Teacher quick start · Was It a Fair Comparison?

Target: identify confounds, specify matched conditions and plan repeated comparisons. Prerequisites: small-table reading, whole-number subtraction, temperature versus change; the prerequisite bridge includes a check and support. Paper and pencil only; optional calculator, number line, pointing or oral explanation. Intended for Grade 6; related NGSS expectations span Grades 6–8. Check local course placement.

Print menu

Print choices: Teach the Lesson includes the prerequisite bridge, model, guided practice and independent task. Use a Quick Check includes the independent task. Reteach: Repair the Comparison and Reteach: Specify the Plan provide focused follow-up tasks. Use a Later Check provides a separate later task. Each selection includes its own setup and data; match its activity titles and question numbers to the teacher key. The full student packet has 7 pages.

Planning estimate: 35–45 min Bridge→Model→Guided practice→the unseen independent task; 10–15 min for focused reteaching; 5–8 min for Later check · Cooling and a repaired test. These are not observed timings. If The independent task was already seen, use Later check · Cooling and a repaired test. Notice goal, relevant conditions, measurement/recording plan and limits separately.

Choose Repair the Comparison if changes are used to excuse unequal conditions; choose Specify a usable plan for vague plans or treating complete designs as material tests. Clarify unclear responses before reteaching. Accept a scientifically adequate labeled diagram, an oral explanation pointing to data, or connected prose. Ask the learner to point to the relevant row, restate the question, and explain a comparison aloud before selecting a reteach. A blank or unclear answer is insufficient evidence of a science difficulty.

No equipment or practical investigation. All setups/results are constructed.

Prerequisite teaching and answers

imagined setup cards, no collected results

Constructed teaching data: these values were invented for this activity, not collected in an experiment.

Condition	Sleeve A	Sleeve B
Water mass	100 g	100 g
Start	10°C	10°C
Cup and lid	Cup C + lid L	Cup C + lid L
Time / room	20 min / 22°C	20 min / 22°C
Reading method	Same method	Same method

Say: “I name the question first: compare complete sleeve designs. Sleeve design is the factor; water temperature after the same interval is the outcome. Mass, initial temperature, cup/lid, time, surroundings and reading procedure can affect the comparison, so we match them.”

Read before evaluating a comparison, question 1: Compared factor sleeve design A/B; measured quantity water temperature at start and end or temperature change over 20 min. Any two listed controls accepted. Read before evaluating a comparison, question 2: No; different mass would be another relevant difference. Do not make one-variable-at-a-time a rule for every scientific investigation; here it serves this question.

Readiness: Ask for a sample 10→14°C change (4°C warming) and a 30→26°C change (4°C cooling), then a pointed-to condition pair. If the data cannot be accessed, use oral/number-line support. If start/change is still confused, use Read Temperature Data Before Explaining It Compare the Measurements or Value or Change before the main critique.

Teacher model and complete key

same 15 minutes, 22°C room, cups/lids and reading method

Constructed teaching data: these values were invented for this activity, not collected in an experiment.

Setup	Water mass (g)	Start (°C)	After 15 min (°C)
Sleeve A	100	8	14
Sleeve B	200	10	13

Think aloud: “A warmed 6°C; B warmed 3°C. That describes these complete setups. A has 100 g starting at 8°C; B has 200 g starting at 10°C. Sleeve, mass and start changed. Smaller warming alone does not identify a sleeve effect. Subtraction does not remove those differences.”

Two things changed, question 1: +6/+3°C; cannot assign to sleeve alone. Two things changed, question 2: Water mass and start are confounds. A defensible repaired plan: both 100 g at 10°C in 22°C room for 15 min, matching cups/lids/position and consistent reading procedure; compare A/B. Accept both 200 g or both 8°C when all other conditions match and the proposed question remains explicit.

Two things changed, question 3: B’s complete setup warmed less in the supplied record; no isolated sleeve preference. Change comparison does not fix confounds. Temperature is a reading, not the total energy stored in a sample. Total energy also depends on the type, state and amount of matter. In these passive water examples, net thermal-energy transfer is from the warmer part toward the cooler part. A temperature record alone does not identify conduction, convection or radiation separately. Less warming does not mean cooling or stopped transfer.

Guided prompts and independent key

proposed setup cards

Constructed teaching data: these values were invented for this activity, not collected in an experiment.

Condition	Sleeve C	Sleeve D
Water mass / start	100 g / 10°C	100 g / 10°C
Cup	Cup C	Cup D
Lid	Lid L	No lid
Time / room	20 min / 22°C	20 min / 22°C
Reading method	Same method	Same method

Ask: “What is our factor? What else changes?” Anticipate “Only sleeves”; point to cup and lid. Then ask for D’s exact replacement conditions. Do not settle for “keep things fair.”

Guided practice · Repair setup cards, question 1: Cup geometry and presence of lid differ, so their effects are mixed with sleeve choice. Guided practice · Repair setup cards, question 2: D uses 100 g at 10°C, Cup C with lid L, 20 min in 22°C room, matching position and reading method. C unchanged. Accept a common alternative cup/lid specified for both.

