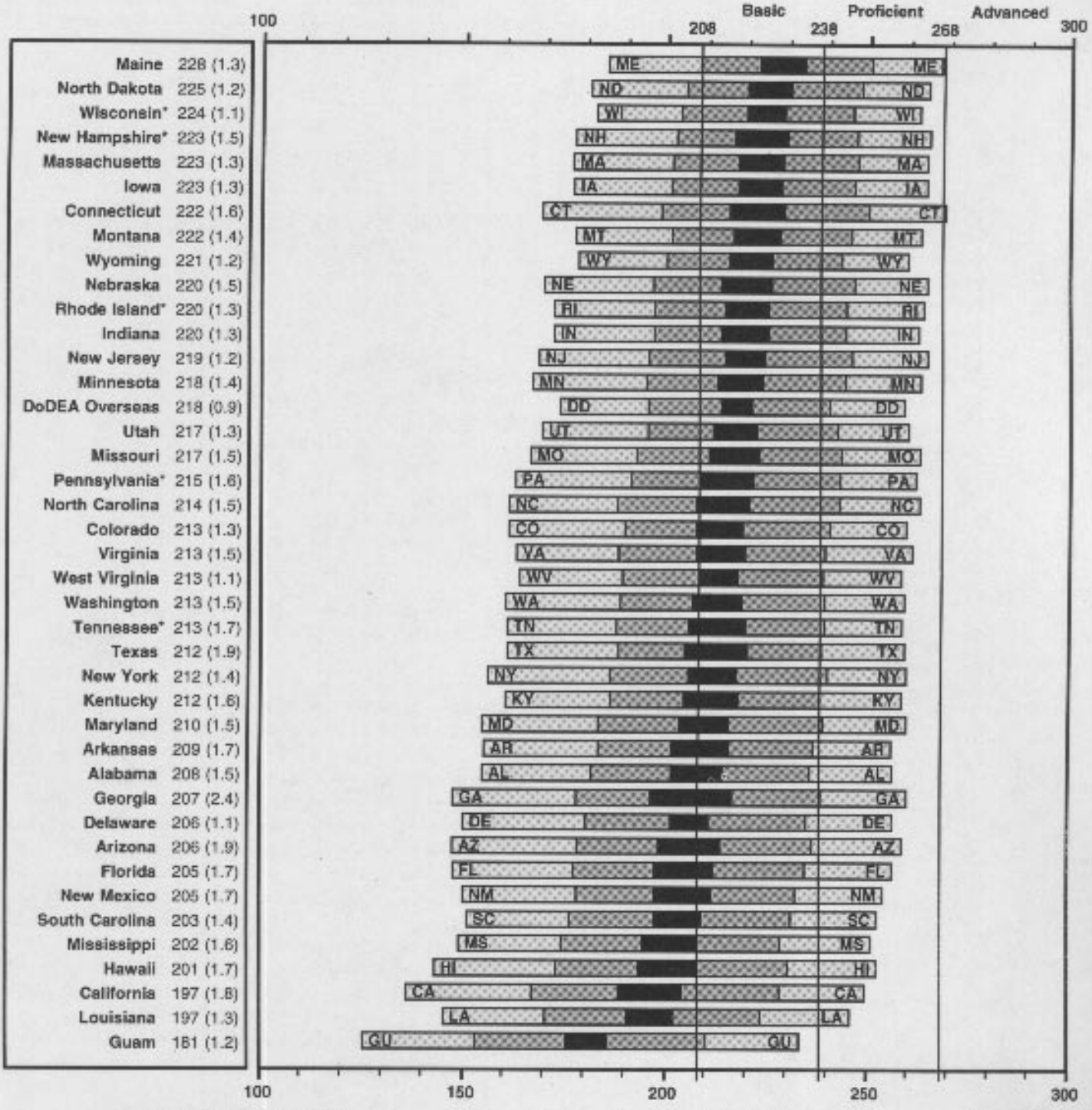
Figure C.2 is provided to help interpret differences in the average proficiencies across states for grade 4 in 1994. The figure provides a method for making appropriate comparisons in average overall reading proficiency across the participating jurisdictions. The figure shows whether or not the differences in average performance between the pairs of jurisdictions are statistically significant.[1]

For example, in Figure C.2, although the average proficiencies in the fourth grade appear to be different between Maine (229) and Montana (223), they in fact are **not** statistically different. The computations underlying Figure C.2 take the sampling and measurement error associated with the estimates of average proficiency into account, as well as controlling for the large number of comparisons that are being made.

As an example of how to read Figure C.2, let us say we are attempting to compare the state of Texas to all other jurisdictions. Reading vertically down the Figure C.2 column labeled Texas, we see that, on average, students in Texas scored lower than did students in all the states listed from Maine through Montana (the dark grey shaded states), about the same, on average, as students in the states listed from Wyoming through South Carolina (the white shaded states), and better, on average, than students in all the states from Mississippi to Guam (the light grey shaded states).

1. The significance tests in Figure C.2 are based on a Bonferroni procedure for multiple comparisons that holds to five percent across all possible comparisons the probability of erroneously declaring the means of any two states to be different when they are not.

