KIN 610: Quantitative Methods in Kinesiology

Chapter 5: Measures of Variability

Ovande Furtado Jr., PhD.

2026-02-03

1 FYI

This presentation is based on the the following books. The references are coming from these books unless otherwise specified.

Main sources:

- Weir, J. P., Vincent, W. J. (2021). Statistics in Kinesiology. Human Kinetics.
- Furtado, O., Jr. (2026). *Statistics for movement science: A hands-on guide with SPSS* (1st ed.). https://drfurtado.github.io/sms

ClassShare App

You may be asked in class to go to the ClassShare App to answer questions.

- https://classsharedrfurtado.netlify.app/

2 Learning Objectives

By the end of this chapter, you should be able to:

- Define variability and explain why it is important in movement science
- Calculate and interpret range, interquartile range, variance, standard deviation, and coefficient of variation
- Choose the appropriate measure of variability for different data types and distributions
- Understand the difference between between-person and within-person variability
- Interpret variability in the context of research findings

3 Symbols

Symbol	Name	Definition
s^2	Sample variance	Average squared deviation from the mean
s	Sample standard deviation	Square root of sample variance
σ^2	Population variance	Average squared deviation from the mean
σ	Population standard deviation	Square root of population variance
$\bar{x}$	Sample mean	Average of sample values
μ	Population mean	Average of population values
n	Sample size	Number of observations in the sample
N	Population size	Number of observations in the population
x_i	Individual observation	A single value in the dataset
Q_1	First quartile	25th percentile
Q_2	Second quartile	50th percentile (median)
Q_3	Third quartile	75th percentile
IQR	Interquartile range	$Q_3 - Q_1$
CV	Coefficient of variation	$(s/\bar{x}) \times 100$

4 Introduction: Describing the Spread

This chapter focuses on **measures of variability** — statistical values that describe how spread out or consistent values are in a data set[1].

- We will explore key measures:
  - **Range**: The distance between minimum and maximum values
  - **Interquartile Range (IQR)**: The spread of the middle 50%
  - **Variance and Standard Deviation**: Typical distance from the mean
  - **Coefficient of Variation (CV)**: Relative variability

! Important

Variability describes how tightly values cluster around the center and how much they differ from one another. In Movement Science, variability is rarely just "messiness" — it can reflect meaningful differences, adaptation, or measurement precision[2].

- Opening: Emphasize that variability answers the question "How consistent are the values?" in a data set.
- Example: Two training groups might both average 3.40 s sprint time, but one group is tightly clustered (3.30-3.50 s) while the other varies widely (2.80-4.00 s).
- Key distinction: Central tendency (where scores cluster) vs. variability (how spread out they are).
- Teaching tip: Ask students to think of scenarios where consistency matters (e.g., athlete reliability, measurement precision).

5 Key Terms

Understanding the terminology is essential for selecting and interpreting the appropriate measure[1]:

- **Variability/Spread**: How much values differ from one another and from the center
- **Range**: The distance between the maximum and minimum values
- **Quartiles**: Values that divide ordered data into four equal parts
- **Interquartile Range (IQR)**: The spread of the middle 50% of data (Q - Q)
- **Variance**: Average squared deviation from the mean
- **Standard Deviation (SD)**: Square root of variance; typical distance from the mean
- **Coefficient of Variation (CV)**: Relative variability expressed as a percentage

> **Note**
>
> Each measure has specific use cases, strengths, and limitations. The choice depends on the data type, distribution shape, and research question.

- Emphasize: These measures complement central tendency — together they provide a complete picture.
- Example: Range for quick sense of extremes, IQR for resistant measure, SD for inference.
- Quick check: "Which measure would you use for skewed data?" (IQR — resistant to outliers)

6 Why Variability Matters in Movement Science

Variability is information, not just noise:

- **Performance consistency**: Two athletes with the same mean sprint time but different variability show different reliability
  - *Example*: Athlete A averages 10.5 s with times ranging 10.4-10.6 s (consistent)
  - Athlete B also averages 10.5 s but ranges 10.0-11.0 s (inconsistent)
  - Who would you select for competition? Consistency matters!
- **Motor behavior**: Variability can reflect exploration, adaptation, fatigue, or loss of control[2]
  - *Healthy variability*: A beginner trying different throwing techniques (exploration)
  - *Adaptive variability*: Adjusting gait on uneven terrain (functional adaptation)
  - *Problematic variability*: Erratic movements in fatigued state (loss of control)
  - Context determines whether variability is beneficial or detrimental
- **Measurement quality**: Observed variability combines biological fluctuation and measurement error[3,4]
  - *Biological*: True day-to-day changes in performance (e.g., recovery status, motivation)
  - *Measurement error*: Imprecision in timing devices, scorer inconsistency
  - Challenge: Separating signal (real change) from noise (measurement error)

> Real example: same mean, different spread
>
> Two groups have identical mean sprint times (3.40 s), but Group A is tightly clustered (SD = 0.08 s) while Group B is scattered (SD = 0.20 s). If you only report the mean, you miss an important performance difference: reliability of execution.

- Key point: "More variability" is not automatically bad, and "less variability" is not automatically good — context matters.
- Example: During early skill acquisition, variability may be high as learners explore strategies; later it decreases as performance stabilizes.
- Teaching prompt: "What does high variability in a balance task under fatigue suggest?" (Control is challenged)

7 Levels of Variability

Variability is often nested in Movement Science studies:

7.1 Between-Person Variability

- How different people are from one another at a given time
- Example: At pre-testing, participants vary widely in peak force or VO max
- Reflects differences in fitness, body size, training history

7.2 Within-Person Variability

- How much the same person varies across trials, sessions, or days
- **Trial-to-trial**: Jump height varies within a single visit
- **Day-to-day**: Sprint time varies due to sleep, soreness, motivation

> **!** Important
>
> Always state which variability you mean (between-person or within-person), over what time scale (trials, sessions, days), and under what conditions (fatigue, learning, dual-task). SEE EXAMPLE IN THE NEXT SLIDE.

- Emphasize: The scientific meaning depends on which level you're studying and which task you're measuring.
- Example: "Trial-to-trial variability in a jump test might reflect warm-up effects, while day-to-day variability might reflect recovery status."
- Research example: "The Branscheidt study shows why we must be precise — fatigue effects on Day 1 changed learning patterns on Day 2, illustrating how conditions and time scales interact."

8 Research Example: Fatigue and Motor Learning

A study by Branscheidt et al.[5] exemplifies the importance of specifying variability type, time scale, and conditions:

Study: "Fatigue induces long-lasting detrimental changes in motor-skill learning"

Key specifications:

- **Variability types examined:**
  - *Within-person*: Performance changes across trials and days for each participant
  - *Between-person*: Comparing fatigued vs. non-fatigued groups
- **Time scales measured:**
  - *Trial-to-trial*: Performance within single practice sessions
  - *Session-to-session*: Learning assessed across multiple days (Day 1, Day 2, etc.)
  - *Long-term effects*: Fatigue on Day 1 impaired learning even on subsequent days without fatigue
- **Conditions specified:**
  - *Fatigue condition*: Participants trained to muscle fatigue (degradation of maximum force)

  - *Non-fatigue condition*: Control group without fatigue
  - *Different contexts*: Skill execution vs. skill acquisition examined separately

Finding: Muscle fatigue impaired motor-skill *learning* (not just execution) in ways that persisted across days, demonstrating that the condition under which practice occurs has long-lasting effects on skill acquisition.

PubMed Link | PMID: 30832766

9 A Quick Visual Start: Variability You Can See

Why start with graphs? Before computing any formula, always visualize your data.

