Day 01 - Lesson 05

Data Language, Needs, Uses, and Relation to Effect Sizes

Ani Ruhil

EBITE 2022

Updated 2022-06-20

Agenda

- The "Typical" student ... no Heffalump for sure!
- Visualize, visualize, visualize
- Crazy little thing called *Confounding*
- *Equivalence*, *Attrition*, & *Bias*

Typical: The Mean

- Numeric data used ... test scores, number of days absent, number of disciplinary incidents, and so on
- Question of interest: How did the "average", the "typical" student do?
- Default answer will be from calculating the Mean ... add all students' scores, divide by the number of students, and voila!
  - Example: 7 students' scores on 10-point quiz: 1, 5, 6, 7, 7, 10, 10
  - Mean = 6.6

Typical: The Median

- Mean intuitive, extensively used, but influenced by a few extreme scores ... notice that 4 out of 7 students scored 7 or higher but the "average" is **< 7**, because of the student who scored **1**
- In these situations the Median better represents the "average", the "typical" score
- Calculated by lining up the scores in ascending or descending order and finding the mid-point
    - Example: 7 students' scores on 10-point quiz: 1, 5, 6, 7, 7, 10, 10
    - Midpoint is highlighted: 1, 5, 6, **[7]**, 7, 10, 10
    - Notice how the "average" now is 7

Typical: The Mode

- Designed for non-numeric data (Gifted, LEP, FRL, Gender, etc)
- Most intuitive of the three because people understand “the typical student is a Gifted student”, “the typical student in our district is rated Above Proficient on the 3rd Grade ELA assessment”, etc.
- You can flip numeric data into non-numeric data by grouping numeric data

Example-01 Example-02

Scale Score Ranges in English Language Arts, 2017-18

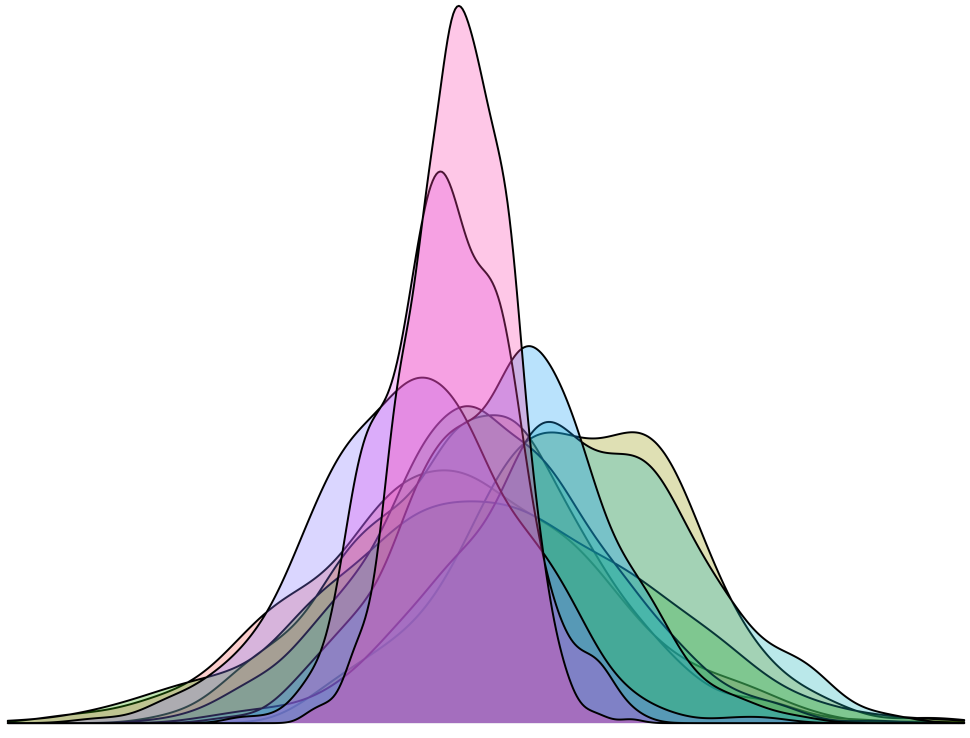
Grade/Subject	Limited	Basic	Proficient	Accelerated	Advanced
Grade 3	545-671	672-699	700-724	725-751	752-863
Grade 4	549-673	674-699	700-724	725-752	753-846
Grade 5	552-668	669-699	700-724	725-754	755-848
Grade 6	555-667	668-699	700-724	725-750	751-851
Grade 7	568-669	670-699	700-724	725-748	749-833
Grade 8	586-681	682-699	700-724	725-743	744-805
English language arts I	606-682	683-699	700-724	725-738	739-800
English language arts II	597-678	679-699	700-724	725-741	742-808

Example-01 Example-02

Third Grade Reading Guarantee Performance Level Scale Score Ranges, 2017-18

Grade and Subject	Does not meet promotion standards	Meets promotion standards
Grade 3 ELA Scale Score	Below 672	At or Above 672

Don't forget variability!