Guided practice · Repair setup cards, question 3: Fresh matched samples for, e.g., three independent trials; same probe placement and stabilization/readout procedure; record trial/setup ID, mass, start, room, elapsed time, every final reading, method and departures. No invented instrument accuracy. For a material-only question, sleeve dimensions, layers/fit/gaps also need matching.

24°C room, cups/lids/position and reading method match

Constructed teaching data: these values were invented for this activity, not collected in an experiment.

Setup	Mass (g)	Start (°C)	Time (min)	Final (°C)
Sleeve E	80	12	10	18
Sleeve F	120	12	20	16

Independent task · Fresh critique, question 1: E warms 6°C, F 4°C. F’s complete setup shows less warming; no isolated sleeve effect. Independent task · Fresh critique, question 2: Mass and elapsed time differ; repair to both 100 g starting 12°C for 20 min in 24°C room with matched cups/lids/position and procedure. Accept both 80 or 120 g, or both 10 min, with explicit matching. Independent task · Fresh critique, question 3: Fresh samples and multiple matched trials, retain raw readings and notes. Accept specific limits about other conditions, small trial count, constructed data, unknown accuracy or mechanism.

Follow-up · Repair the comparison

22°C room, 20 minutes, same cups/lids and method

Constructed teaching data: these values were invented for this activity, not collected in an experiment.

Setup	Mass (g)	Start (°C)	Final (°C)
Sleeve H	60	10	17
Sleeve J	120	10	15

Think aloud: “Starts match, but masses do not. H warms 7°C, J 5°C. I can describe these setups; I cannot assign the difference only to the sleeve. I choose one common mass for a retest.”

Subtraction does not fix a confound, question 1: H +7°C, J +5°C. No isolated sleeve effect. Accept both 60, both 120 or another equal mass, same 10°C start, 20 minutes, 22°C room, cups/lids/position and reading procedure, with fresh repeated samples.

Data fresh follow-up check · matched samples; 23°C room, 15 minutes

Constructed teaching data: these values were invented for this activity, not collected in an experiment.

Setup	Mass (g)	Start (°C)	Final (°C)
Sleeve U	100	12	17
Sleeve V	100	12	16

Subtraction does not fix a confound, question 2: U warms 5°C, V 4°C; V limits warming more in this one supplied trial. This is a nominal record comparison, not a reliable or universal guarantee. More matched trials and information about accuracy/conditions would support a stronger claim. Accept cautious withholding of a confident preference because the difference is small. Do not dismiss that alternative.

Follow-up · Specify the plan

three fresh 100 g trials, 10°C start, 22°C room, 20 minutes; matched conditions

Constructed teaching data: these values were invented for this activity, not collected in an experiment.

Trial	Sleeve J final (°C)	Sleeve K final (°C)
1	15	16
2	16	17
3	15	16

Think aloud: “J warms 5,6,5°C; K warms 6,7,6°C. J has 1°C less warming in each row. These invented records favor J under these conditions; they do not establish a specific material property. A usable plan says what to match and what to record.”

Specify a usable plan, question 1: J lower in each trial. No isolated material property or exact transfer pathway. Accept a cautious preference or a qualified request for accuracy/repeats instead of a confident ranking.

Specify a usable plan, question 2: Example paper plan: compare complete J/K sleeves; fresh equal 100 g samples at 10°C, identical cups/lids and positions in 22°C room; 20 min; consistent probe position/readout stabilization; three fresh matched trials. Record trial ID, sleeve version, mass, initial and final raw temperatures, room temperature, elapsed time, instrument/method ID and departures. Mark missing readings as missing. Three is a planning choice, not guaranteed sufficient evidence. No need to name a thermometer specification that was not supplied.

Fresh Data fresh comparison check · equal 100 g, 23°C room, matching cups/lids and method

Constructed teaching data: these values were invented for this activity, not collected in an experiment.

Setup	Start (°C)	Time (min)	Final (°C)
Sleeve R	11	20	16
Sleeve S	11	10	15

Specify a usable plan, question 3: Match duration, e.g. 20 min for both, and repeat. R warms 5°C over 20 min; S 4°C over 10 min. Do not attribute this to the sleeve or simply divide to invent a constant warming rate.

Later check · Key and limits

22°C room, 15 minutes, cups/lids and method match

Constructed teaching data: these values were invented for this activity, not collected in an experiment.

Setup	Mass (g)	Start (°C)	Final (°C)
Sleeve P	100	36	29
Sleeve Q	150	36	32

Later check · Cooling and a repaired test, question 1: P falls 7°C; Q falls 4°C. Mass differs (100/150 g); higher final Q reading does not isolate sleeve effect.

Data Tr · fresh matched 100 g samples, 36°C start, 22°C room, 15 minutes

Constructed teaching data: these values were invented for this activity, not collected in an experiment.

Trial	Sleeve P final (°C)	Sleeve Q final (°C)
1	30	32
2	31	32
3	30	33

Later check · Cooling and a repaired test, question 2: P falls 6,5,6°C; Q falls 4,4,3°C. Q limits cooling more in each matched trial; endpoints or changes accepted. Bound to these constructed records and conditions; no universal, actual product or exact mechanism claim. Later check · Cooling and a repaired test, question 3: No, analysis of supplied values is not physical construction/testing.

