Data Visualization for Impact and Implementation Findings

Carter Epstein and RJ de la Cruz

October 18-19, 2022

Asking Questions

- Participant microphones are automatically muted upon entry to limit background noise, unintentional feedback, or interference.
- During this session, you may ask the questions using the chat feature in Zoom.

- If your question is specific to your situation, or you need technical assistance, then the facilitator may respond to you directly in the chat panel.
- If your question is more general, then the facilitator will share your question with the presenter(s) so it can be answered at the appropriate time with the entire group.
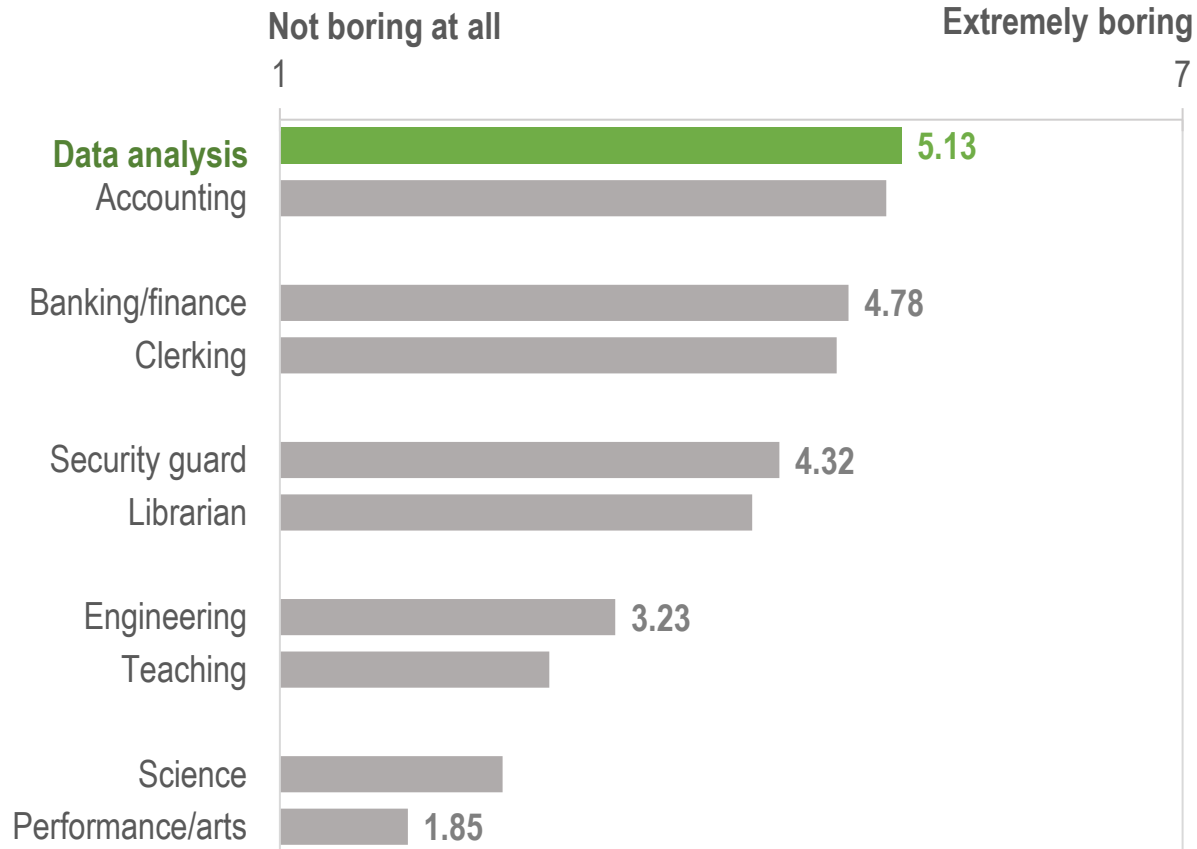
Agenda

- Benefits of effective data visualization

- What makes a "good" data display?

- How you can use data displays to

  - Demonstrate impact

  - Describe implementation

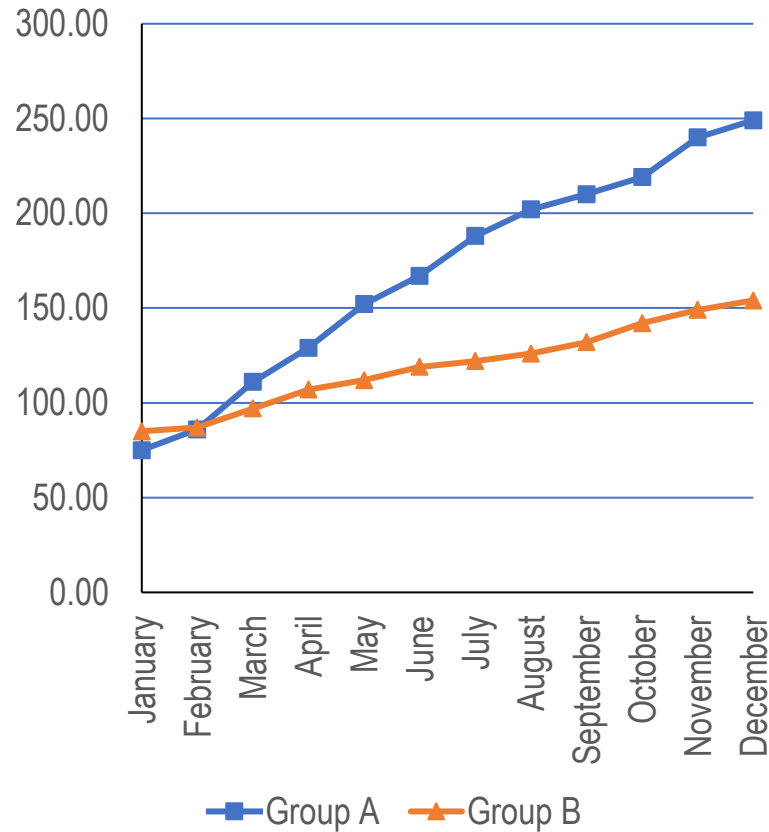- Resources

Benefits of effective data visualizations

Not boring at all
1

Extremely boring
7

Data analysis	5.13
Accounting	
Banking/finance	4.78
Clerking	
Security guard	4.32
Librarian	
Engineering	3.23
Teaching	
Science	
Performance/arts	1.85

- **Grab attention**

- **Convey key findings more quickly**

- **Make key points more memorable**

van Tilburg, W. A. P., Igou, E. R., & Panjwani, M. (2022). Boring people: Stereotype characteristics, interpersonal attributions, and social reactions. *Personality and Social Psychology Bulletin.* https://doi.org/10.1177/01461672221079104
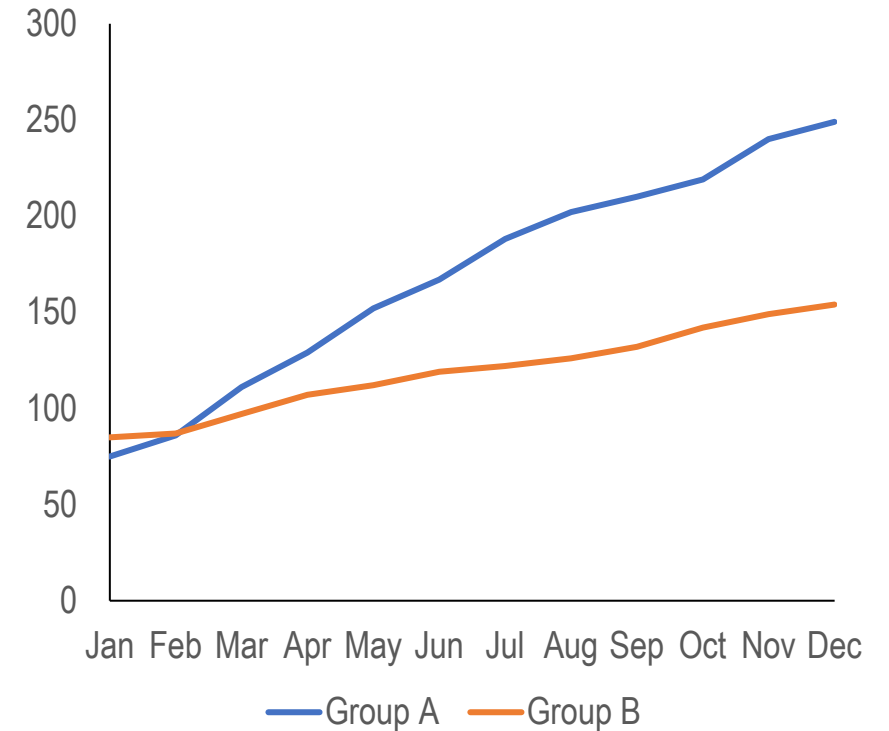
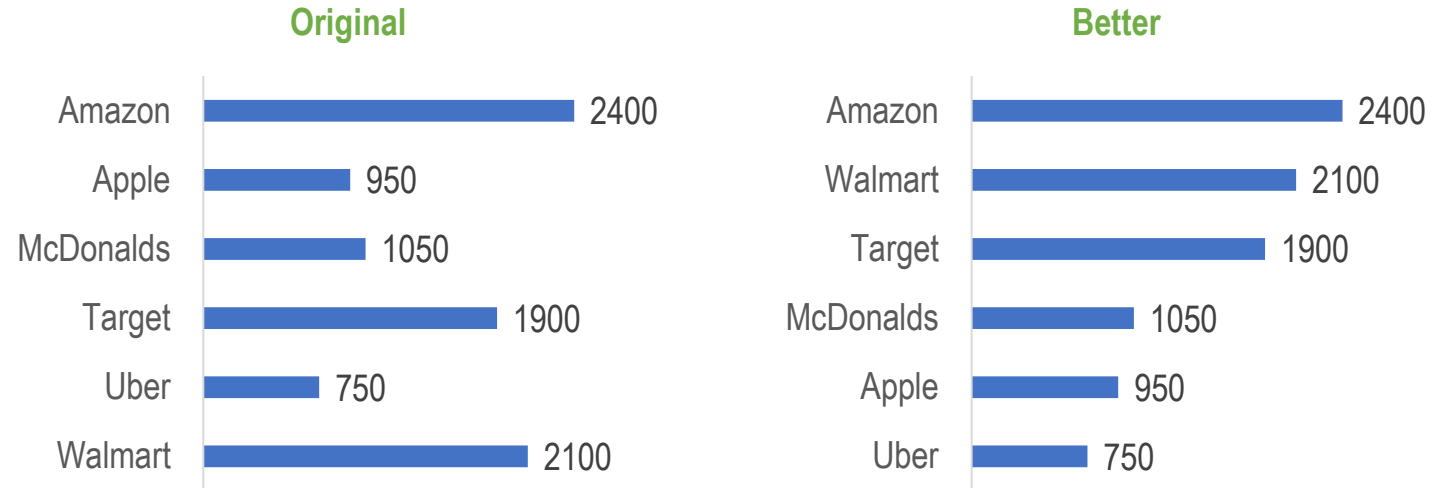# What makes a "good" data display?