Different plots highlight different aspects of variability. Let's explore each one:

- Key teaching point: "Visual inspection helps you detect outliers, identify skewness, and choose the right spread measure."
- Example: "A histogram might reveal right-skew that makes median/IQR more appropriate than mean/SD."

10 Dot Plots: See Every Data Point

What they show:

- Every individual data point
- Concentration of values (stacking)
- Extreme points and outliers
- Overall spread at a glance

Best for:

- Small to moderate datasets
- Showing exact values
- Comparing groups side-by-side

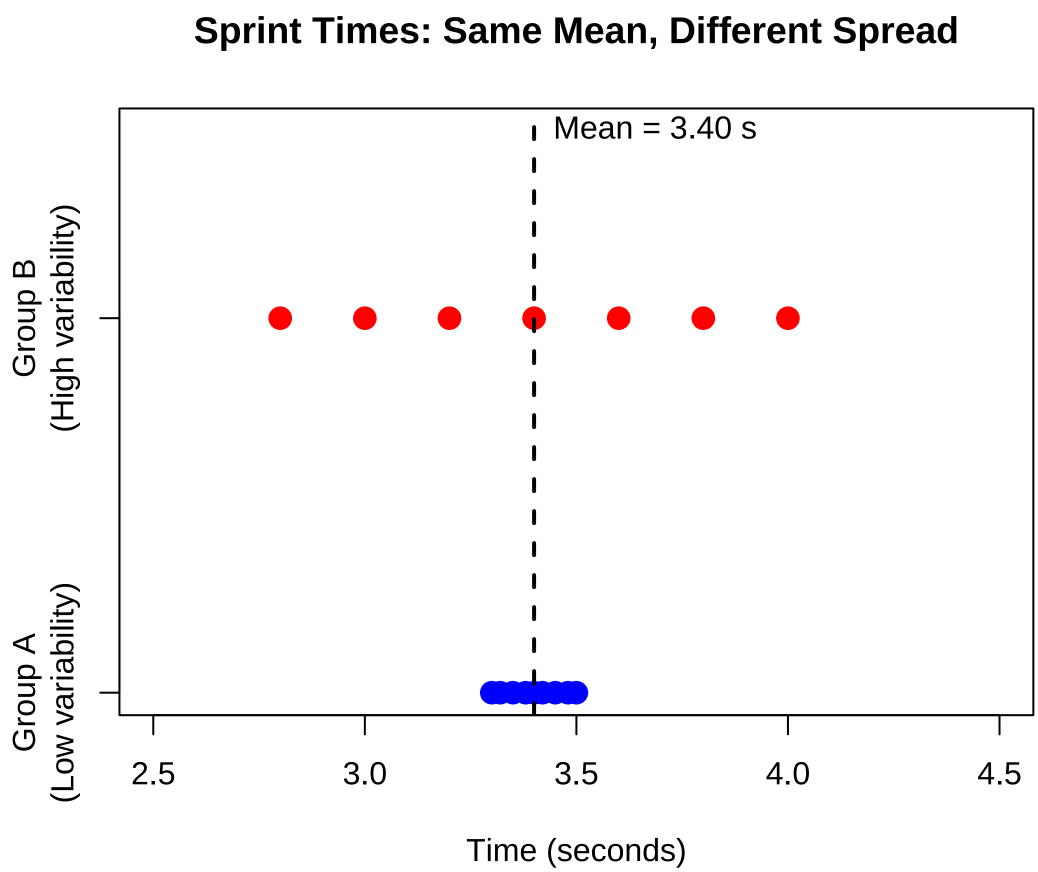


Figure 1: Dot plot showing sprint times

11 Box Plots: Focus on the Middle 50%

What they show:

- **Box**: Middle 50% of data (IQR)
- **Line in box**: Median
- **Whiskers**: Typical range
- **Dots beyond whiskers**: Outliers

Best for:

- Comparing multiple groups
- Identifying outliers
- Resistant to extreme values

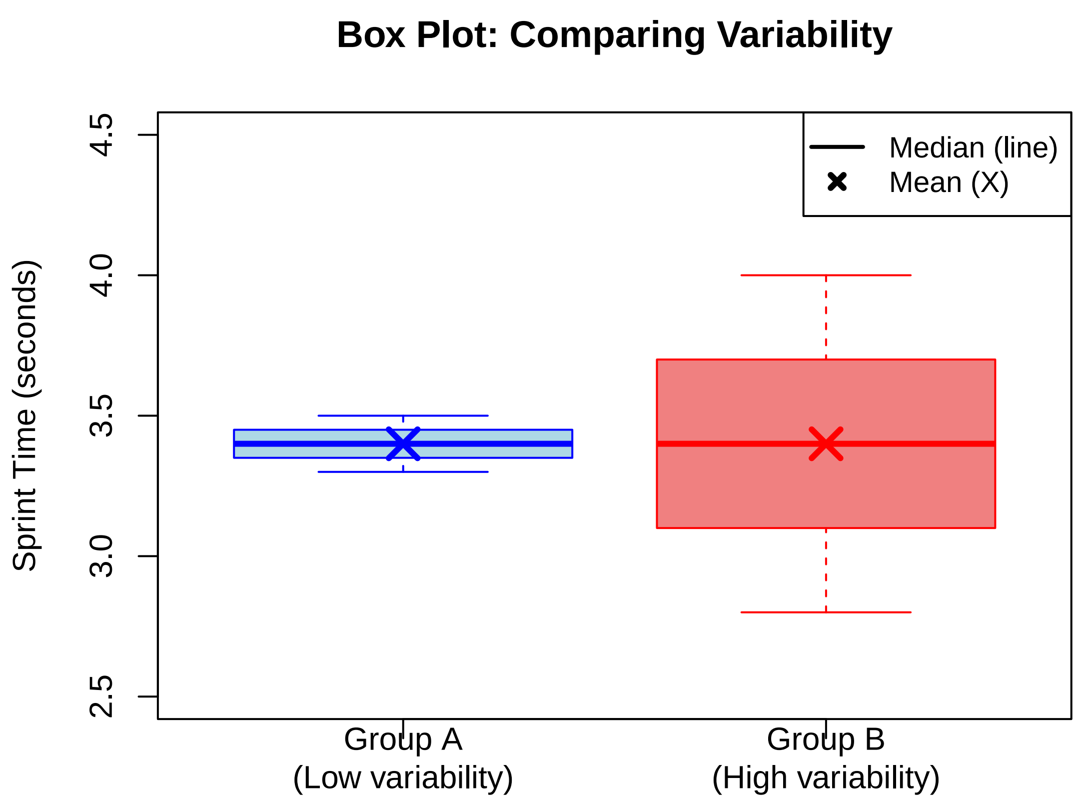


Figure 2: Box plot comparing two groups

> **Note**
>
> Notice how the box for Group B is much taller than Group A, showing greater spread in the middle 50% of the data.

12 Histograms: See Shape and Spread

What they show:

- **Shape** of the distribution
- **Spread** across the range
- **Skewness** (asymmetry)
- **Modality** (peaks)

Best for:

- Large datasets
- Assessing normality
- Identifying distribution shape

> **Note**
>
> Group A's histogram is tall and narrow (low variability), while Group B's is short and wide (high variability).

13 Error Bars: Showing Spread vs. Uncertainty

Two types of error bars:

1. SD bars (Standard Deviation) - Show **spread** in the sample - Describe variability of individuals - Longer bars = more variable data

2. SE/CI bars (Standard Error/Confidence Interval) - Show **uncertainty** in the mean - Describe precision of the estimate - Longer bars = less certain about mean

> **Warning**
>
> **Critical distinction**: Don't confuse SD bars (spread) with SE/CI bars (uncertainty)!

- Emphasize: "SD bars are much longer than SE bars because they describe different things."