Range = (Maximum - Minimum) a useful rough measure of variability

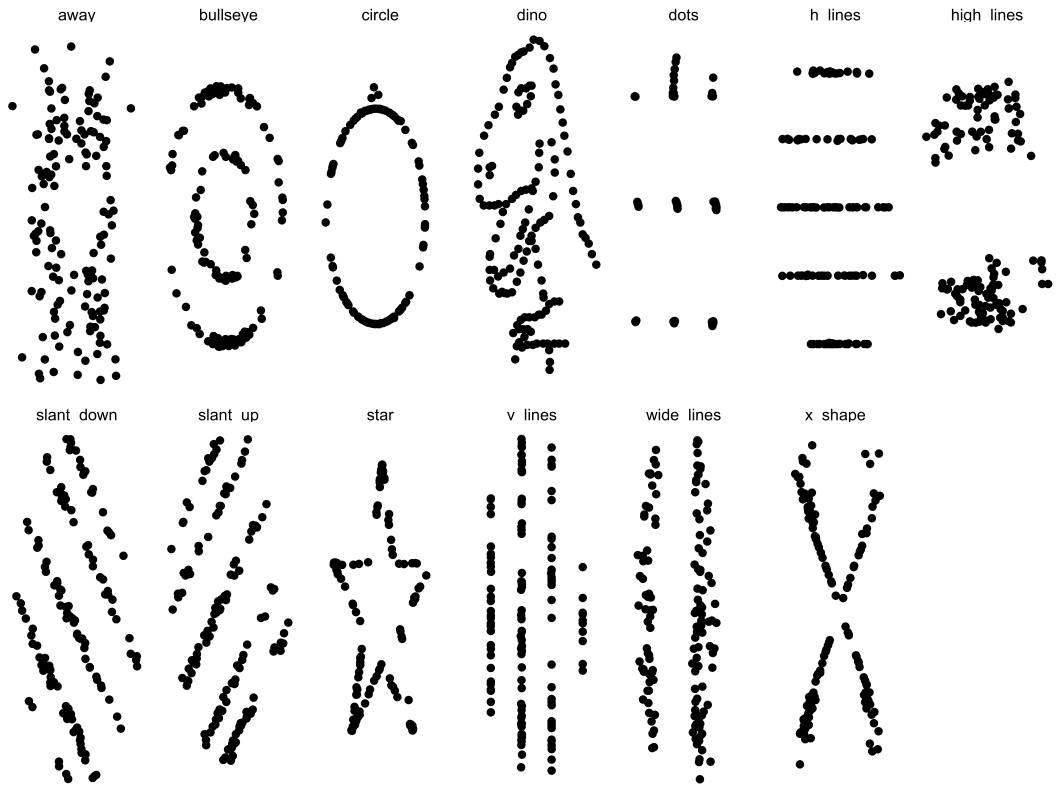
Always look at your data!!

How things look | How things really are

dataset	Mean of x	Mean of y	Variation of x
away	54.3	47.8	26.9
bullseye	54.3	47.8	26.9
circle	54.3	47.8	26.9
dino	54.3	47.8	26.9
dots	54.3	47.8	26.9
h_lines	54.3	47.8	26.9
high_lines	54.3	47.8	26.9
slant_down	54.3	47.8	26.9
slant_up	54.3	47.8	26.9
star	54.3	47.8	26.9
v_lines	54.3	47.8	26.9
wide_lines	54.3	47.8	26.9
x_shape	54.3	47.8	26.9

Always look at your data!!

How things look | How things really are



Confounding

Image Source: Quality Matters

Confounding commonly seen by WWC reviewers ...