Reduce clutter

Original

Better

What makes a "good" data display?(2)

Reduce clutter

Order data intentionally
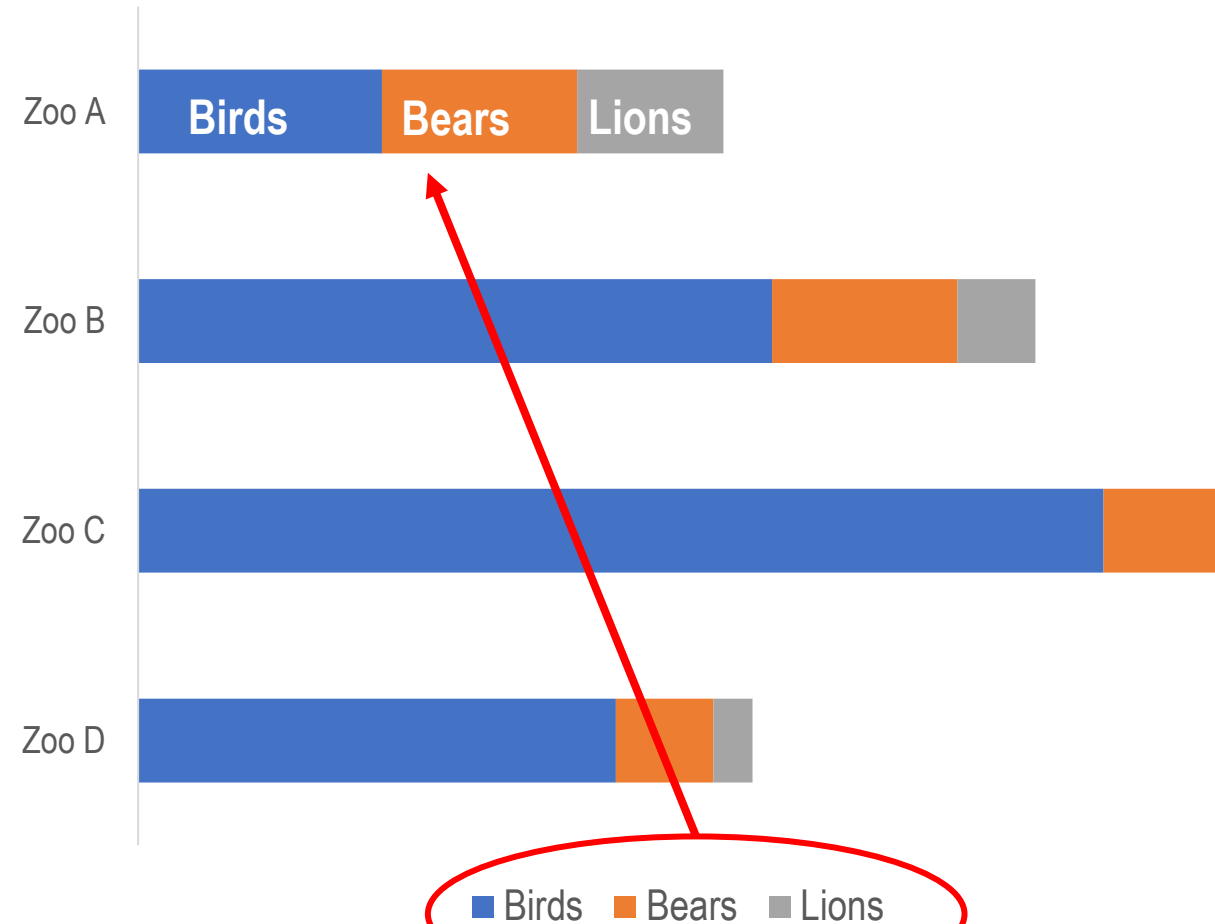
What makes a "good" data display?(3)

Reduce clutter

Order data intentionally

Place text and labels near the data

What makes a "good" data display?(4)
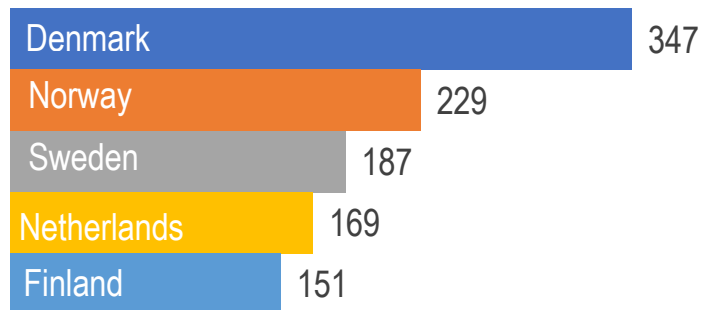
Reduce clutter

Order data intentionally

Place text and labels near the data

Use color strategically

Original

Average number of minutes per week
teaching reading

Denmark	347
Norway	229
Sweden	187
Netherlands	169
Finland	151

Better

Average number of minutes per week
teaching reading

Denmark	347
Norway	229
Sweden	187
Netherlands	169
Finland	151

What makes a "good" data display? (5)

Reduce clutter

Order data intentionally

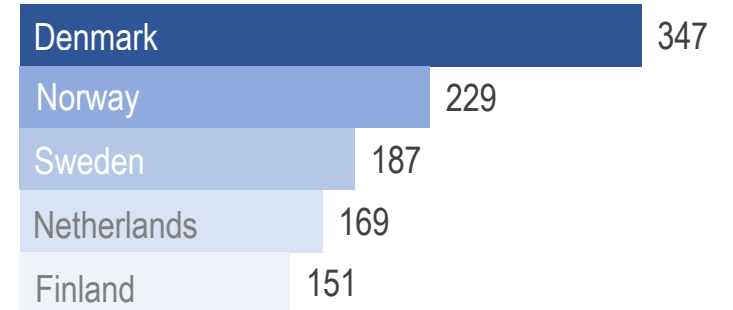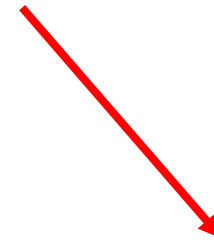Place text and labels near the data

Use color strategically

State the key point succinctly in a title

Europe's Greenhouse Gas Emitters

By a wide margin, Germany emitted the most greenhouse gas among European countries in 2017.

What makes a "good" data display?(6)

Reduce clutter

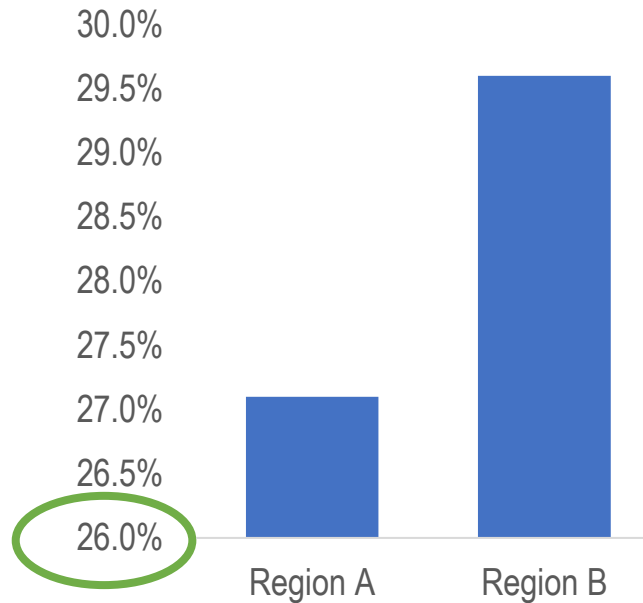Order data intentionally

Place text and labels near the data
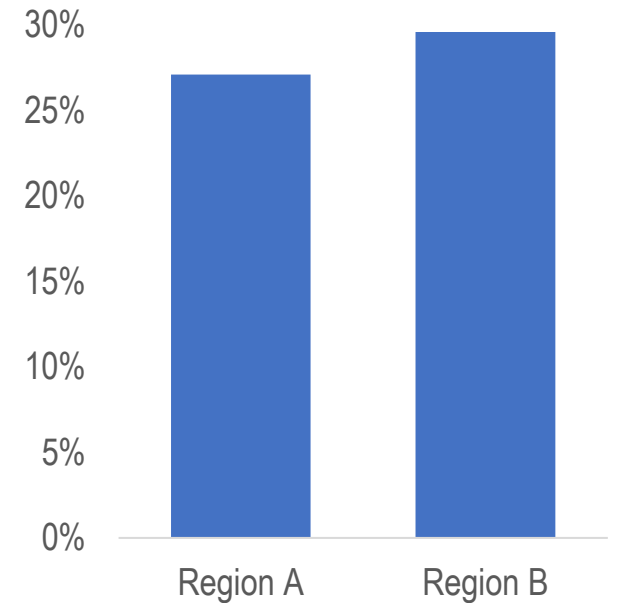
Use color strategically