Figure C.1 Distribution of Overall Reading Proficiency Organized by Average Proficiency for the 1994 Trial State Reading Assessment, Grade 4, Public Schools Only



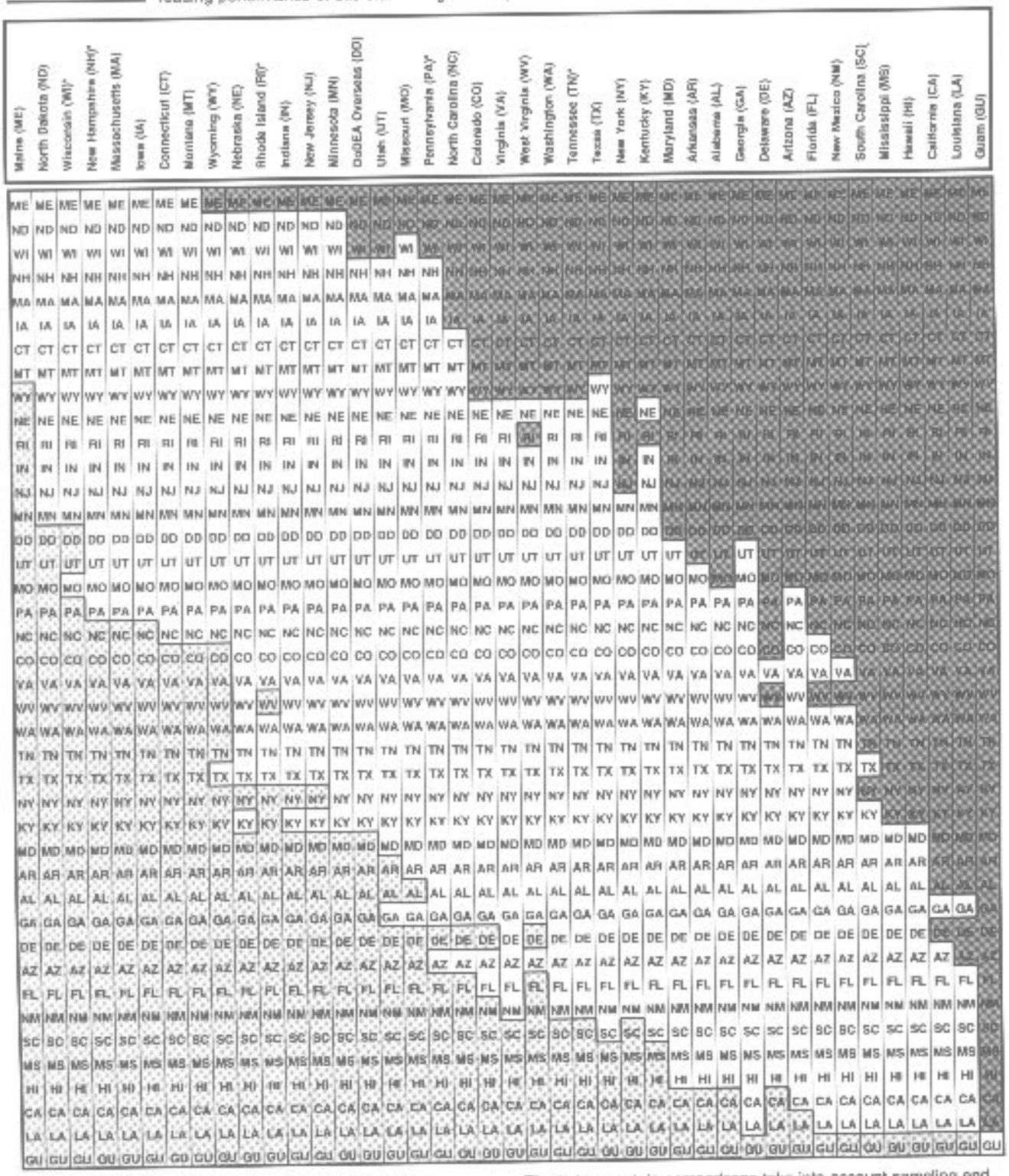
The center ***darkest*** box indicates a simultaneous confidence interval around the average reading proficiency for the state based on the Bonferroni procedure for multiple comparisons. The ***darker shaded*** boxes indicate the ranges between the 25th and 75th percentiles of the reading proficiency distribution, and the ***lighter shaded*** boxes the ranges between the 10th to 25th percentiles and the 75th to 90th percentiles of the distribution.

*Did not satisfy one or more of the guidelines for sample participation rates (see Appendix for details).

THE NATION'S REPORT CARD
1992
1994
Reading Assessment

Figure C.2 Comparisons of Overall Reading Average Proficiency for the 1994 Trial State Reading Assessment, Grade 4, Public Schools Only

INSTRUCTIONS: Read *down* the column directly under a state name listed in the heading at the top of the chart. Match the shading intensity surrounding a state postal abbreviation to the key below to determine whether the average reading performance of this state is higher than, the same as, or lower than the state in the column heading.



State has statistically significantly higher average proficiency than the state listed at the top of the chart.

No statistically significant difference from the state listed at the top of the chart.

State has statistically significantly lower average proficiency than the state listed at the top of the chart.

The between state comparisons take into account sampling and measurement error and that each state is being compared with every other state. Significance is determined by an application of the Bonferroni procedure.

*Did not satisfy one or more of the guidelines for sample participation rates (see Appendix for details).