Exhibit 3: Examples of potential confounding factors

Example	Intervention	Potential confounding factor
In a study of a new reading curriculum, three new teachers volunteer to try the new program, while three teachers with 20 or more years of experience stick with the curriculum they've used for years.	New reading curriculum	Teacher experience (intervention group)
Intervention group students receiving a new math curriculum also receive additional tutoring that is not part of the curriculum. Comparison group students use the standard curriculum and receive no tutoring.	New math curriculum	Tutoring (intervention group)
All of the students in Mrs. Jones's and Mr. Wright's classes use a new software package to work on spelling (the intervention group), while Mrs. Smith's students continue to work only with pencils and paper (the comparison group).	New software package	Single teacher (comparison group)

Source: What Works Clearinghouse Standards Brief: Confounding

Baseline Equivalence

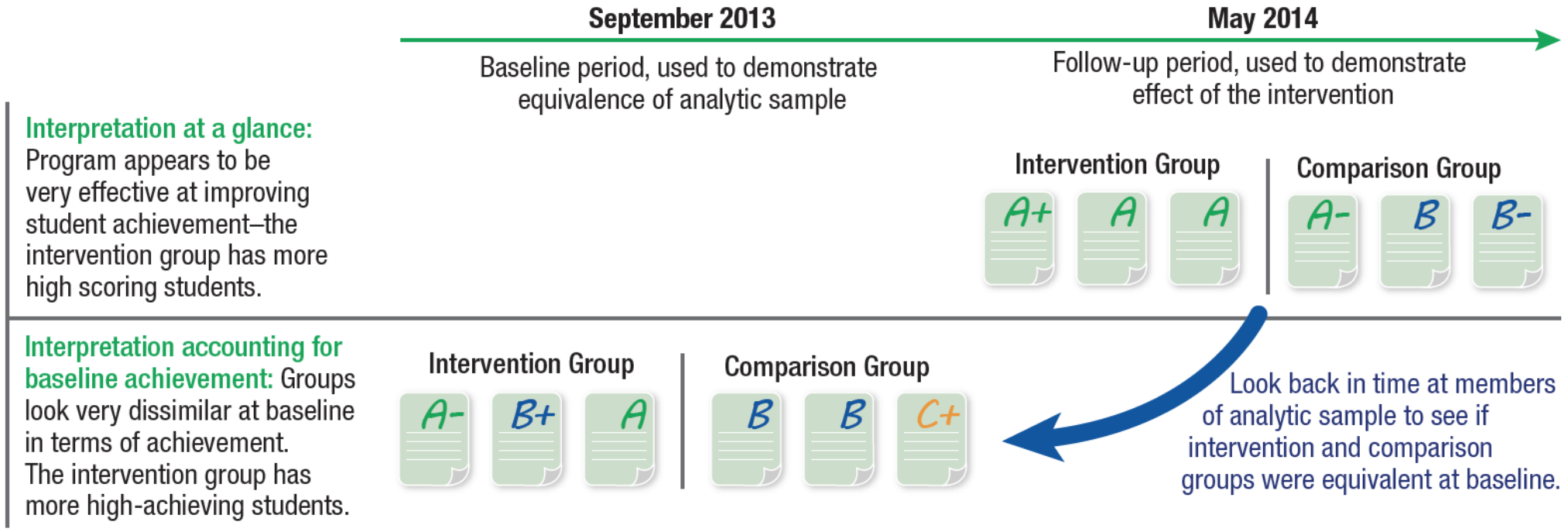
- When two groups are similar at the start of a study (baseline) and, after that, the only difference between the groups is that one receives the intervention and the other does not, it is reasonable to conclude that any differences in the outcomes that are measured at the end of a study (follow-up) are caused by the intervention.

- However, if the two groups are **different at baseline on key characteristics that could influence the outcomes**, the effect found at the end of the study might be due to the differences that already existed at the beginning.

- Demonstrating baseline equivalence is important in studies that did not assign participants randomly to the intervention and comparison groups.

- It is also important in **random assignment** studies with high **attrition**.