State the key point succinctly in a title

Avoid misleading design choices

Truncated axis exaggerates difference

Better

Displaying demographics data

	High schools in the study	All high schools in the state
Female	51%	50%
Male	49%	50%
White	72%	74%
Black/African-American	21%	19%
Asian	5%	6%
Race other than Black, White, Asian	2%	1%
Hispanic/Latino	22%	18%
Eligible for FRPL	87%	45%
English language learners	31%	21%
Low achievement	69%	40%
Dropout rate	33%	15%

Easily add a default chart
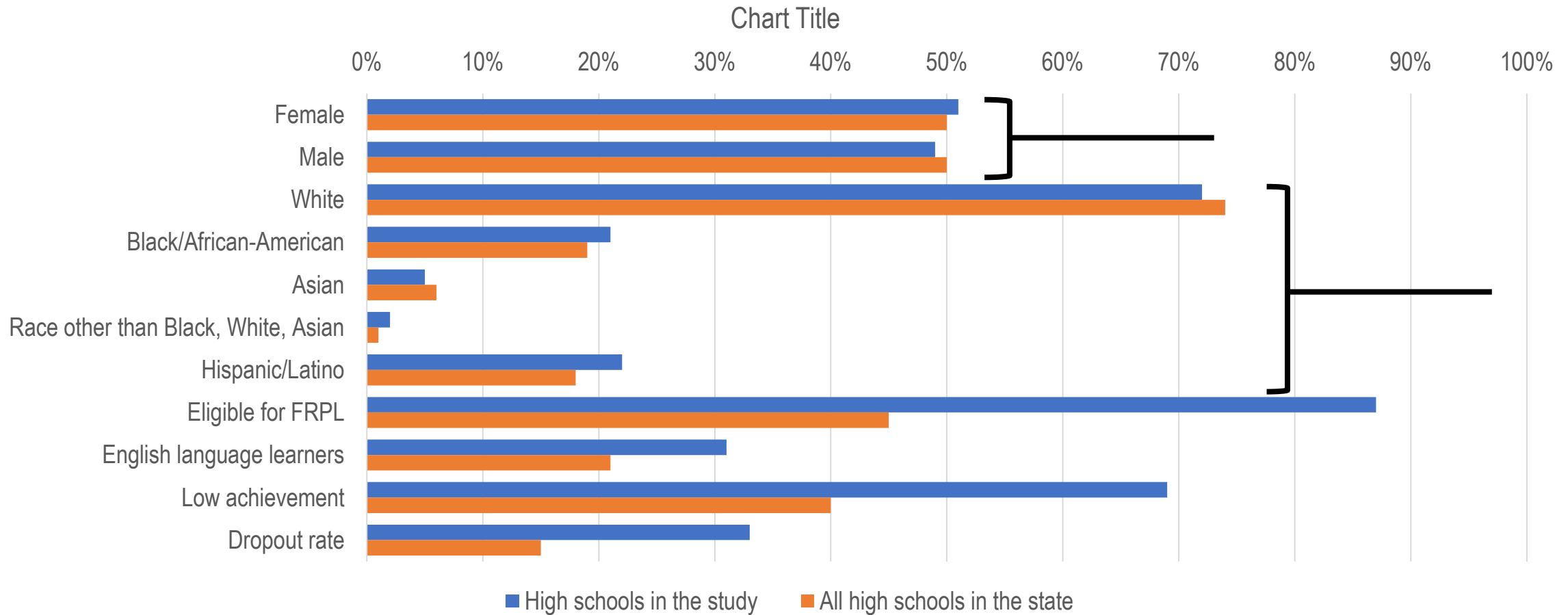
Keep the data that tells your story

Chart Title

High schools in the study ■ All high schools in the state ■
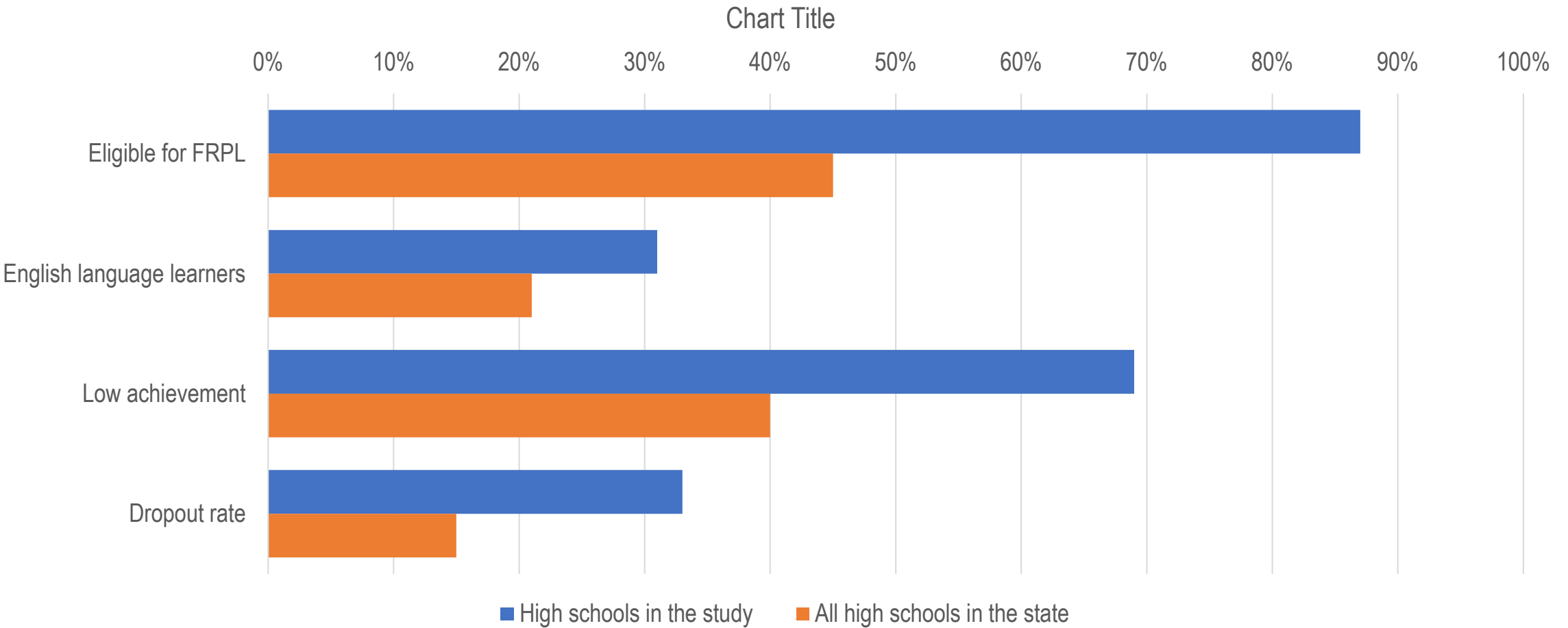
Remove clutter

Chart Title

Eligible for FRPL
- Study high schools: 87%
- All high schools: 45%

English language learners
- 31%
- 21%

Low achievement
- 69%
- 40%

Dropout rate
- 33%
- 15%
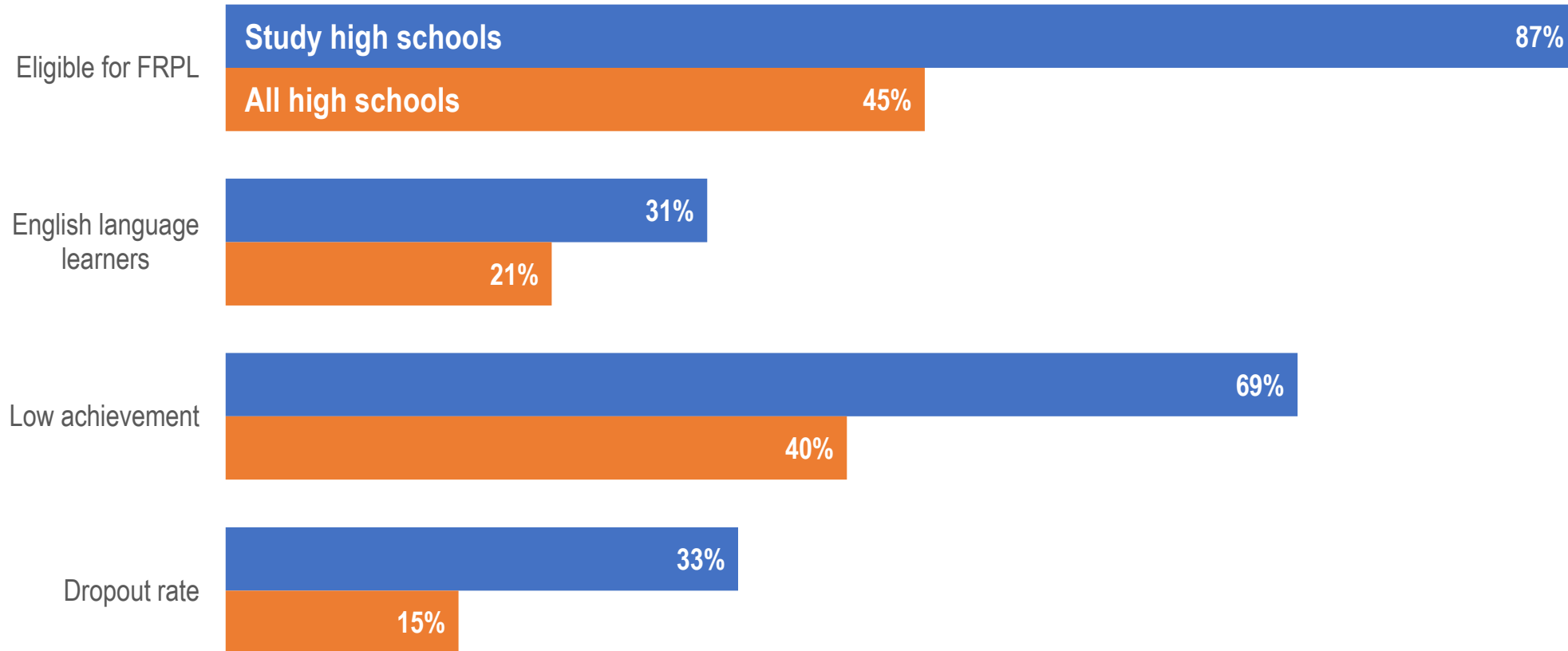
Add emphasis with color

Chart Title

Pick a simple, descriptive title

High needs schools were recruited for the study.