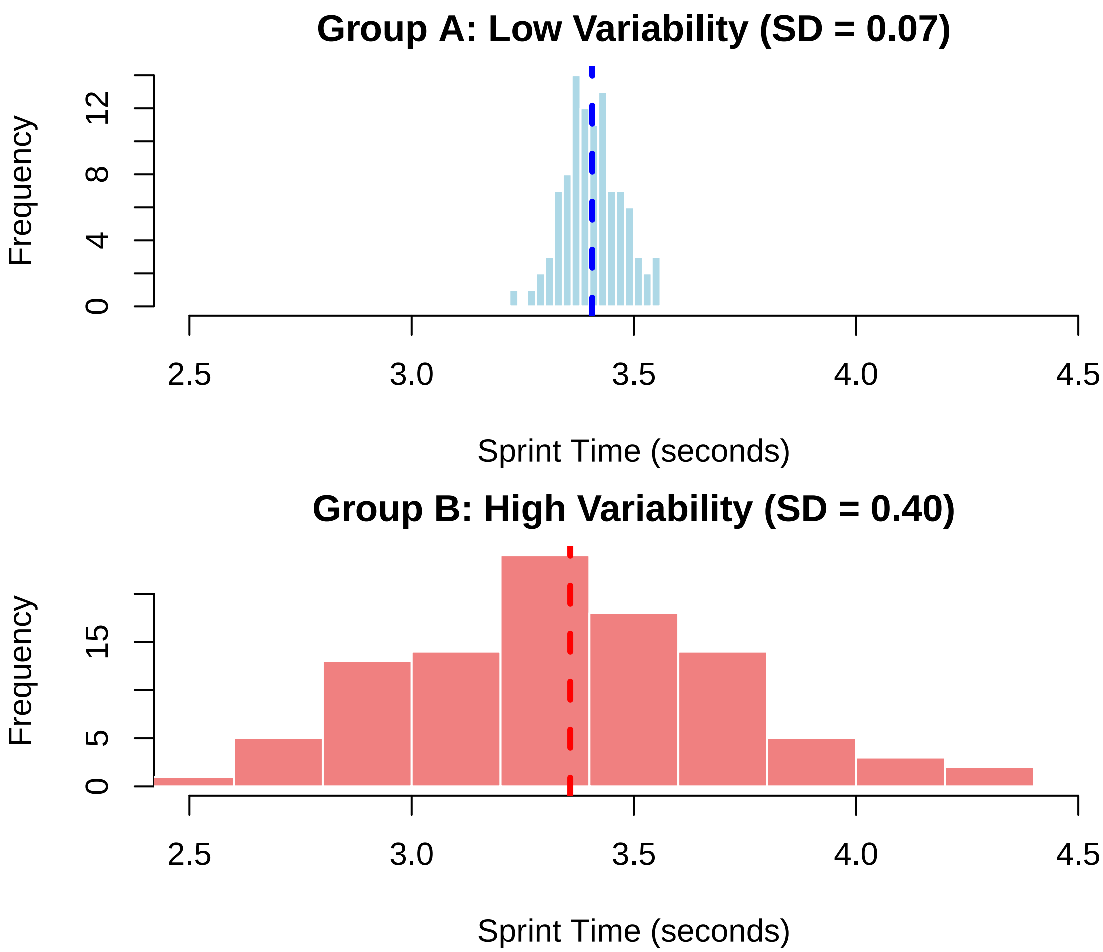


Figure 3: Histogram showing distribution shape

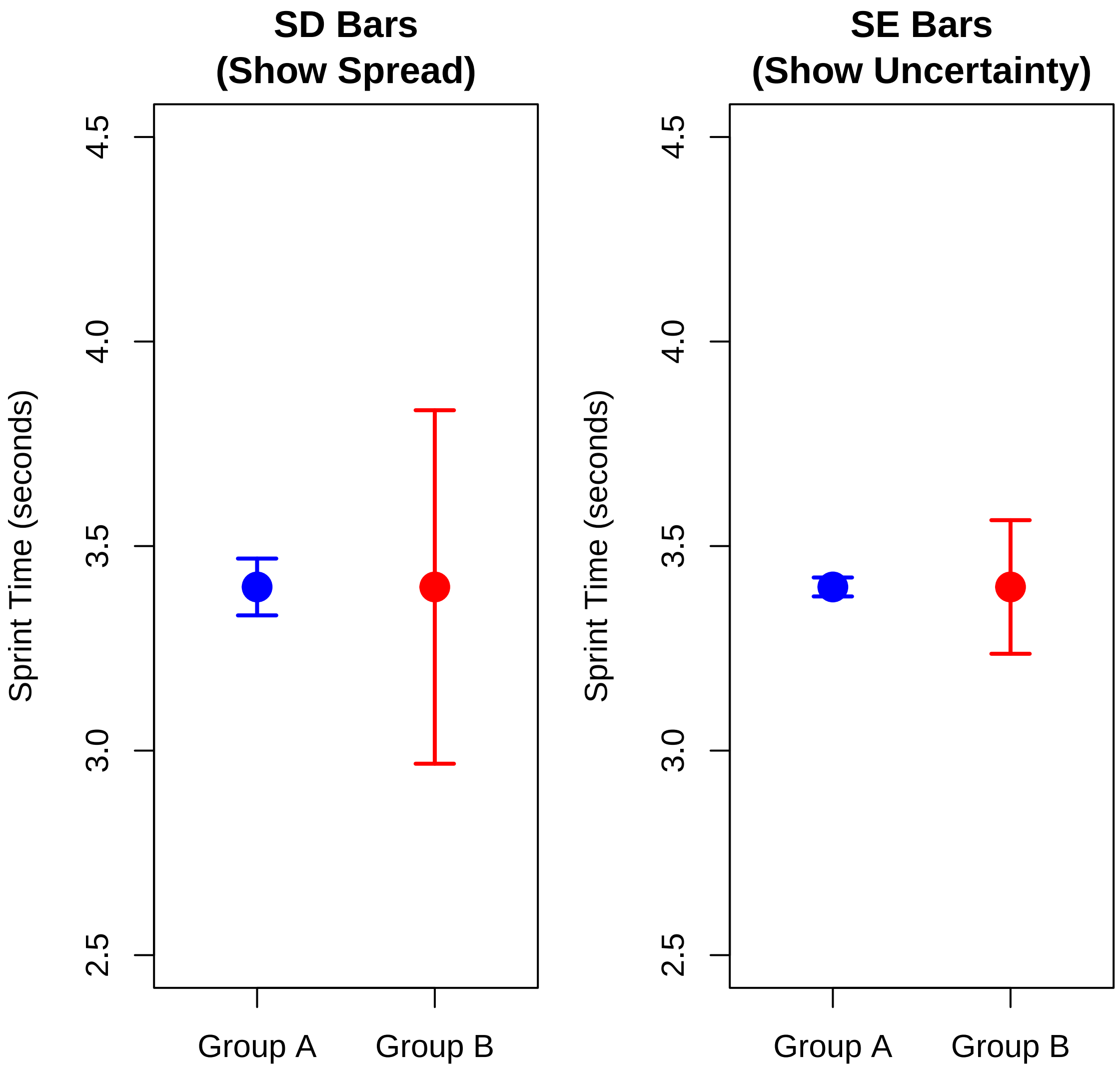


Figure 4: Error bars: SD vs SE

- Common mistake: Using SE bars to make data look less variable than it really is.

