Experimental programme

Experimental programme: design, measurement, statistics, evaluation and collaboration

A population association

> A cohort association is not a personal exercise prescription.

- Lee et al. 2019: step activity and mortality in older women.
- Participants were not randomly assigned to step counts.
- Identify the population, measures and design.

A focused question

> Define the comparison, outcome, participants and conditions.

- Distinguish manipulated variables from observed exposures.
- A prediction with a reason is not an established mechanism.
- The null specifies the parameter tested.

Sampling and allocation

Process	Purpose and limitation
Random sampling	Select participants; representation still matters
Random allocation	Assign conditions; small groups may still differ
Repeat trials	Improve a planned within-person summary

Order and expectation

> Controls reduce particular biases; they do not remove every alternative.

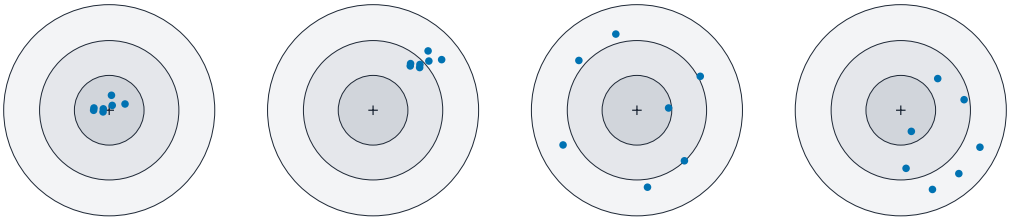
- Familiarisation can reduce learning; counterbalance condition order.
- Pilot the method and consider carry-over/recovery.
- State who is blinded and whether blinding is feasible.

Safety, choice and confidentiality

Obtain appropriate school approval before collecting data.

- Explain consent/assent, data use and withdrawal arrangements.
- A screening questionnaire is not medical clearance.
- Codes with a name key are pseudonymised, not **anonymous**.

Bias and scatter are different



A target analogy does not establish construct **validity**.

Validity and intended use

> A repeatable result can still miss the intended construct.

- Criterion: comparison with an appropriate reference.
- Internal: causal inference; external: application beyond the sample.
- Ecological: relevance to the real task.

Choose a suitable test

> Laboratory location does not guarantee **validity** or **reliability**.

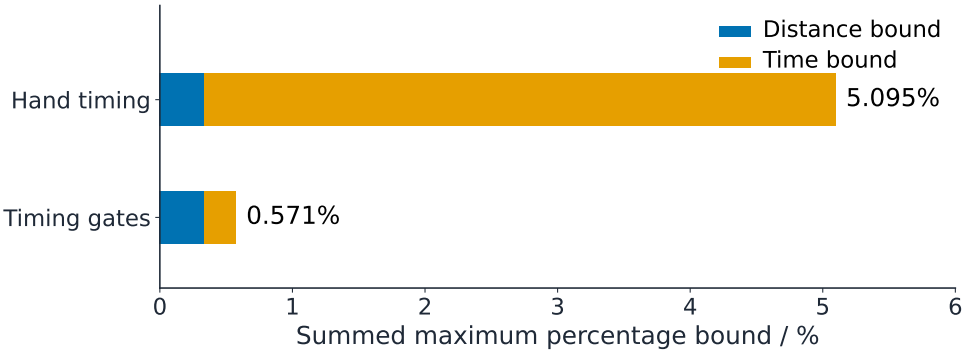
- Sit-and-reach reflects a specific task, not every joint.
- A planned change-of-direction course is not reactive agility by itself.
- Norms need a relevant population and matching protocol.

Error and calibration

Calibration compares with a reference; a justified correction retains uncertainty.

- Independent repeats can reduce random uncertainty of a **mean**.
- Repeats do not automatically remove systematic bias.
- Readout resolution is only one uncertainty component.

An uncertainty budget



Maximum-bound approximation; distance uncertainty remains.

Speed with supplied maximum bounds

1. **Speed.** $30.0 \,/\, 4.20 = 7.142857\ldots$ m/s
2. **Fraction.** $0.1 \,/\, 30 + 0.20 \,/\, 4.20 = 0.05095238\ldots$
3. **Absolute.** $7.142857\ldots \times 0.05095238\ldots = 0.363945\ldots$ m/s

Decision

Report 7.1 ± 0.4 m/s with a stated rounding convention.