Eligible for FRPL
- Study high schools: 87%
- All high schools: 45%

English language learners
- 31%
- 21%

Low achievement
- 69%
- 40%

Dropout rate
- 33%
- 15%

Before and after

Chart Title

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

- Female
- White
- Asian
- Hispanic/Latino
- English language learners
- Dropout rate

■ High schools in the s...

High needs schools were recruited for the study.

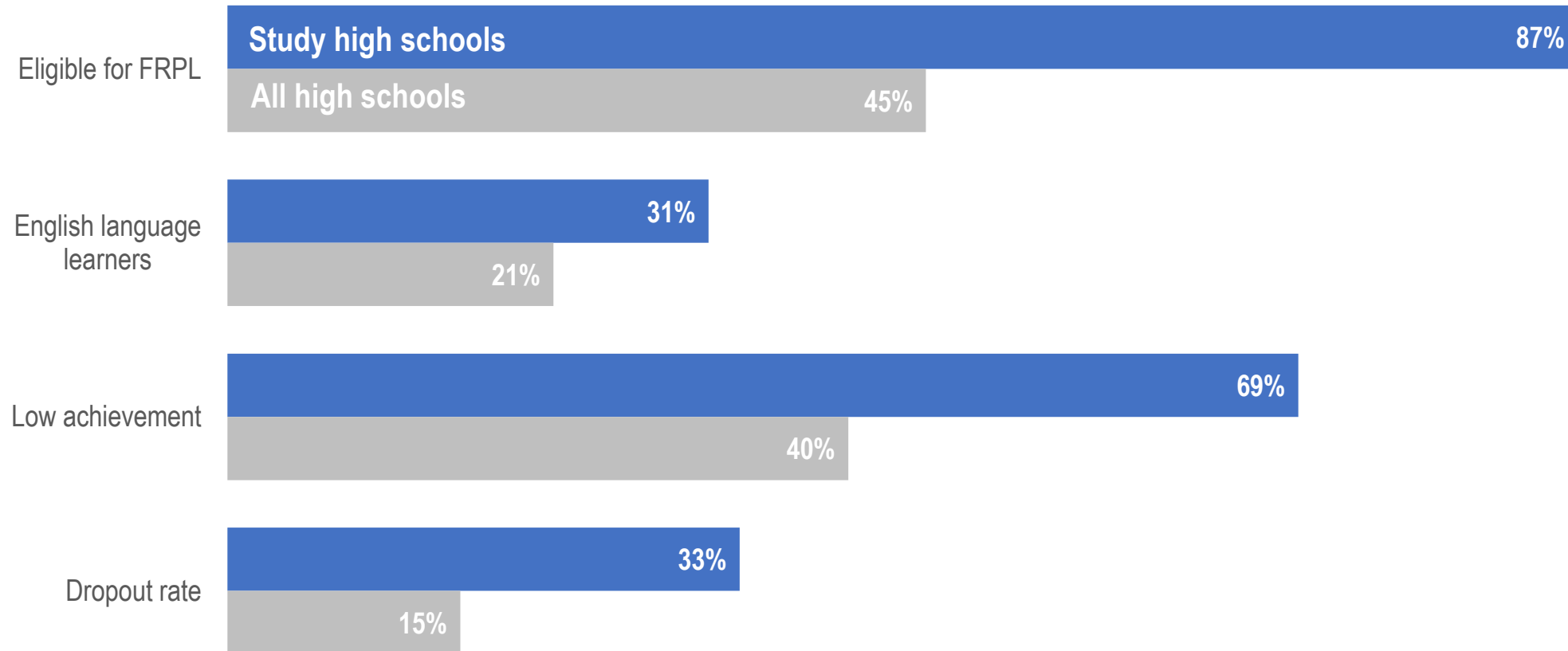|	Study high schools	All high schools
Eligible for FRPL	87%	45%
English language learners	31%	21%
Low achievement	69%	40%
Dropout rate	33%	15%

Visualizing impact findings

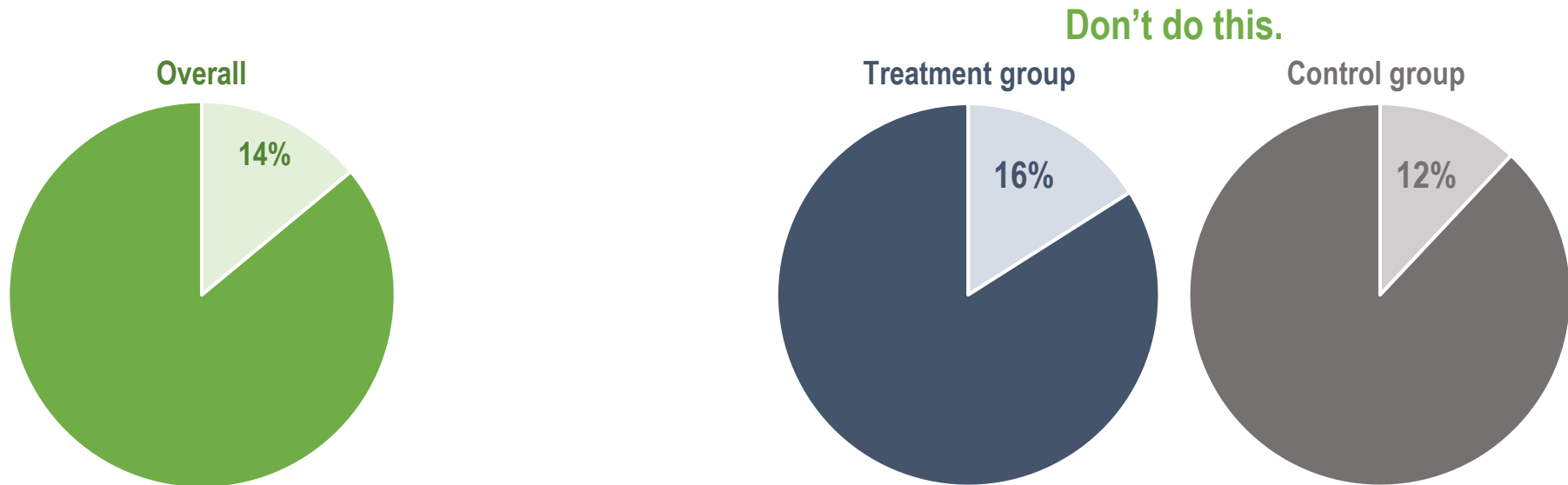- Displaying part-to-whole relationships (attrition)

- Comparing two quantities (baseline, outcome means)

  - Baseline equivalence

  - Outcomes

Use pie charts for one, but not multiple, part-to-whole relationships

Student attrition was **low overall** and differed little between the **treatment** and **control** groups.

Don't do this.

Overall

14%

Treatment group

16%

Control group

12%

Consider overlapping bar charts for comparing part-to-whole relationships

In both treatment and control groups, **math scores were available at follow-up for most of the students who were randomly assigned**.

Chart legend: ■ At random assignment ■ At followup

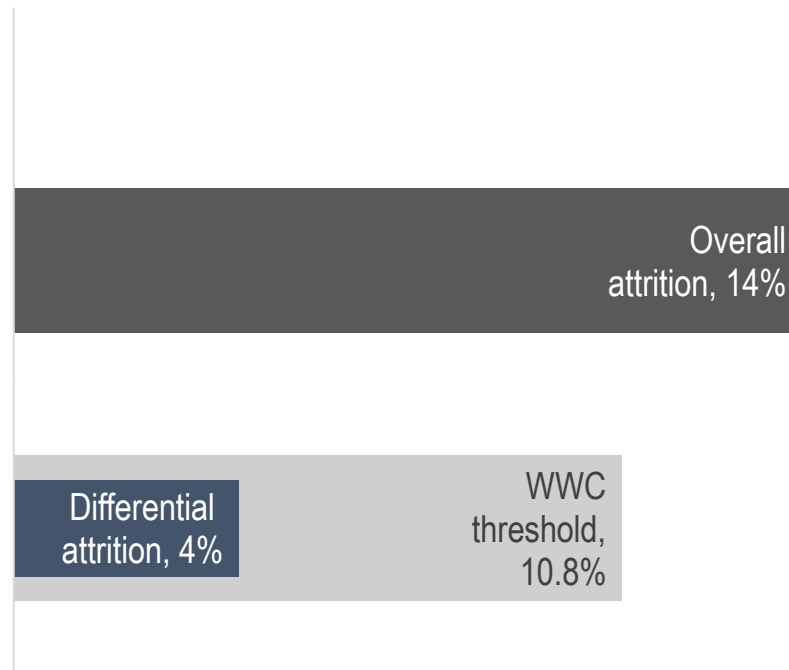| Group	At followup	Sample size
Total	86%	Students randomly assigned, N= 693
Treatment	84%	n=348
Control	88%	n=345