Figure 1: Why demonstrating baseline equivalence is important



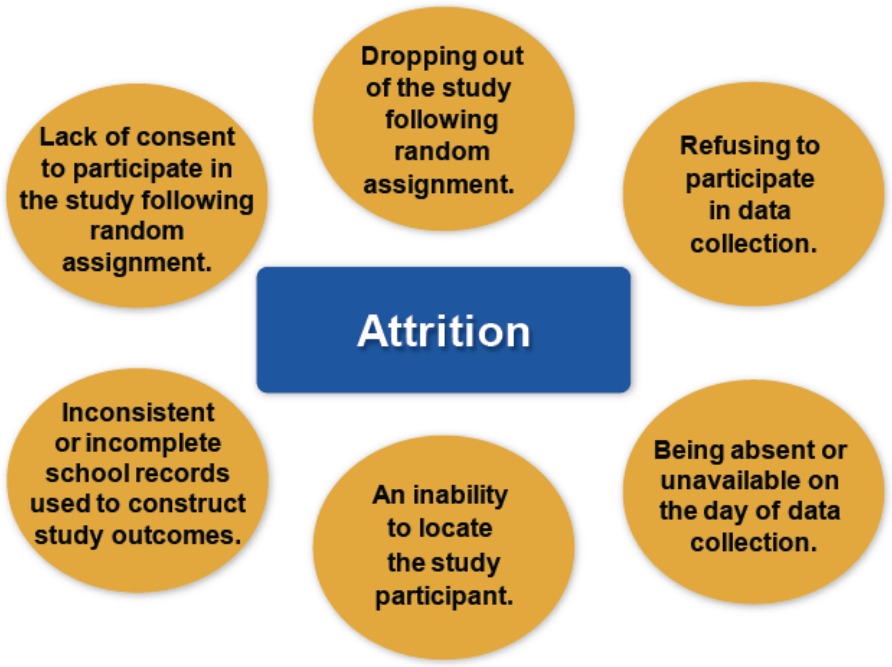
Conclusion: While the intervention group has higher achievement in May 2014, its members started off with higher achievement at baseline in September 2013. If the groups' baseline achievement levels are very different from each other (Figure 2), then the WWC will conclude that the groups are not equivalent and the study *Does Not Meet WWC Group Design Standards*.

Source: What Works Clearinghouse Standards Brief: Baseline Equivalence

Attrition: Losing participants over time

Sources of Attrition | Potential Consequences of Attrition | Potential Bias

Figure 1: Common causes of attrition

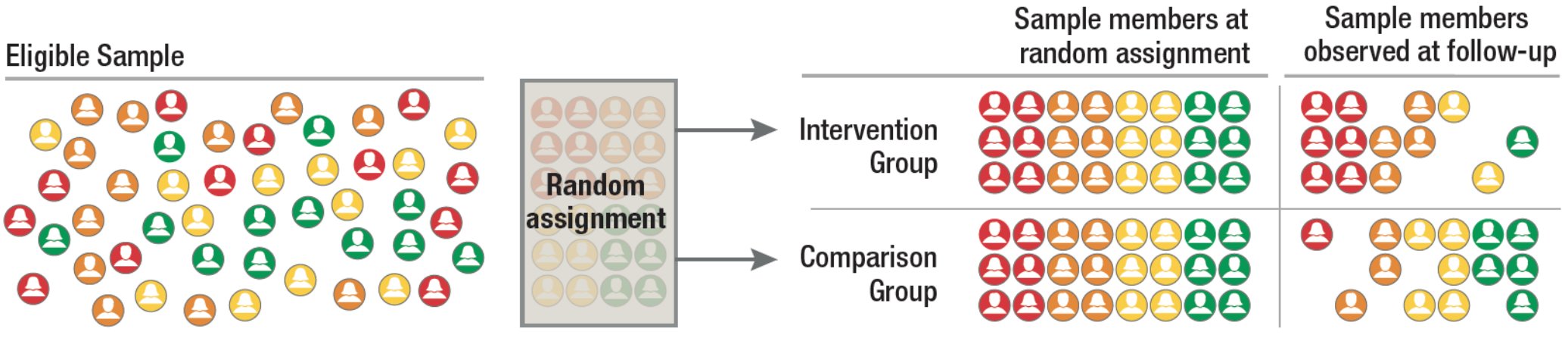


Source: What Works Clearinghouse Standards Brief: Attrition

Attrition: Losing participants over time

Sources of Attrition | Potential Consequences of Attrition | Potential Bias

Figure 2: Illustration of non-equivalence of baseline characteristics due to sample attrition



Source: What Works Clearinghouse Standards Brief: Attrition

Attrition: Losing participants over time

Bias .. the difference between

- the impact estimated using data from a sample experiencing attrition **versus ...**
- the true impact that would have been estimated had there been no attrition

Questions??