14 Visual Example: Same Mean, Different Spread

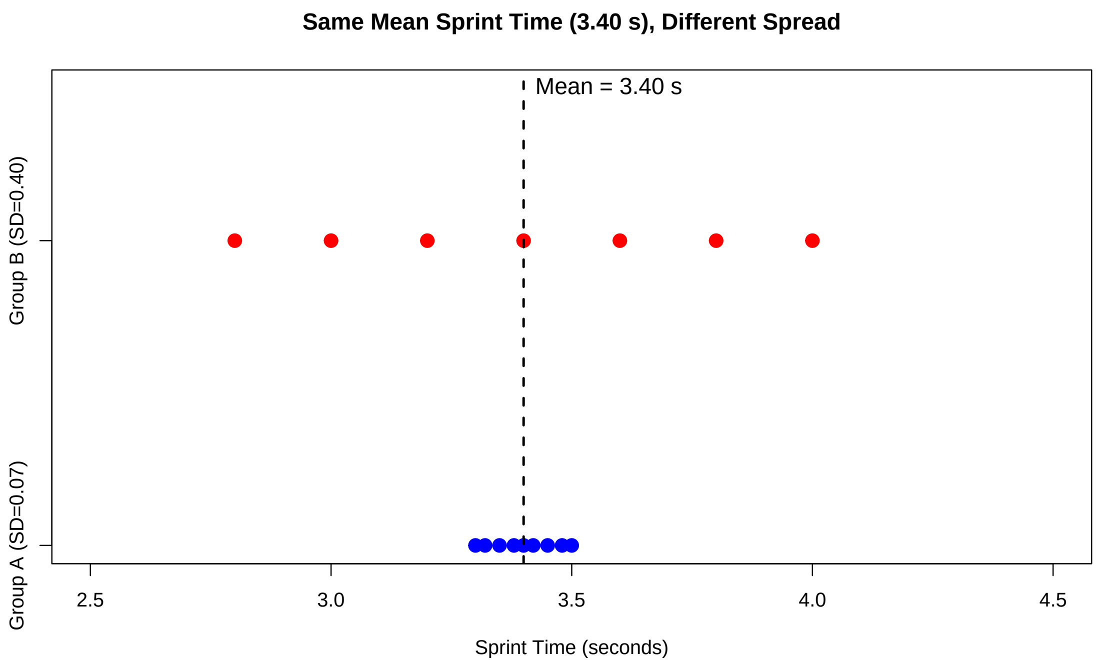
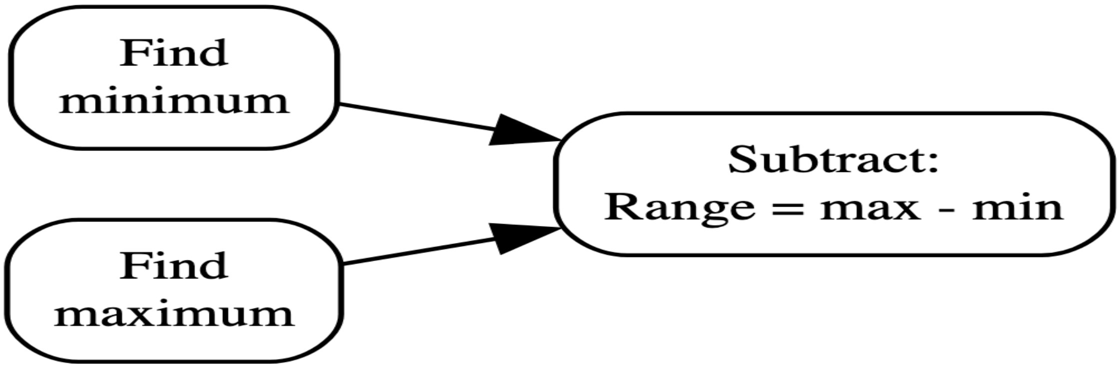


Figure 5: Dot plots showing two groups with identical means but different variability

- Walk through the plot: "Both groups have the same mean, but Group A shows high consistency while Group B shows poor reliability."
- Emphasize: "In Movement Science, this distinction matters for athlete selection, training effectiveness, and interpreting interventions."

15 Range: The Simplest Spread Measure



$$\text{Range} = x_{max} - x_{min}$$

- **Definition**: The distance between the smallest and largest values
- **Strengths**: Simple, quick sense of extremes
- **Limitations**:
  - Unstable — depends on only two observations
  - One unusual value can dominate it
  - Does not reflect typical spread

> ⚠ Warning
>
> The range often reflects outliers and extremes more than it reflects the typical participant[1].

- Example: "If nine sprint times fall between 3.30 and 3.50 s but one is 4.20 s (a slip or timing error), the range is driven by that single value."
- Teaching prompt: "When would the range be misleading?" (When outliers are present)

16 Range: Worked Example

Use this dataset of 10 observations (unsorted):

$$12,\ 9,\ 15,\ 8,\ 10,\ 14,\ 7,\ 11,\ 13,\ 9$$

Step 1: Identify minimum and maximum

- Minimum = 7

- Maximum = 15

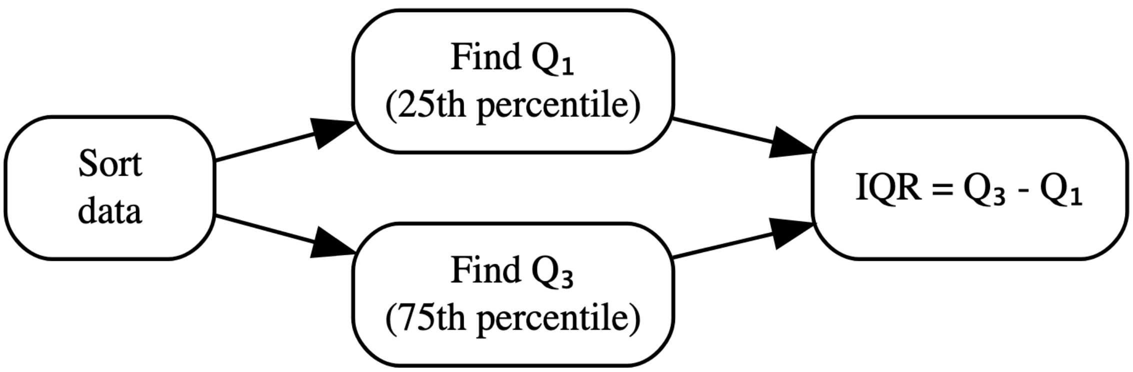
Step 2: Calculate range

$$\text{Range} = 15 - 7 = 8$$

Sensitivity example: If the maximum changed from 15 to 25 (an extreme value or error), the range would become 25 - 7 = 18. The range changed dramatically even though only one value changed.

- Walk through: "First find the smallest and largest values, then subtract."
- Emphasize sensitivity: "This shows why range is unstable — one outlier can change it completely."

17 Interquartile Range (IQR): A Resistant Measure



$$\text{IQR} = Q_3 - Q_1$$

- **Definition**: The spread of the middle 50% of the data
- **Characteristics**:
  - Based on ranks, not actual values
  - Resistant to extreme values
  - Pairs naturally with the median

> **Note**
>
> IQR is especially useful for skewed variables and ordinal scales[1]. It tells you how spread out the central bulk of your data is.

- Key point: "Unlike the range, IQR focuses on the typical spread rather than the extremes."
- Example: "If one athlete's jump height is measured incorrectly, the IQR won't change much, but the range will."

18 IQR: Worked Example

Use the same dataset, now sorted:

$$7,\ 8,\ 9,\ 9,\ 10,\ 11,\ 12,\ 13,\ 14,\ 15$$

Step 1: Split into lower and upper halves

- Lower half: 7, 8, 9, 9, 10
- Upper half: 11, 12, 13, 14, 15

Step 2: Find quartiles

- Q (median of lower half) = 9
- Q (median of upper half) = 13

Step 3: Calculate IQR

$$\text{IQR} = 13 - 9 = 4$$

Interpretation: The middle 50% of observations fall within a 4-unit span. Compared with the range (8), the IQR reflects typical spread rather than extremes.