0%   100%
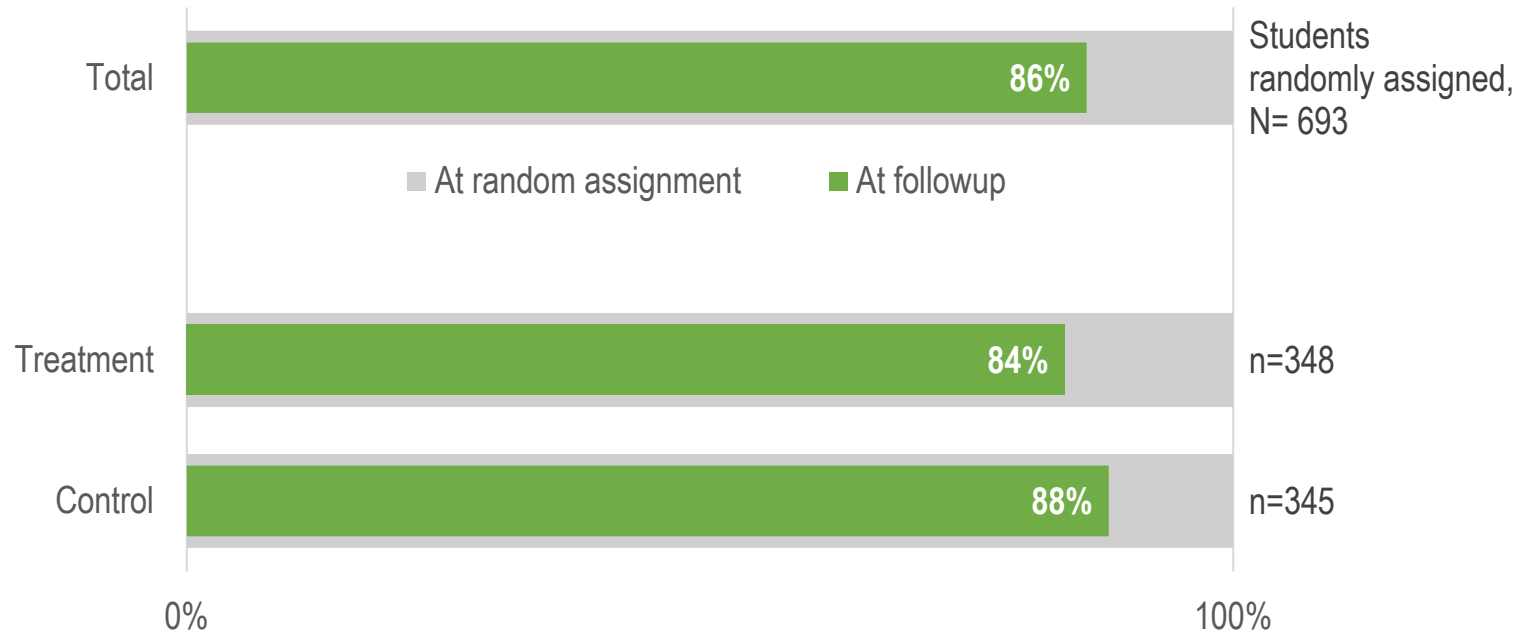
Or use overlapping bars to show performance relative to a benchmark

Student attrition was low overall and **differed little** between the treatment and control groups.

Overall attrition, 14%

WWC threshold, 10.8%

Differential attrition, 4%

Consider not using a graph at all

Sometimes, no data display is the better option

Eligible students included those who had scored in the "below basic" proficiency range on the spring 2016 state mathematics assessment and were in grades 6 and 7 at the time of this baseline testing. All eligible students were randomly assigned within school by grade level blocks. Across the 10 schools in the study, 348 students were assigned to the treatment group and 345 to the control group. At the start of the 2016-17 school year, we administered the IAAT to students once they had returned to school. The study team coordinated testing schedules and logistics with each school. All testing sessions, including one make-up session per school, occurred during a 3-week window in September 2016.

Attrition was within the What Works Clearinghouse boundary under the cautious assumption. The fall 2016 IAAT was completed by 292 students in the treatment group and 304 students in the control group. Of these 596 students, a spring 2016 state math score was available for 584 students. Pretest scores were imputed for the

Math scores were available at follow-up for
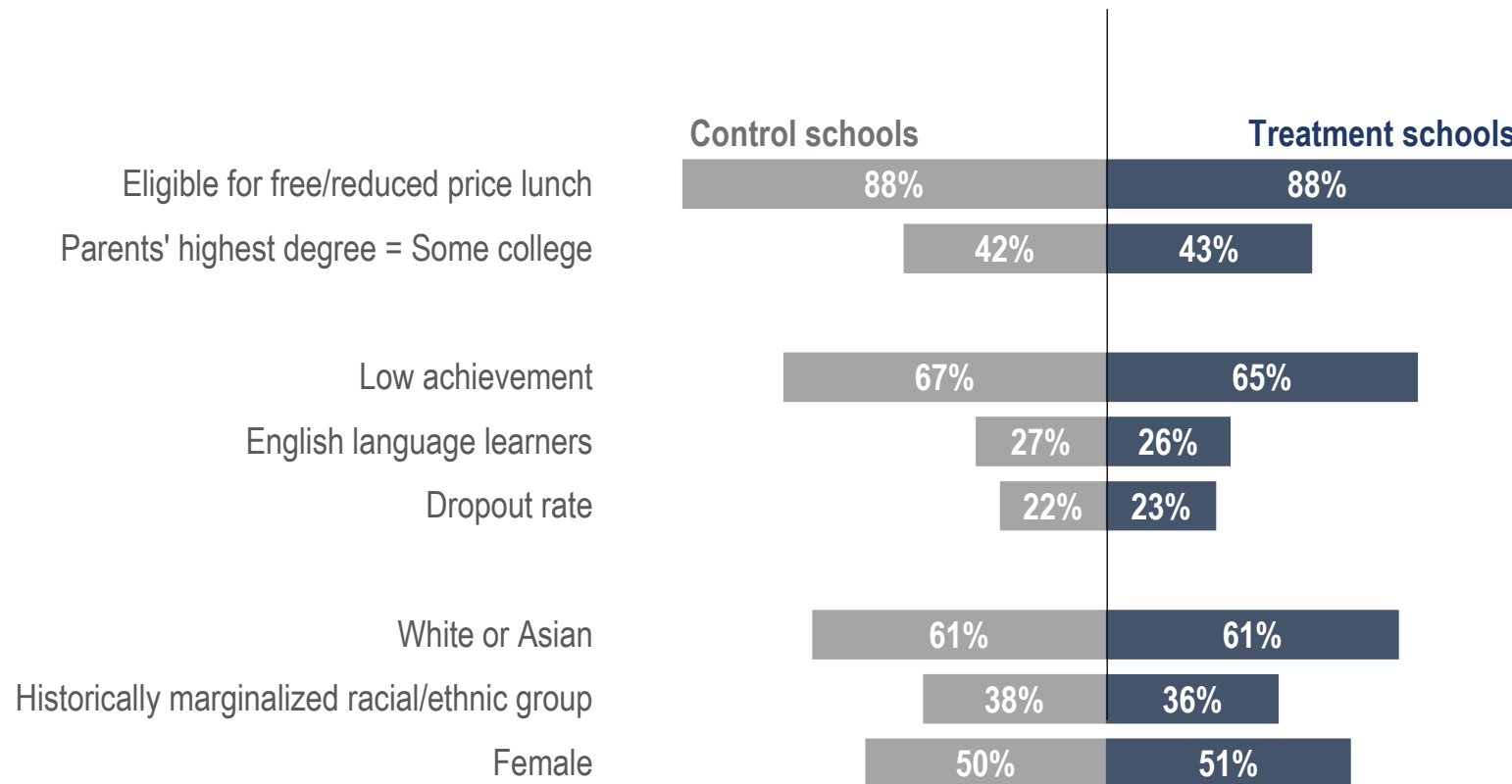
86%

of students overall

84% of the
Treatment
group

88% of the
Control
group

Consider using symmetry to show similarities and differences

After random assignment, schools in the **treatment** and **control** groups were similar on baseline characteristics.

|	Control schools	Treatment schools
Eligible for free/reduced price lunch	88%	88%
Parents' highest degree = Some college	42%	43%
Low achievement	67%	65%
English language learners	27%	26%
Dropout rate	22%	23%
White or Asian	61%	61%
Historically marginalized racial/ethnic group	38%	36%
Female	50%	51%

Avoid error bars on column/bar charts

Students in the **treatment** and **comparison** groups were equivalent on prior math achievement.

