04/16/2004 AERA 1 Updated 5/30/2004

Frequency of Simpson's Paradox in NAEP Data

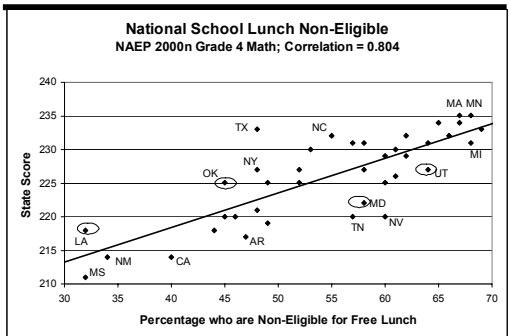
James Terwilliger
NAEP Coordinator, Minnesota Department of Education
Jim.Terwilliger@state.mn.us

Milo Schield
Augsburg College Department of Business Administration
Director, W. M. Keck Statistical Literacy Project
schield@augsburg.edu www.Augsburg.edu/ppages/~schield

American Educational Research Association
April 16, 2004
www.StatLit.org/PDF/2004TerwilligerSchieldAERA.pdf

04/16/2004 AERA 2 Updated 5/30/2004

State Scores by Family Income



04/16/2004 AERA 3 Updated 5/30/2004

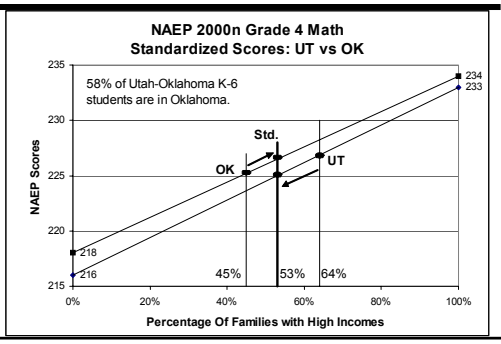
Simpson's Reversal by Family Income
OK overtakes UT; LA overtakes MD

State	All	High $	Low $
UT	227	233	216
OK	↓225↓	↑234↑	↑218↑

State	All	High $	Low $
MD	222	233	207
LA	↓218↓	233	↑211↑

04/16/2004 AERA 4 Updated 5/30/2004

Simpson's Reversal by Family Income: Oklahoma overtakes Utah



04/16/2004 AERA 5 Updated 5/30/2004

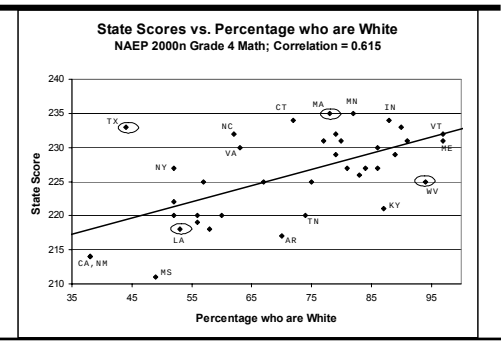
Simpson's Reversal by School Location:
NY overtakes MO; TN overtakes GA

State	All	City	Non-City
MO	229	216	233
NY	↓227↓	216	↑236↑

State	All	City	Non-City
GA	220	208	222
TN	220	↑213↑	↑224↑

04/16/2004 AERA 6 Updated 5/30/2004

State Scores by Race



04/16/2004 AERA 7 Updated 5/30/2004

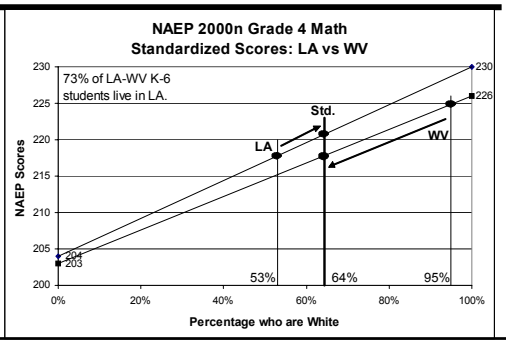
Simpson's Reversal by Race/Ethnicity: TX overtakes MA; LA overtakes WV

State	All	White	Black	Hispanic	Asian
MA	235	241	210	208	237
TX	↓233↓	↑243↑	↑220↑	↑224↑	↑247↑

State	All	White	Black
WV	225	226	203
LA	↓218↓	↑230↑	↑204↑

04/16/2004 AERA 8 Updated 5/30/2004

Simpson's Reversal by Race: Louisiana Overtakes West Virginia



04/16/2004 AERA 9 Updated 5/30/2004

Conditions for Simpson's Reversal

'*Simpson's reversal*' occurs (state A overtakes B) if:

1. Overall score in state A is lower in than in state B.
2. All subgroup scores in state A are at least as high as in state B.
3. At least one subgroup score in state A is higher than in state B.

A '*change*' occurs when condition (1) is replaced by:

4. Overall score is no higher in State A than in B

04/16/2004 AERA 10 Updated 5/30/2004

4% to 10% of Statistically Significant Differences are Reversed by Race

2002 Grade 8 Reading		Statistically Significant		
Confounder	**States**	**Pairs**	**Reverse**	**%**
School Lunch	40	505	1	0.2%
School Location	39		0	0%
Race/Ethnicity: All	40	505	52	10%
Race: White vs. Non	40	505	18	4%

52 reversals of statistically significant differences shows Simpson's Paradox is not a rare phenomena.

Statistical significance is obtained from the NAEP Data Tool.
'Non-white' includes blacks, Hispanics and Asians.
The 'all four groups' results are more disputable than the white/non-white.

04/16/2004 AERA 11 Updated 5/30/2004

Simpson's Paradox is not rare in NAEP data

NAEP 2002 Grade 8 Reading data:

- Absolute: 52 statistically significant differences are reversed by race
- Relative: 10% of statistically significant differences are reversed by race.

Some score differences are quite large:

- 14 points: California overtakes West Virginia

All Simpson's reversals are '**journalistically significant**'

04/16/2004 AERA 12 Updated 5/30/2004

Simpson's Paradox and NAEP: Recommendations

1. List state scores or ranks within relevant subgroups (e.g., school lunch, race/ethnicity)
2. Adjust state scores for non-school factors (other than race/ethnicity) such as student socio-economic status
3. Adjust state scores after controlling for the race/ethnicity of students
4. Increase sample sizes so a two point difference is statistically significant