This checks critique and bounded comparison in a cooling context. It does not prove mastery or establish equivalence to the independent task. If the learner explains the confound but cannot write a plan, clarify the expression demand and offer an oral plan before Specify the Plan.

Response guidance and sources

Accept a scientifically adequate labeled diagram, an oral explanation pointing to data, or connected prose. Ask the learner to point to the relevant row, restate the question, and explain a comparison aloud before selecting a reteach. A blank or unclear answer is insufficient evidence of a science difficulty.

Keep separate notes for stated question, relevant controls, recording/repetition, and limits. Fictional “F always works better” needs attribution limits; “we will keep everything the same” needs actual conditions; “throw away an odd trial” needs raw-record retention and a documented reason. More trials do not automatically remove systematic bias. Display precision is not accuracy.

Official sources checked 2026-09-29: NGSS MS-PS3-4 (https://www.nextgenscience.org/pe/ms-ps3-4-energy) supports temperature/energy distinctions and investigation planning; MS-PS3-3 (https://www.nextgenscience.org/pe/ms-ps3-3-energy) supports transfer direction and requires actual design, construction and testing. These paper tasks address selected components, not complete performance expectations. Original text/data/graph; no third-party passages or graphics reproduced. Classroom copying/adaptation follows https://mentorteaching.com/terms-of-use/.

Was It a Fair Comparison? addresses planning components of MS-PS3-4 and evidence reasoning connected to MS-PS3-3. It does not cover all mass/material/temperature relationships or fulfill actual device construction/testing.

Select response practice and keep its check separate

The current preparation menu supplies Repair the comparison and Specify a usable plan as response-practice files, with each Fresh check in a separate reserved file. Use only the practice selection while explaining and practicing. Then give the matching reserved fresh-check file, retaining the original question numbers so the matching answer guidance still applies. Keep the separate later task for a different occasion. The earlier combined sheets are historical; do not distribute them during response teaching.

A different prompt is not a validated equivalent test. Record which materials and access route were used, and clarify unclear wording or access before choosing further teaching.

Fictional teaching response — sound for the named task

Named task: Independent task: Fresh critique, questions 1–3

Fictional response — sound for the named task: “E rises 6°C and F 4°C, so F warms less in these particular setups. It does not isolate sleeve design: 80 versus 120 g and 10 versus 20 minutes also differ. A repaired comparison could use 100 g per sample, 12°C starts, a 24°C room and 15 minutes with matching cups, lids, exposure and reading method. Repeat with three fresh matched pairs, recording actual starts, masses, times, final readings and anomalies. This still supports only those designs/conditions; accuracy information and further conditions would matter before broader claims.”

Visible evidence: Correct calculations, identified confounds, a specified comparison/recording plan and a limit.

Not established: A measured repaired outcome or physical suitability of an apparatus.

Clarify: Ask which original differences the proposed constants remove and which questions remain open.

Next teaching/evidence: Use a separate current fresh check after practice or the Later assessment on another occasion. Never claim the repaired result was measured.

This response is invented for teacher interpretation. Accept defensible alternatives. Record access, prior exposure and the actual response; do not assign an ability label or claim mastery from this example.

Later assessment

Fictional teaching response — partially developed

Named task: Independent task: Fresh critique, questions 1–3

Fictional response — partially developed: “F warmed less, but the masses and times differ. Make them the same.”

Visible evidence: The observed ordering and both confounds.

Not established: The actual changes, specified comparison values or repeat/recording detail.

Clarify: Request both subtractions and concrete masses, starts and duration, then a record of repeated fresh trials.

Next teaching/evidence: Use Specify a usable plan, practice only, followed by its separate reserved fresh check. Save Later assessment.

This response is invented for teacher interpretation. Accept defensible alternatives. Record access, prior exposure and the actual response; do not assign an ability label or claim mastery from this example.

Specify a usable plan

Reserve until after this practice: Specify a usable plan — fresh check

Reserve: Later assessment

Fictional teaching response — plausible error

Named task: Independent task: Fresh critique, questions 1–3

Fictional response — plausible error: “E warms 0.6°C per minute and F 0.2°C per minute, so F is three times the better sleeve. Dividing by time fixes the comparison.”

Visible evidence: Correct average change-per-minute calculations for the two records.

Not established: An isolated sleeve effect: dividing does not remove the mass difference or establish identical temperature-change behavior over unequal intervals.

Clarify: Ask which setup conditions still differ after that arithmetic and what ‘three times better’ would measure.

Next teaching/evidence: Use Repair the comparison, practice only, to inspect conditions and specify a matched comparison. Then give its separate fresh-check file; do not reveal it during practice.

This response is invented for teacher interpretation. Accept defensible alternatives. Record access, prior exposure and the actual response; do not assign an ability label or claim mastery from this example.

Repair the comparison

Reserve until after this practice: Repair the comparison — fresh check

Reserve: Later assessment