Spring 2016 mean mathematics scores and standard deviations

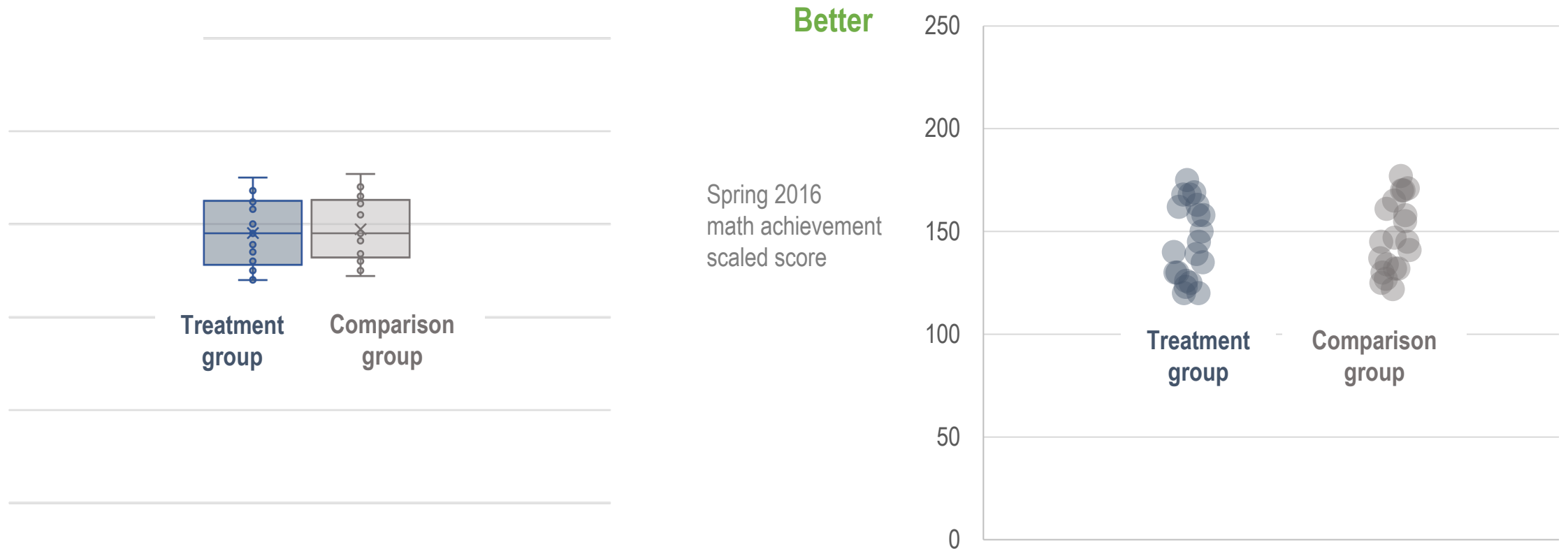**Often misinterpreted**
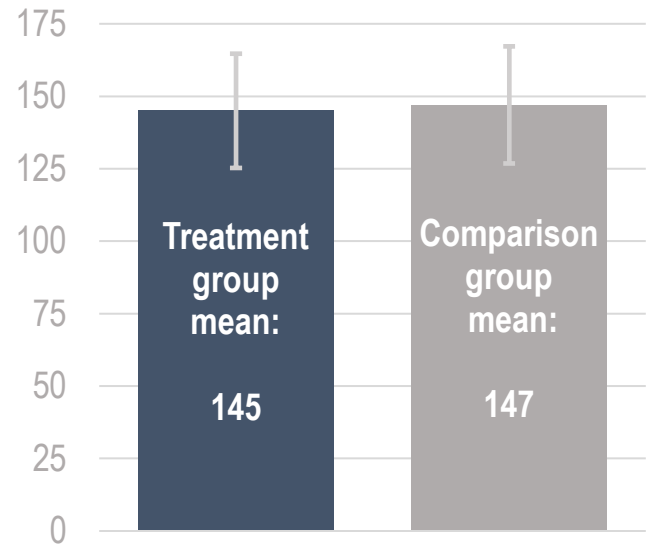
Better
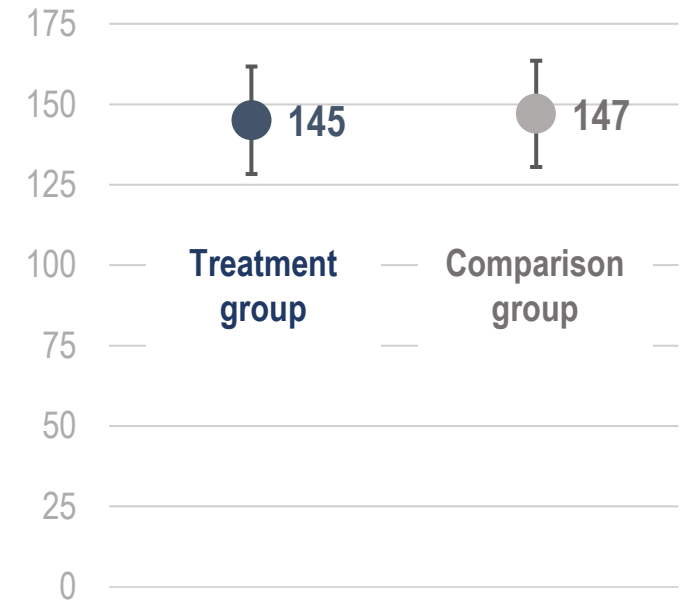
Avoid box plots, try jittered strip plots

Students in the **treatment** and **comparison** groups were equivalent on prior math achievement.

Bar or column charts work well for comparing two numbers

In each grade, students in the **program** group outperformed those in the **business-as-usual comparison** group.
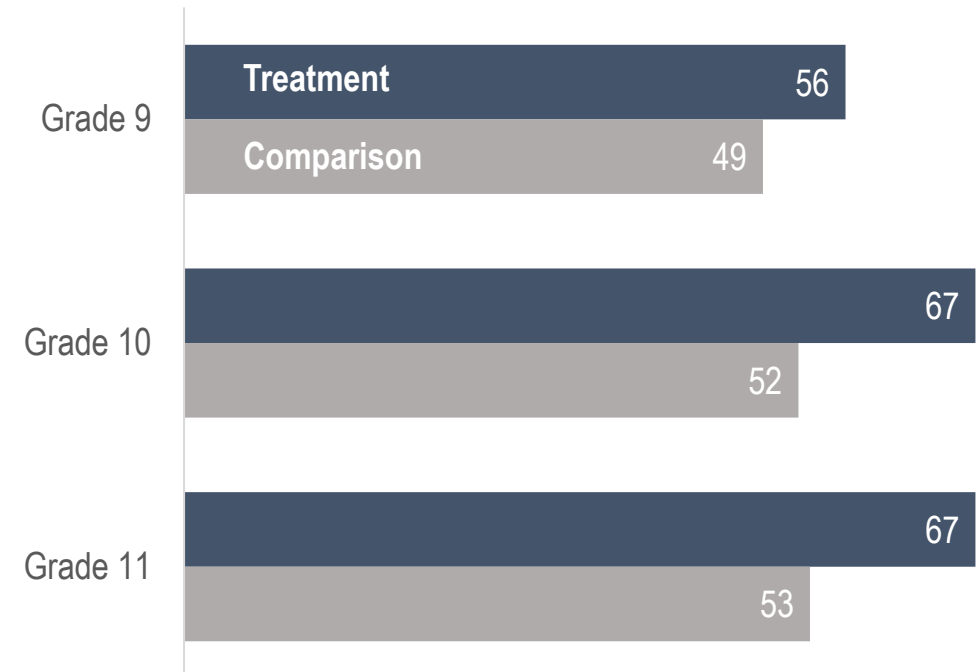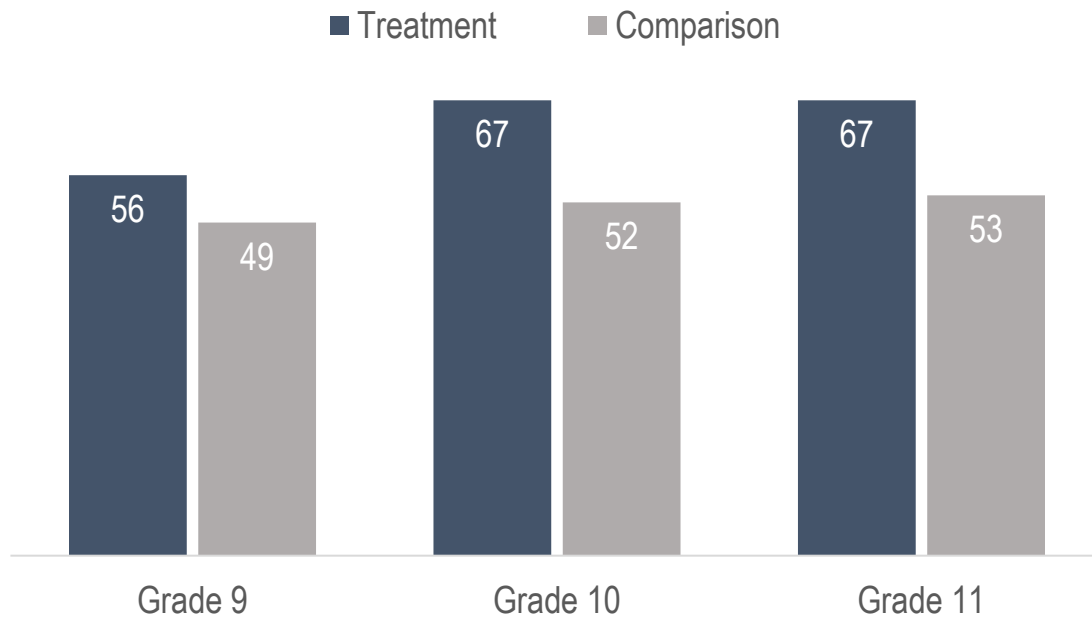
Columns or bars work equally well

Consider line graphs for pre-post differences

Students in **treatment schools** saw **greater improvement** in math achievement than students in control schools.
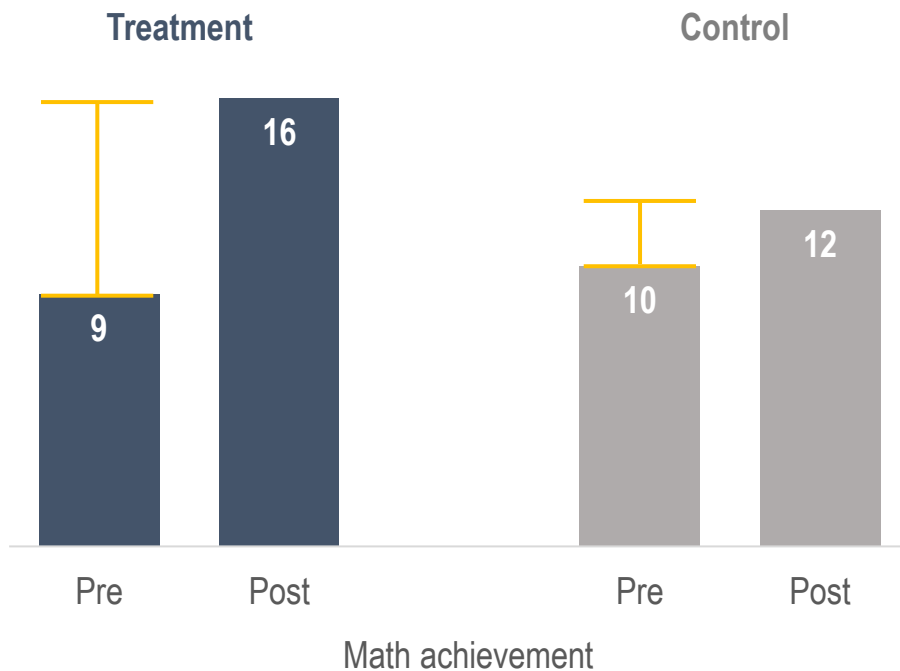
Don't crowd subgroup findings

Students in grades 9 and 10 in **treatment schools** saw **greater improvement in math achievement** than students in these grades in control schools.