> SPSS
>
> Use the "Frequencies" (`Analyze` > `Descriptive Statistics` > `Frequencies`) procedure and click the "Statistics" button. Check the box for "Quartiles" under "Percentile Values" to get the 25th and 75th percentiles. The IQR is then calculated as the difference between these two values: IQR = Q3 (75th percentile) − Q1 (25th percentile).
> For more details, refer to: [6].

- Walk through: "Sort the data, split into halves, find the median of each half, then subtract."
- Emphasize: "This gives us a robust measure that isn't affected by that extreme value at 15."

19 Box Plot: Visualizing IQR and Outliers

Figure 6: Box plot showing the interquartile range and outliers

Anatomy of a Box Plot

A box plot (or box-and-whisker plot) provides a visual summary of the **five-number summary**:

1. **The Box**: Spans from Q_1 (25th percentile) to Q_3 (75th percentile). The height or length of the box represents the **IQR**.
2. **The Median**: The line inside the box marks the 50th percentile.
3. **Outlier Detection (1.5 × IQR Rule)**:
   - **Lower Fence**: $Q_1 - (1.5 \times \text{IQR})$

- **Upper Fence**: $Q_3 + (1.5 \times \text{IQR})$
- **Outliers**: Any data points falling **outside** these fences (as shown by the red label in the graph above).

4. **Whiskers**: Extend from the box to the most extreme data points that are **not** outliers.

- Key teaching point: "Box plots help quickly assess data symmetry, central tendency, and potential anomalies."
- Example: "In Movement Science, outliers might indicate measurement issues or exceptional cases."

20 Variance and Standard Deviation: Why Use Deviations?

For many methods, we want to describe how far values tend to fall from the mean.

The problem with simple deviations: $(x_i - \bar{x})$

- Deviations from the mean always sum to zero
- Example: Mean = 10, observations = 8, 10, 12
- Deviations: -2, 0, +2
- Sum: -2 + 0 + 2 = 0 (even though values clearly vary!)

The solution: Square the deviations

- Makes every value nonnegative
- Gives larger departures more weight
- Leads to variance (average squared deviation) and standard deviation (square root of variance)

- Key point: "We can't just average deviations because they cancel out. Squaring them solves this problem."
- Teaching prompt: "Why square instead of using absolute values?" (Squaring has better mathematical properties for inference)

21 Sample Variance

Sample variance is the average squared deviation from the mean:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

Components:

- s^2 = Sample variance
- $\sum(x_i - \bar{x})^2$ = Sum of squared deviations
- $n - 1$ = Degrees of freedom

❗ Why n-1 instead of n?

The key difference: Sample vs. Population

- **Population variance** (σ^2): When you have **all** the data, divide by N (the total population size)
  - Formula: $\sigma^2 = \frac{\sum(x_i-\mu)^2}{N}$
  - You're describing the **actual** variability in your complete population
- **Sample variance** (s^2): When you have only a **sample**, divide by $n - 1$ (not n)
  - Formula: $s^2 = \frac{\sum(x_i-\bar{x})^2}{n-1}$
  - You're **estimating** the population variability from incomplete data

Why the correction?
When we calculate the sample mean ($\bar{x}$), we're using the same data to estimate both the center AND the spread. This creates a problem: our sample mean is always closer to our sample data than the true population mean would be.
Example: - Imagine the true population mean sprint time is 4.0 seconds - Your sample of 10 athletes has a mean of 3.8 seconds (by chance, a bit faster) - If you calculate deviations from 3.8 instead of 4.0, they'll be artificially smaller - Using n would **underestimate** the true population variance - Using $n - 1$ corrects for this bias and gives a better estimate
Bottom line: Use $n-1$ when calculating variance from a **sample** to estimate **population** variability. This makes your estimate unbiased.
Units problem: Variance has squared units (e.g., seconds²), which makes it hard to interpret directly. This is why we typically report standard deviation instead.

- Emphasize: "Variance is mathematically important, but hard to interpret because of squared units."
- Example: "If measuring sprint time in seconds, variance is in seconds-squared — what does that mean?"

22 Standard Deviation

Standard deviation is the square root of variance:

$$s = \sqrt{s^2}$$

- **Definition**: A typical distance from the mean, in the **same units** as the original variable
- **Interpretation**: Most interpretable when the distribution is roughly symmetric and unimodal
- **Caution**: Under strong skew or outliers, SD can be inflated in ways that do not match the typical participant

> **A useful interpretation sentence**
>
> **Think of SD as the "average spread," not a fence.**
> The standard deviation tells you how far values **typically** are from the mean—it's not a strict boundary. Some values will be closer, some farther.
> **Example:** If sprint times have a mean of 10 s and SD of 1 s, most athletes will be within about 1 second of 10 s (so between 9-11 s), but some might be at 8 s or 12 s. The SD just describes the typical spread.
> **Key point:** Always look at your actual data distribution, not just the SD number[1].

> **Good to know**
>
> How to get the variance from the standard deviation?
> Click to expand
> The answer is simple: square the standard deviation.
>
> $$s^2 = s^2$$

- Key point: "SD returns us to the original units, making it much easier to interpret than variance."
- Example: "If mean sprint time is 3.40 s with SD = 0.20 s, the typical distance from the mean is 0.20 s."

23 Variance and SD: Worked Example

Use the same dataset:

$$12,\ 9,\ 15,\ 8,\ 10,\ 14,\ 7,\ 11,\ 13,\ 9$$

Step 1: Compute the mean

$$\bar{x} = \frac{108}{10} = 10.8$$

Step 2: Compute squared deviations and sum them

Each squared deviation is $(x_i - 10.8)^2$. Summing all squared deviations yields:

$$\sum (x_i - \bar{x})^2 = 63.6$$

- Walk through: "First calculate the mean, then find how far each value is from the mean, square those distances, and add them up."

24 Variance and SD: Worked Example (continued)

Step 3: Compute sample variance

$$s^2 = \frac{63.6}{10 - 1} = \frac{63.6}{9} = 7.07$$

Step 4: Compute sample standard deviation

$$s = \sqrt{7.07} = 2.66$$

Interpretation: If these values were a performance outcome (e.g., seconds), the typical distance from the mean is about 2.66 units. Whether that is "large" depends on the context and the measurement scale.

> ⚠ Common mistake
>
> Interpreting SD as "most values are within one SD" without checking the distribution[1]. That rule-of-thumb only behaves well under roughly symmetric, bell-shaped distributions.

> SPSS
>
> To calculate the variance and standard deviation in SPSS, you can use the `Descriptives` procedure.
>
> 1. Go to `Analyze` > `Descriptive Statistics` > `Descriptives`.
> 2. Move your variable into the `Variables(s)` box.
> 3. Click `Options`.
> 4. Check `Variance` and `Standard Deviation`.

5. Click `Continue`, then `OK`.

- Emphasize: "The interpretation of SD depends on context — 2.66 might be large for one measure and small for another."
- Teaching prompt: "How would you know if an SD is 'large'?" (Compare to the mean, look at the distribution, check literature values)

25 Interpreting Variability in Motor Performance

Variability does not have a single meaning across tasks:

- **Reduced variability** can reflect:
  - Consistency and skill (good)
  - Rigidity (potentially problematic)
- **Increased variability** can reflect:
  - Loss of control (problematic)
  - Exploration during learning (adaptive)

Examples:

- **Balance task under fatigue**: Sway variability increases (control is challenged)
- **Early skill acquisition**: Trial-to-trial variability high (exploring strategies), then decreases (performance stabilizes)

> **Optimal Movement Variability Framework**
>
> Healthy, skilled movement needs the right amount of variability—like a "Goldilocks zone" (not too much, not too little, just right):
>
> - **Too little variability** = Rigid, can't adapt (think: walking on ice, very stiff and careful)
> - **Too much variability** = Unstable, unpredictable (think: loss of control when fatigued)
> - **Just right** = Flexible enough to adjust to changes, stable enough to be reliable
>
> **Example:** A skilled basketball player varies their shot slightly based on defender position (adaptive), but not so much that accuracy suffers (stable)[2,7].