Changing only the timing bound

1. **Time.** Change ± 0.20 s to ± 0.01 s
2. **Summed.** $0.333333\ldots\% + 0.238095\ldots\% = 0.571428\ldots\%$
3. **Absolute.** $\pm 0.0408163\ldots$ m/s; factor $8.91666\ldots$ improvement

Decision

Report 7.14 ± 0.04 m/s; the factor is not exactly ten.

Centre and sample spread

1. **Data.** 42, 45, 44, 43, 46 cm; **mean**/median 44 cm
2. **Spread.** Squared deviations sum 10; SD $= \sqrt{(10/4)} = 1.581\ldots$ cm
3. **Relative.** CV $= 1.581\ldots$ / 44 $\times$ 100 $= 3.593\ldots\%$

Decision

CV needs a suitable ratio scale and nonzero mean.

SD, SE and half-range

1. **SD.** 1.5811388… cm: spread of the observations
2. **SE.** SD / $\sqrt{5} = 0.7071068…$ cm under assumptions
3. **Range.** Range 4 cm; half-range 2 cm

Decision

Repeated trials from one person are not independent students.

Median and quartiles

1. **Sorted.** 42, 43, 44, 45, 46 cm; no single mode
2. **Convention.** Exclude the median from each half
3. **IQR.** Q1 = 42.5; Q3 = 45.5; IQR = 3 cm

Decision

State the quartile convention; software methods can differ.

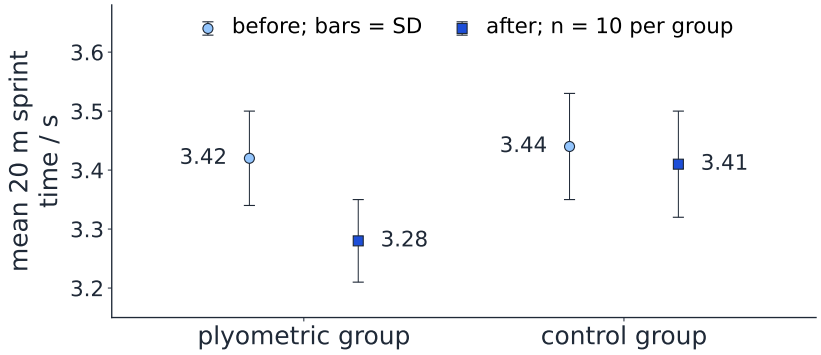
Which percentage?

Question	**For positive values 40 and 44**
Change from 40	$4 \,/\, 40 \times 100 = 10\%$
Error relative to a 40 reference	$4 \,/\, 40 \times 100 = 10\%$
Symmetric difference	$4 \,/\, 42 \times 100 = 9.5238\ldots\%$

Choose a useful display

Data or purpose	Suitable display
Categorical counts	Bar chart
Quantitative distribution	Histogram with suitable bins
Paired quantities	Scatter plot and justified model
Time course	Line plot

SD bars do not supply a paired p-value



Paired-change uncertainty needs within-person information.

Choose the test for the design

The sampling unit and assumptions matter before a test name.

- Paired: one signed difference per participant.
- Independent groups: use a suitable unpaired method.
- For a small paired sample, examine the distribution of differences.

Apply a paired t-test

1. **Reductions.** 0.10, 0.12, 0.14, 0.16, 0.18 s; **mean** 0.14 s
2. **SE.** SD = 0.0316228… s; SE = 0.0141421… s
3. **Statistic.** t = 0.14 / SE = 9.89949…; df = 4

Decision

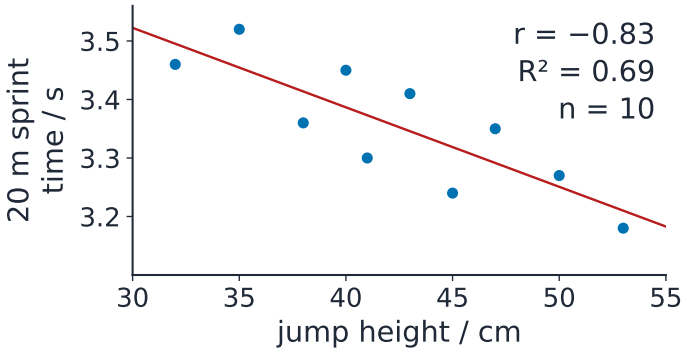
Above supplied two-sided 0.05 critical magnitude 2.776: reject the stated null.

What p = 0.03 means

> Probability is conditional on the specified null model and test assumptions.

- It concerns a statistic at least as incompatible with the null as observed.
- It is not the probability the null is true or the result was caused by chance.
- Significance does not prove cause or practical importance.