Don't do this

Chart legend:
- Grade 9 treatment
- Grade 9 control
- Grade 10 treatment
- Grade 10 control
- Grade 11 treatment
- Grade 11 control

Y-axis: Unadjusted mean score

X-axis: Pre, Post

Data values — Pre: 11, 10, 9; Post: 18, 16, 15, 14, 12, 11
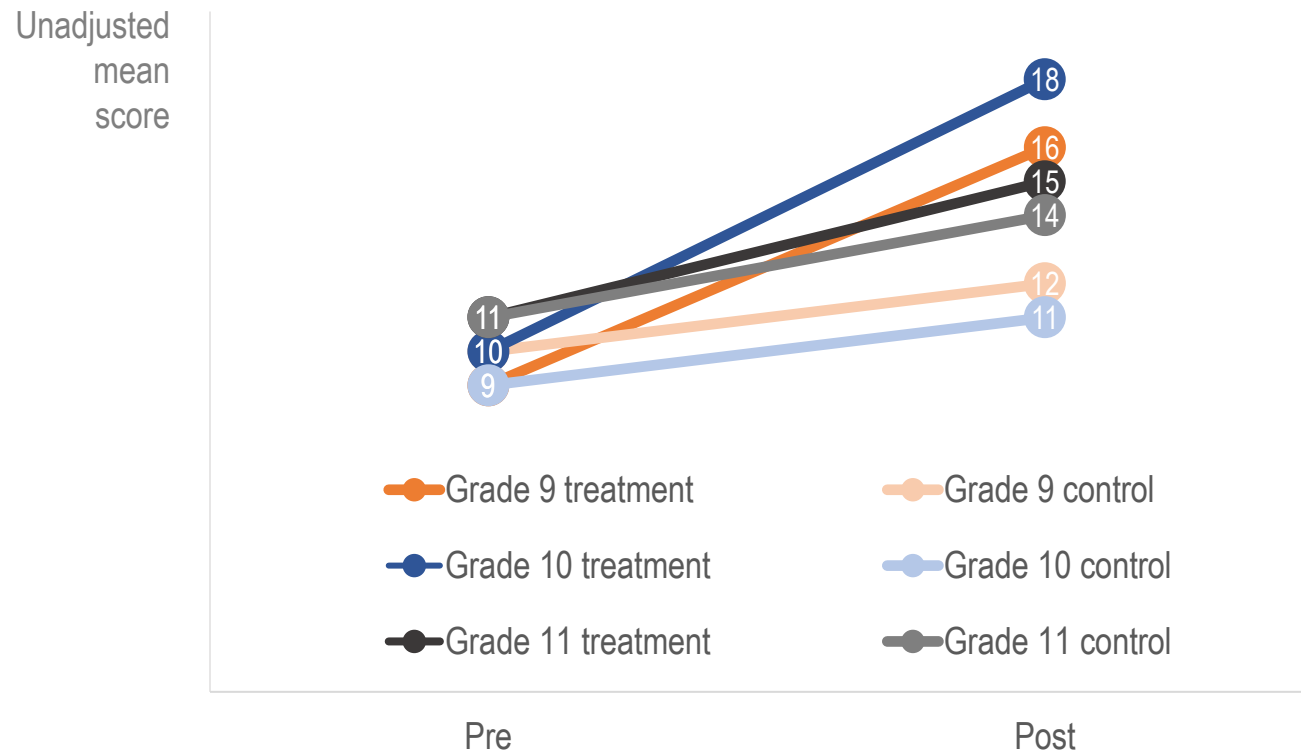
Give each subgroup its own pre-post graph

Students in grades 9 and 10 in **treatment schools** saw **greater improvement in math achievement** than students in these grades in control schools.

Consider dumbbell plots for pre-post differences

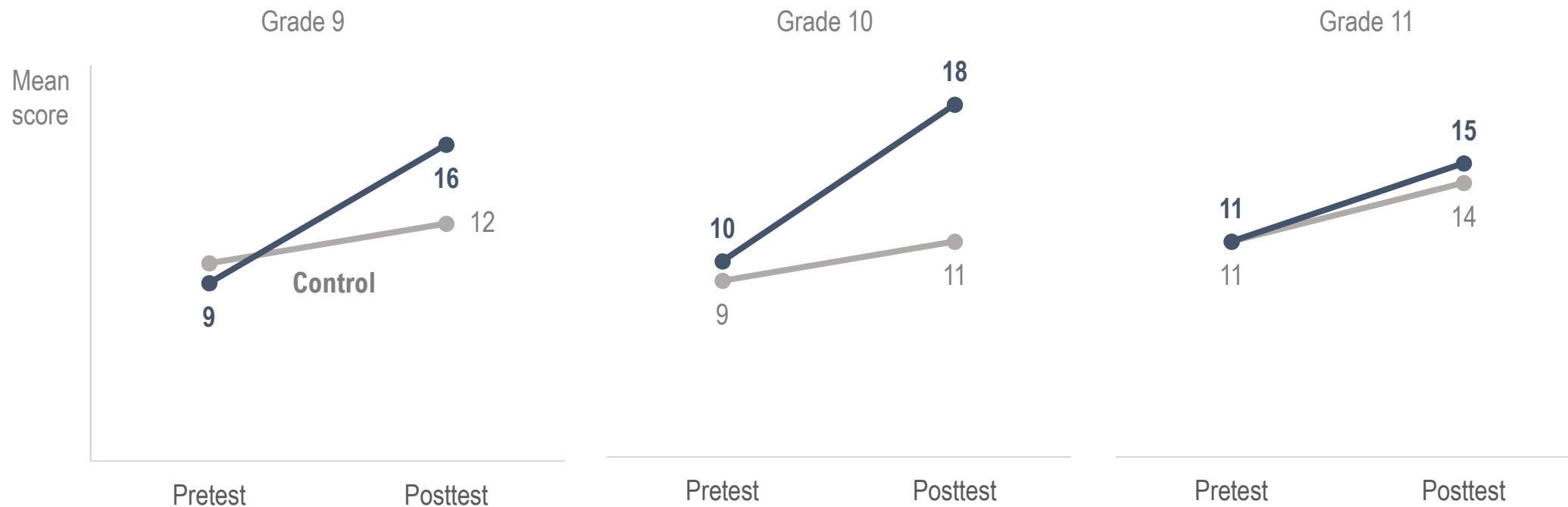Students **in grades 9 and 10** in treatment schools saw **greater improvement in math achievement** than students in these grades in control schools.

Grade 9*
- Treatment: 9.2 → 15.0
- Control: 9.5 → 11.2

Grade 10*
- 15.5 → 18.9
- 14.8 → 16.0

Grade 11 (ns)
- 18.4 → 21.4
- 18.1 → 19.7

○ Pretest   ● Posttest

Grade 9*
- Treatment: 9 → 15
- Control: 10 → 11

Grade 10*
- 16 → 20
- 15 → 16

Grade 11
- 18 → 21
- 18 → 20
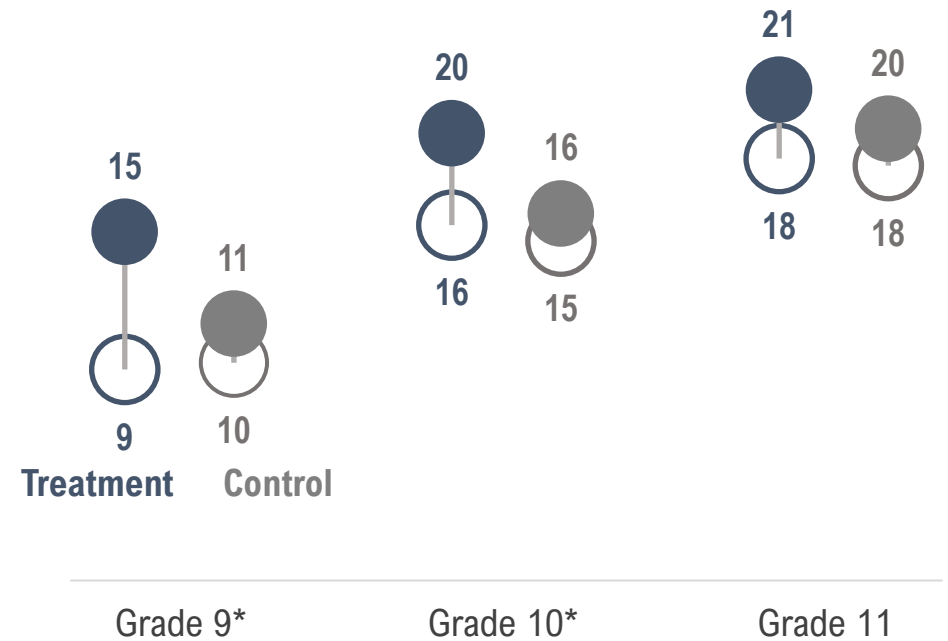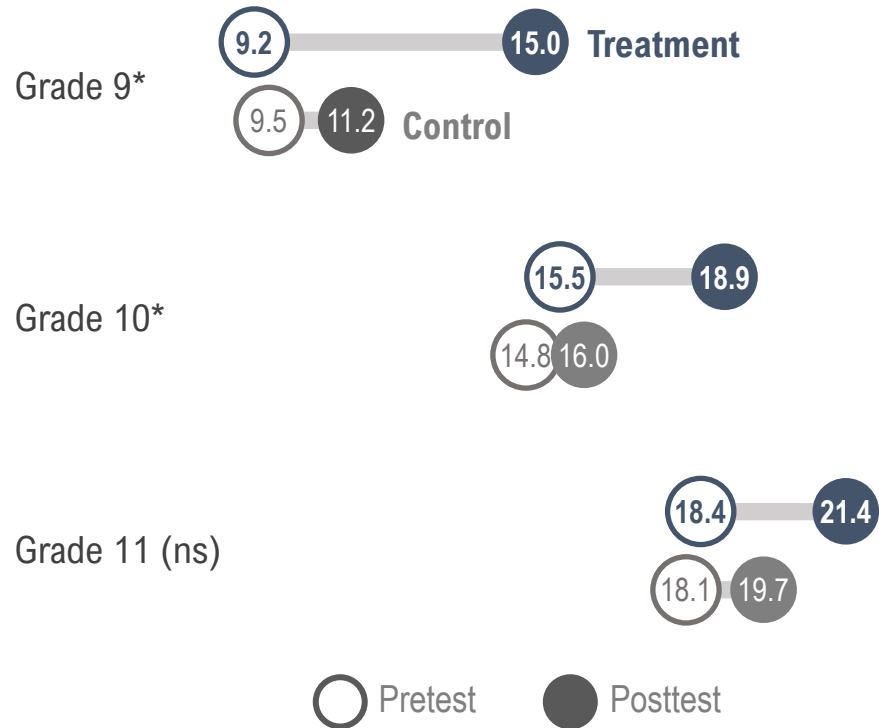
Visualizing implementation findings