- Key teaching point: "Context determines whether variability is good or bad."
- Example: "In a strength training study, day-to-day variability may remain while the mean improves."

26 Coefficient of Variation (CV)

The coefficient of variation expresses variability relative to the mean:

$$\text{CV} = \frac{s}{\bar{x}} \times 100\%$$

- **Purpose**: Compare variability across outcomes with different units or magnitudes
- **Interpretation**: Expresses SD as a percentage of the mean
- **Use cases**: Especially useful when measurement error scales with magnitude[3]
    - **What this means**: In some measurements, larger values naturally have more variability
    - **Example 1 (Strength)**: A person who lifts 200 kg might vary by ±10 kg, while someone lifting 50 kg might vary by ±2 kg. The absolute variability (10 vs. 2) looks different, but the CV shows they're both about 5% variable, making them comparable
    - **Example 2 (FitnessGram)**: Comparing test-retest reliability across different tests—PACER has SD = 8 laps (mean = 80 laps, CV = 10%), while sit-and-reach has SD = 2 cm (mean = 20 cm, CV = 10%). Even though the units differ (laps vs. cm), CV reveals both tests have similar relative consistency

> ⚠ Cautions
>
> - CV behaves poorly when the mean is near zero
> - Not appropriate for ordinal scales or bounded scales where "zero" is not a true absence
> - In those settings, an absolute spread measure (IQR or SD) is often more interpretable

- Example: "Comparing variability in sprint time (seconds) vs. jump height (cm) — CV allows meaningful comparison."
- Teaching prompt: "When would CV be inappropriate?" (Pain ratings, temperature in Celsius, any scale without a true zero)

27 CV: Worked Example

From the earlier worked example:

- $\bar{x} = 10.8$
- $s = 2.66$

$$\text{CV} = \frac{2.66}{10.8} \times 100\% = 24.6\%$$

Interpretation: A CV of 24.6% means the standard deviation represents about one quarter of the mean value.

Movement Science context:

- Highly reliable measurements (force platforms): CV $< 5\%$
- Field-based performance tests (sprint, jump): CV = 3-8%
- More variable outcomes (EMG, subjective ratings): CV = 10-30%
- Values $> 30\%$: Often indicate measurement issues or highly heterogeneous samples

Whether 24.6% is "acceptable" depends on the measurement context and typical error ranges in the literature[3].

- Emphasize: "CV provides context — is this variability typical for this type of measurement?"
- Example: "For EMG, 24.6% might be reasonable; for precision timing, it could signal problems."

28 Test Your Knowledge: Calculating Variability

You measure vertical jump height (in cm) for 5 athletes:

Dataset: 45, 50, 55, 60, 40

Calculate:

1. Range
2. Mean
3. Standard deviation (use calculator: s 7.91 cm)
4. Coefficient of variation

Answers

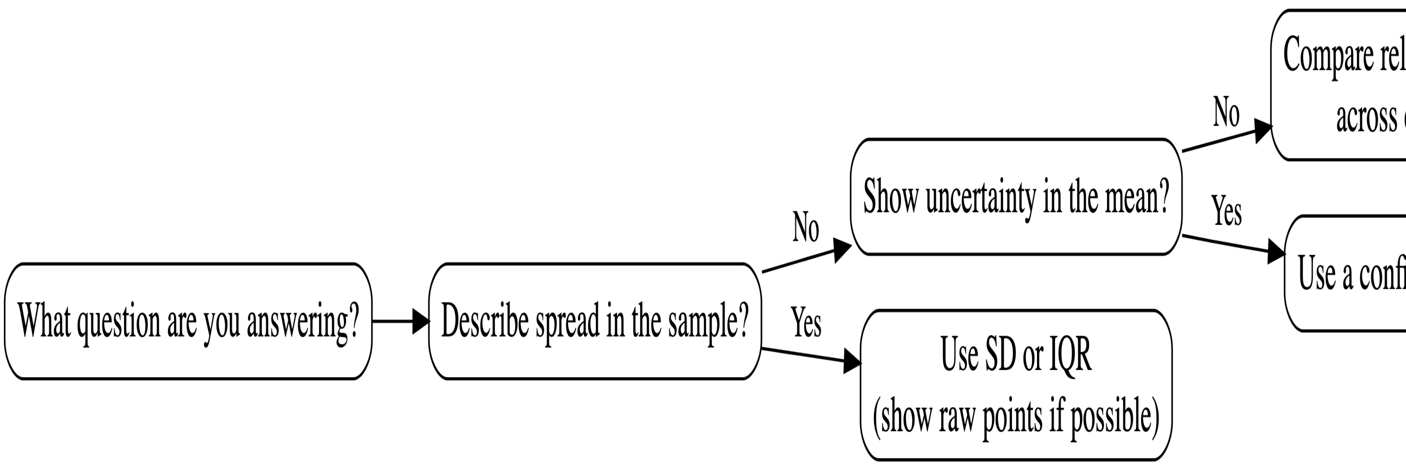
1. **Range**: 60 - 40 = 20 cm
2. **Mean**: (45 + 50 + 55 + 60 + 40) / 5 = 250 / 5 = 50 cm
3. **Standard deviation**: s 7.91 cm
4. **CV**: (7.91 / 50) × 100% = 15.8%

SPSS

In SPSS, you can calculate the range, mean, standard deviation, and CV using the Descriptive Statistics procedure (Analyze > Descriptive Statistics > Descriptives).

29 Graphical Depiction of Variability

Graphs often convey variability more clearly than a single number:



❗ Important

If the goal is to describe performance consistency, SD or IQR is usually appropriate. If the goal is to show the precision of a mean estimate, confidence intervals are usually more meaningful.

30 Effective Visualization Options

- **Boxplots**: Show the IQR directly
  - Tall box = middle 50% spread out
  - Short box = tightly clustered
- **Dot plots**: Display every observation

  - Excellent for small/moderate samples
  - Clearly show clusters and outliers
- **Histograms**: Illustrate distribution shape and spread
  - Important because variability measures behave differently under skew
- **Line plots across time**: Reveal whether variability changes during training
  - Can add individual trajectories or plot spread at each time point
- **Error bars**: Require careful choice of type
  - SD bars: Describe spread among individuals
  - SE/CI bars: Describe uncertainty in the mean estimate
- Key point: "Choose the visualization that best answers your research question."
- Example: "For showing individual consistency, dot plots or line plots work well. For comparing groups, box plots are efficient."

31 Variability as a Data Quality Check

Variability estimates serve as diagnostic tools for data quality:

When variability is unusually large:

- Real heterogeneity (distinct subgroups mixed together)
- Measurement inconsistency (protocol variations, different testers)
- Device or calibration issues
- Data entry errors

When variability is unusually small:

- Constrained measurement range (insufficient resolution)
- Ceiling or floor effects (values pile up at boundaries)
- Rounding or truncation (e.g., recording 10.2, 10.4, 10.7 all as "10" reduces apparent variability)
- Overly homogeneous sample (e.g., studying only elite athletes aged 20-22 may show less variability than the general population)

> ❗ Important
>
> Variability should be evaluated alongside the research context, expected ranges for similar measurements, and visual inspection of distributions.