A fitted linear association



Sample $R^2 \approx 0.686$ accounts for variation in this fitted model, not cause.

Correlation and limits

> r near zero does not exclude a curved relationship.

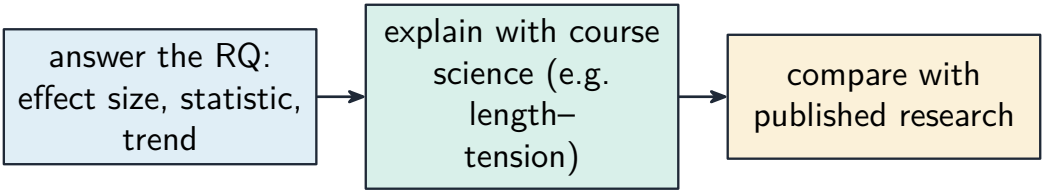
- Use the stated critical magnitude, n and test assumptions.
- Do not infer direction, intervention benefit or absence of confounding.
- One significant change and one non-significant change do not prove their difference.

Evaluate the actual limitation

Give the weakness, its consequence and a feasible improvement.

- Sample size does not repair selection or measurement bias.
- Use comparable definitions, familiarisation and justified repeats.
- Credit relevant research without selecting only agreement.

Sizing the conclusion



12 classmates cannot prove a law for all athletes

> Twelve classmates cannot prove a law

The scientific investigation

Criterion	**What the evidence should support**
Research design	Focused question and justified method
Data analysis	Correct processing and uncertainties
Conclusion	An answer proportionate to the evidence
Evaluation	Consequences and realistic improvements

Collaboration and individual work

> Do not present another student's same raw data set as your investigation.

- Optional groups of at most three; each has a unique question.
- Use distinct variables or allowed distinct selections from communal data.
- All report authoring, including methodology, is individual.

Experimental-programme hours

1. **SL.** Practical 20 + project 10 + investigation 10 = 40 h
2. **HL.** Practical 40 + project 10 + investigation 10 = 60 h
3. **Tools.** Techniques, technology and mathematics throughout inquiry

Decision

Do not add the two ten-hour components again to the totals.

Units and direct/inverse models

1. **Prefixes.** 3500 ms $=$ 3.5 s; 2.5×10^3 J $=$ 2.5 kJ
2. **Reciprocal.** 120 beats/min $=$ 2 beats/s $\rightarrow$ 0.5 s/beat
3. **Conditions.** At fixed distance, doubling time halves speed

Decision

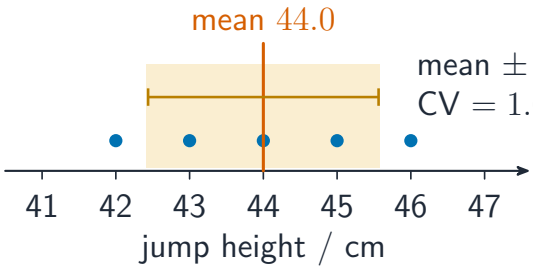
Convert units and keep the model's fixed conditions explicit.

Video, sensors and model outputs

A device or simulation needs a valid measurement model.

- Frames 120→240 at constant 240 frames/s: $120/240 = 0.50$ s.
- Check timestamps, **calibration** plane, parallax and sampling.
- Preserve raw inputs; label model-generated data and verify spreadsheet formulas.

Spread and reliability



mean $\pm$ SD: 44.0 ± 1.6 cm

CV $= 1.6/44 \times 100 \approx 3.6\%$

A small CV means a reliable measure

Mean, median, mode and range



$$\text{median} = 3$$

$$\text{mean} = \frac{2+3+3+5+7}{5} = 4$$

$$\text{mode} = 3 \text{ (most common)}$$

$$\text{range} = 7 - 2 = 5$$

The **mean** is pulled up by the 7

Testing components and perceived effort

> Name the component, protocol, endpoint and intended interpretation.

- Endurance, balance, reaction and coordination need defined tests.
- Use appropriate safe tasks and distinguish reactive agility from a planned course.
- An exertion scale needs its actual anchors; 0–10 is not interchangeable with 6–20.

Exam tips

State the measurement, model and limits of the inference.

- Keep SD, SE, half-range and instrument bounds distinct.
- Show denominators, units and unrounded working.
- No p-value or causal effect follows from SD-bar overlap alone.

Check yourself

Try these

Why are five people with ten trials not fifty independent participants?

Calculate the swimmer's speed and maximum-bound uncertainty.

Explain why a significant paired change does not establish a training effect.