- Were key components implemented with fidelity?

  - What did the intervention developer do?

  - What did schools and teachers do?

- How satisfied were schools and teachers?

Consider overlapping bars for fidelity of implementation findings

The program met its goals for rates of teacher participation in the summer training institute and school year coaching but fell short in other areas.

Average percentage of teachers across schools implementing the program

Chart showing threshold (light gray) and actual percent (dark blue) bars:

- Attended 3 days of summer training institute: **92.3%**
- Passed 3 knowledge checks: **86.5%**
- Participated in at least 6 of 8 data meetings: **76.9%**
- Participated in at least 2 of 3 coaching sessions: **86.5%**

Legend: Threshold, Actual percent
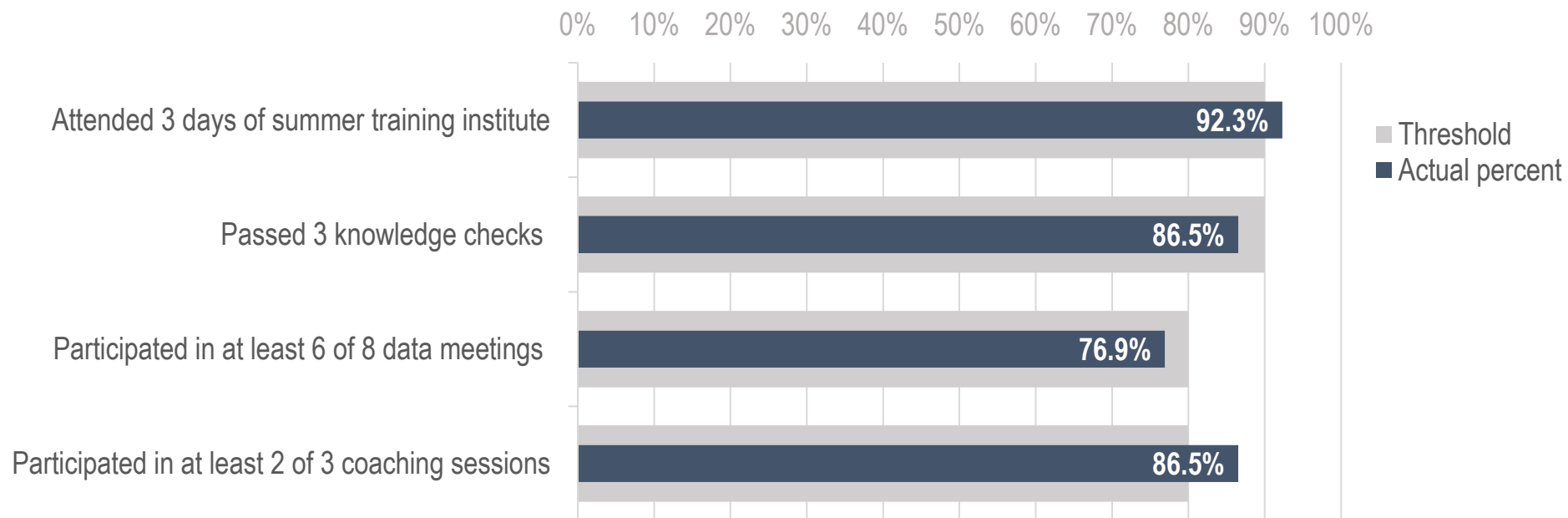
Consider overlapping bars for fidelity of implementation findings(2)

The program met its goals for rates of teacher participation in the summer training institute and school year coaching but fell short in other areas.

Average percentage of teachers across schools implementing the program

Legend:
- Threshold
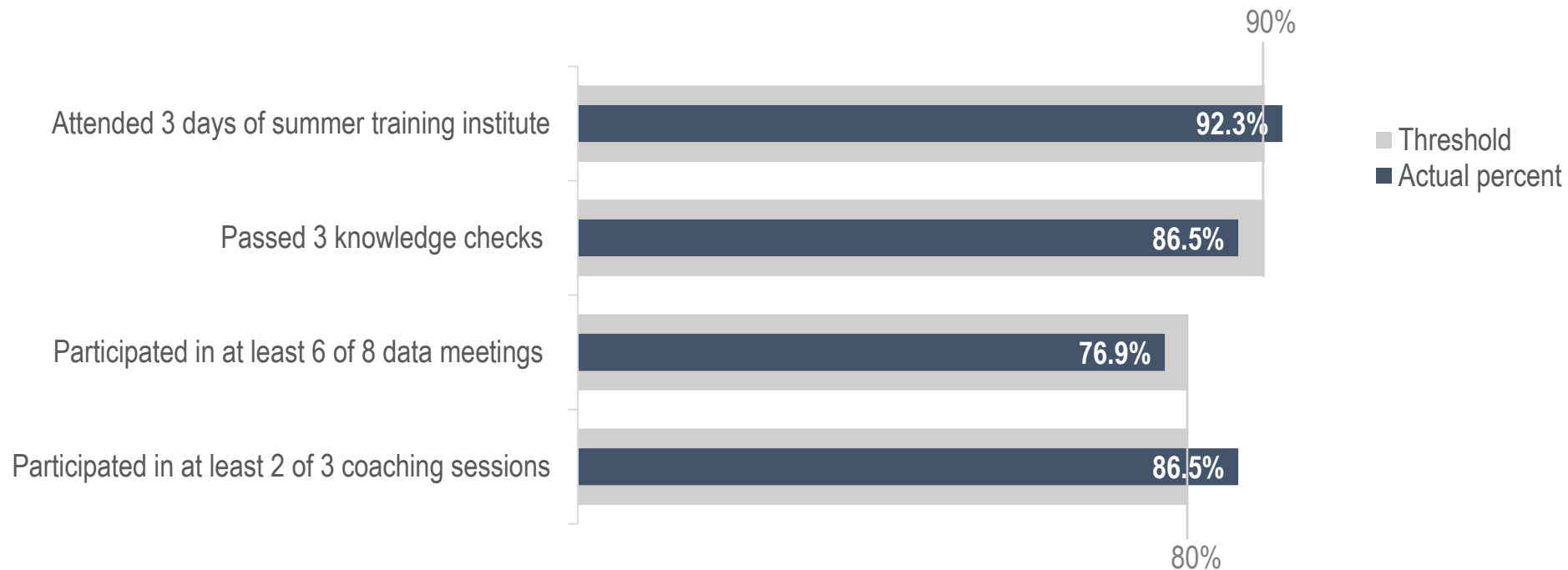- Actual percent

Chart values:
- Attended 3 days of summer training institute: 92.3%
- Passed 3 knowledge checks: 86.5%
- Participated in at least 6 of 8 data meetings: 76.9%
- Participated in at least 2 of 3 coaching sessions: 86.5%

Axis: 80% – 90%

Consider simple bar charts for Likert scale means

On average, teachers rated **reviewing videos of their instruction** as a highly effective strategy

	1= Highly ineffective	4= Highly effective
Reviewing videos of my instruction with my coach		3.7
Summer in-person training	2.4	
Summer reflection assignments	1.9	
Mid-year refresher training	1.3	
Online community of practice	1.2	

Try diverging stacked bars to show percent "agreement" on Likert scales

Most teachers felt that the **summer training** and **reviewing videos of their own instruction** helped them implement the program curriculum in the classroom

	Highly ineffective	Ineffective	Effective	Highly effective
Reviewing videos of my instruction with my coach		1%	10%	83%
Summer in-person training		2%	18%	45%	35%
Summer reflection assignments	18%	42%	25%	15%
Mid-year refresher training	10%	52%	25%	13%
Online community of practice	20%	50%	15%	15%

Effective   Highly effective

Highly ineffective   **Ineffective**

Resources

- Color:
  - https://giggster.com/guide/complementary-colors/
  - https://www.accessibilitychecker.org/color-contrast-checker/
  - https://toolness.github.io/accessible-color-matrix/
- How to:
  - https://stephanieevergreen.com/diverging-stacked-bars (and her website in general)
  - https://flowingdata.com/ (for R users)
- General data viz
  - https://nightingaledvs.com/tag/data-visualization/
  - http://www.perceptualedge.com/ (and Stephen Few's books)
  - https://www.storytellingwithdata.com/blog
  - https://policyviz.com/blog/
- Equity in data science:
  - https://weallcount.com/

Questions?

Contact

EIR Evaluation TA Team

EIREvalTA@AbtAssoc.com

BOLD
THINKERS
DRIVING
REAL-WORLD
IMPACT

www.abtassociates.com