- Key teaching point: "Extreme variability in either direction should prompt investigation."

- Example: "If SD is suspiciously small, check for rounding or ceiling effects. If suspiciously large, check for outliers or mixed subgroups."

32 Variability Audit Workflow

When you calculate variability measures, always ask: **"Does this make sense?"**

Step 1: Plot the distribution

- Create a histogram, dot plot, or box plot
- Visualize the spread before trusting numbers

Step 2: Assess plausibility

- Is the spread reasonable for this measurement?
- Compare to expected ranges or literature values

Step 3: If spread seems unusual:

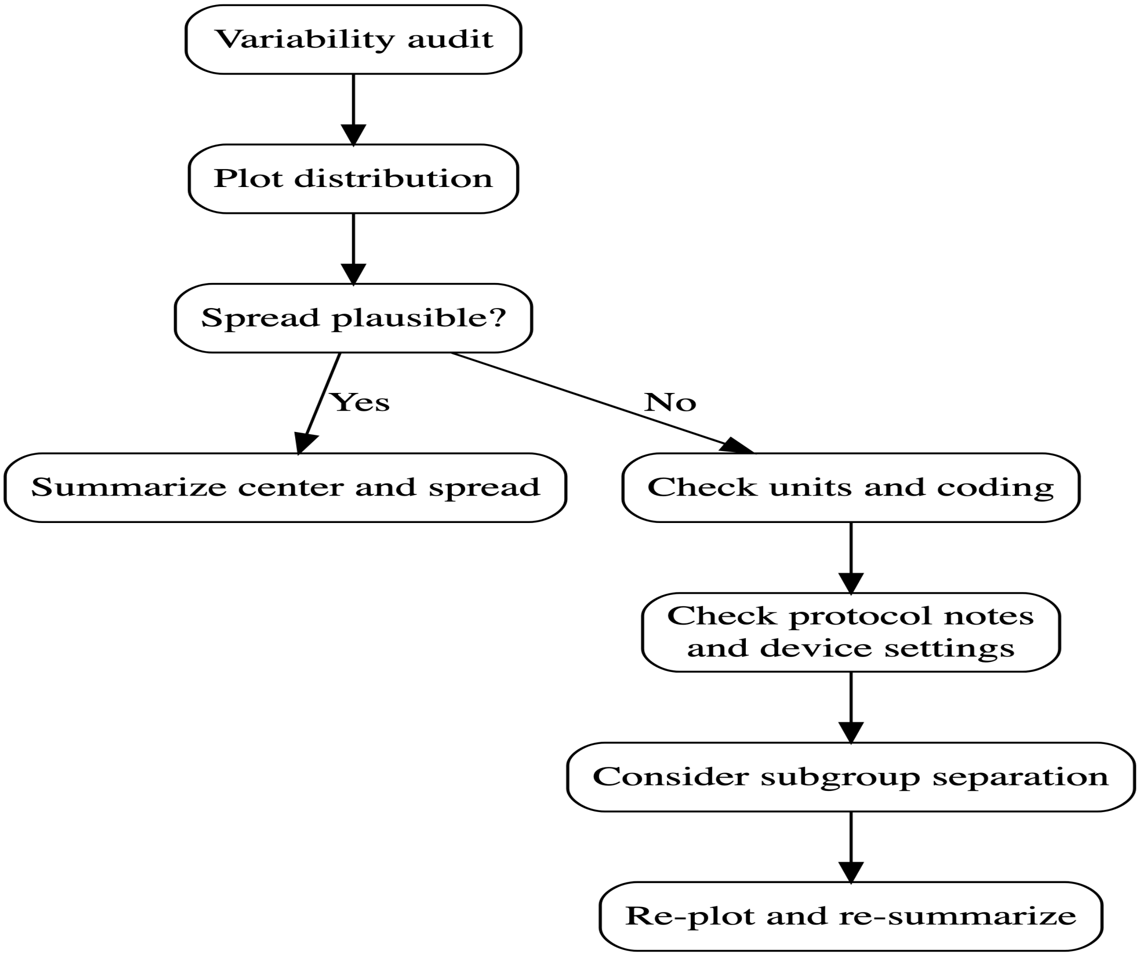
- **Too large?** Check for outliers, data entry errors, mixed subgroups
- **Too small?** Check for rounding, ceiling/floor effects, restricted range

Step 4: Investigate and fix

- Review data collection protocols
- Check device settings and calibration
- Consider if subgroups need separate analysis

Step 5: Re-analyze

- After corrections, re-plot and recalculate
- Verify the variability now makes sense



- Walk through the workflow: "Always start by plotting, then assess plausibility, then investigate if something looks wrong."
- Teaching prompt: "What would you do if you found an SD of 0.01 for sprint times?" (Check for rounding, data entry errors, or ceiling effects)

33 Choosing the Right Variability Measure

For common Movement Science variables:

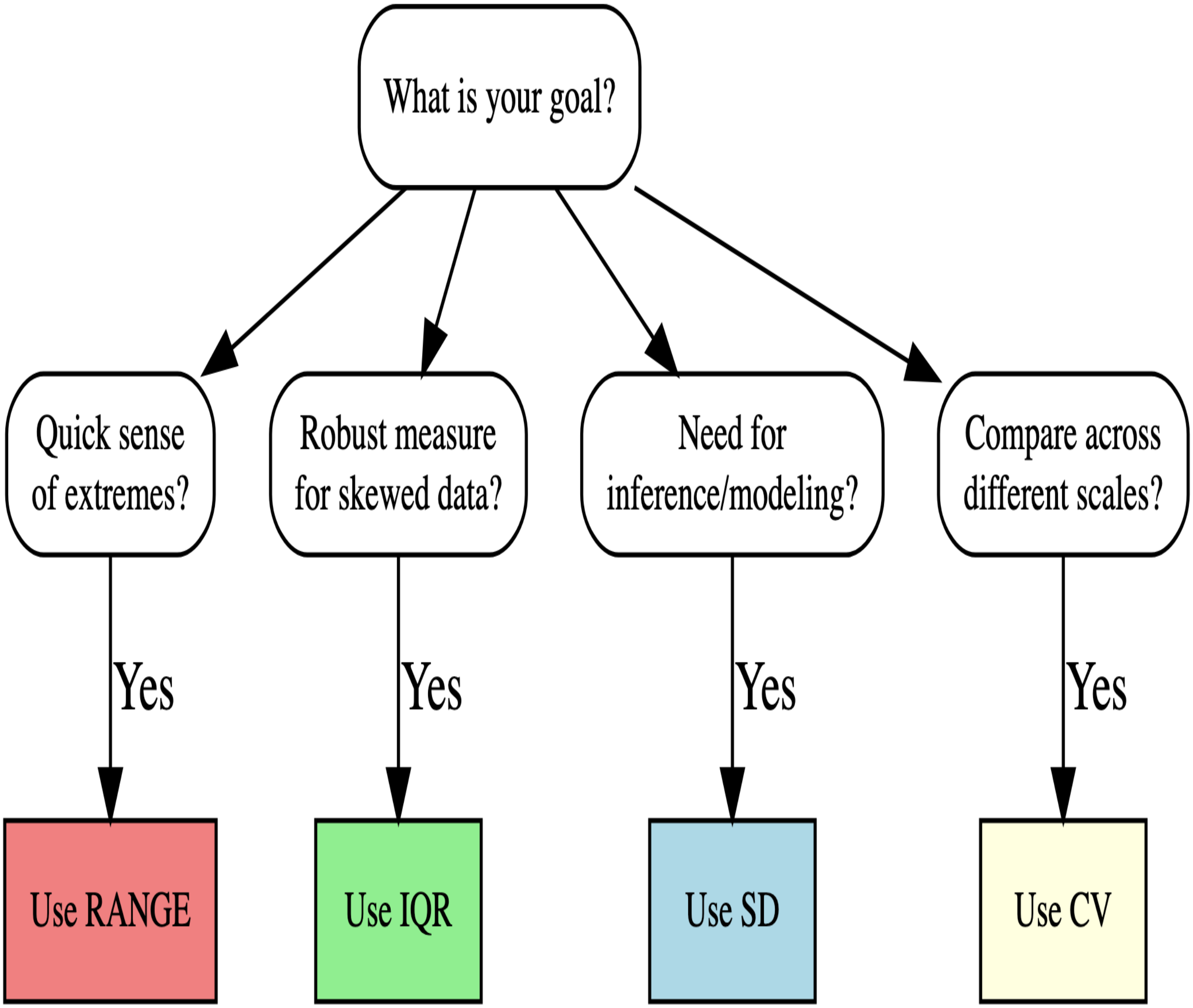
- **Sprint time**: If roughly symmetric, mean and SD are reasonable. Confirm with a dot plot.
- **Sway area or EMG amplitude**: If right-skewed, median and IQR often describe spread more faithfully. Consider log transform if modeling.

- **Pain rating or RPE**: For ordinal scales, report median with IQR. SD and CV can be hard to interpret.
- **Function score**: Watch for ceiling effects. Use median, IQR, and a plot. Be cautious with small SD values that may reflect a ceiling.

> **Note**
>
> Match the spread measure to the outcome, distribution shape, and research question.

- Emphasize: "The choice of measure depends on your data characteristics and goals."
- Example: "For skewed EMG data, IQR is more robust than SD."

34 Decision Tree: Selecting the Right Measure



> **! Important**
>
> This decision tree provides a practical guide, but always consider the research question, data distribution, and context when selecting a measure of variability.

35 Test Your Knowledge: Choosing the Right Measure

For each scenario, identify which measure of variability is most appropriate:

1. **Reaction time** (in ms) with one extreme outlier — which spread measure?
2. **EMG amplitude** with right-skewed distribution — which spread measure?
3. **Sprint times** from two different age groups (adults vs. children) — how to compare variability?
4. **Jump height** for normally distributed sample, need to calculate z-scores later — which spread measure?

Using the ClassShare App, submit your answers.

Answers

1. **IQR** — resistant to the outlier
2. **IQR** — robust for skewed data
3. **CV** — allows comparison across different magnitudes
4. **SD** — needed for z-score calculations and inference

36 Practical Application: Real Data Example

Scenario: You measure countermovement jump height (in cm) for 10 athletes:

Table 2: Vertical Jump Data for 10 Athletes

Athlete	Jump Height (cm)
1	45
2	48
3	46
4	50
5	45
6	52
7	47
8	45
9	51
10	65

Questions:

1. What is the range?
2. What is the IQR?

3. What is the mean and SD?
4. Which measure best represents typical variability? Why?

- Walk through calculations together.
- Range: 65 - 45 = 20 cm
- Ordered: 45, 45, 45, 46, 47, 48, 50, 51, 52, 65
- Q1 = 45, Q3 = 51, IQR = 6
- Mean 49.4, SD 5.9 (without outlier: SD 2.8)
- Discussion: The IQR (6 cm) best represents typical variability because the SD is inflated by the outlier (65 cm).

37 Practical Application: Answers

Click to reveal answers

1. **Range**: 65 - 45 = 20 cm

2. **IQR**:

   - Ordered data: 45, 45, 45, 46, 47, 48, 50, 51, 52, 65
   - Q = 45, Q = 51
   - IQR = 51 - 45 = 6 cm

Click to reveal answers (continued)

3. **Mean and SD**:

   - Sum = 494
   - Mean = 494 / 10 = 49.4 cm
   - SD 5.9 cm (inflated by outlier)
   - Without outlier (65): SD 2.8 cm

4. **Best measure**: The **IQR (6 cm)** best represents typical variability because the SD is inflated by the outlier at 65 cm. The IQR focuses on the middle 50% and is resistant to this extreme value.

> **Note**
>
> This example demonstrates why understanding distribution shape and outliers is crucial for selecting the appropriate measure of variability.

38 Key Takeaways

> ❗ Remember These Core Concepts:
>
> 1. **Range**: Simple but unstable — dominated by extremes
> 2. **IQR**: Resistant to outliers — best for skewed or ordinal data; describes middle 50%
> 3. **Variance**: Average squared deviation — mathematically important but hard to interpret (squared units)
> 4. **Standard deviation**: Typical distance from the mean — most interpretable for symmetric distributions
> 5. **Coefficient of variation**: Relative variability — useful for comparing across different scales
> 6. **Context matters**: In Movement Science, variability can reflect performance structure, adaptation, or measurement precision
> 7. **Always visualize**: Graphs reveal patterns that numbers alone cannot show

- Closing: "Measures of variability complete the picture started by measures of central tendency. Together, they describe your data distribution."
- Final prompt: "What questions do you have about when to use each measure?"

39 Chapter Summary

Measures of variability describe how spread out values are and how consistent performance appears:

- **Range** is simple but unstable and dominated by extremes
- **IQR** focuses on the middle 50% and is resistant to outliers
- **Variance and SD** quantify typical deviation from the mean and are widely used in mean-based methods
- **CV** expresses relative spread and is useful when variability scales with magnitude

In Movement Science, variability can reflect both performance structure and measurement precision, so it should be interpreted in context and supported by graphs[2,3].

> ℹ Workflow for summarizing spread
>
> 1. Identify variable type and measurement scale
> 2. Visualize the distribution
> 3. Choose a spread measure that matches the distribution and purpose
> 4. Pair spread with an appropriate measure of center
> 5. Write a one-sentence justification for your choice

40 Next Steps: Chapter 6

In the next chapter, we will explore **the normal distribution and standard scores**:

- What is the normal curve and why does it matter?
- How do we use z-scores to standardize values?
- What is the relationship between variability and probability?

> **Note**
>
> Understanding variability is essential for interpreting standard scores, which express how far a value falls from the mean in standard deviation units.

- Preview: "Variability (SD) is the foundation for z-scores, which tell us how unusual a value is relative to the distribution."
- Example: "If an athlete's VO max is 2 SD above the mean, we can use the normal curve to estimate how rare that is."

References

1. Moore, D. S., McCabe, G. P., & Craig, B. A. (2021). *Introduction to the practice of statistics* (10th ed.). W. H. Freeman; Company.
2. Stergiou, N., & Decker, L. M. (2011). Human movement variability, nonlinear dynamics, and pathology: Is there a connection? *Human Movement Science*, *30*, 869–888. https://doi.org/10.1016/j.humov.2011.06.002
3. Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine*, *30*(1), 1–15. https://doi.org/10.2165/00007256-200030010-00001
4. Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, *26*(4), 217–238. https://doi.org/10.2165/00007256-199826040-00002
5. Branscheidt, M., Kassavetis, P., Anaya, M., Rogers, D., Huang, H., Lindquist, M. A., & Celnik, P. (2019). Fatigue induces long-lasting detrimental changes in motor-skill learning. *eLife*, *8*, e40578. https://doi.org/10.7554/eLife.40578
6. Furtado, O., Jr. (2026). *Statistics for movement science: A hands-on guide with SPSS* (1st ed.). https://drfurtado.github.io/sms/
7. Stergiou, N., Harbourne, R. T., & Cavanaugh, J. T. (2006). Optimal movement variability: A new theoretical perspective for neurologic physical therapy. *Journal of Neurologic Physical Therapy*, *30*(3), 120–129.