

Two Creams, One Choice: Confirm the New Substance, Then Judge the Design

Learning target: I can confirm from property data that a reaction made a new substance, then evaluate competing designs against stated criteria and constraints using test results, choose one, and justify the choice including its trade-off.

Transfer and design

About 50 min

Pairs

MS-ETS1-2 · practice

MS-PS1-2 · practice

engineering design

data analysis

writing

Standards connections

MS-ETS1-2 (practice, evaluating three competing designs against criteria and a constraint with a systematic table): Students receive trial results for three cream designs, convert them to percentages, check each against two criteria and one cost constraint in a table, and justify a choice with its trade-off. The data are supplied, so this builds the systematic-evaluation practice toward the PE rather than fully assessing it.

MS-PS1-2 (practice, using melting range, solubility, density and odor to decide whether Compound S is a different substance from Compound R): Part A puts before-and-after property data in front of students, including three tests that do not separate them, and asks whether the reaction produced a new substance. It transfers the chapter's analysis to a new context inside the PE's assessment boundary.

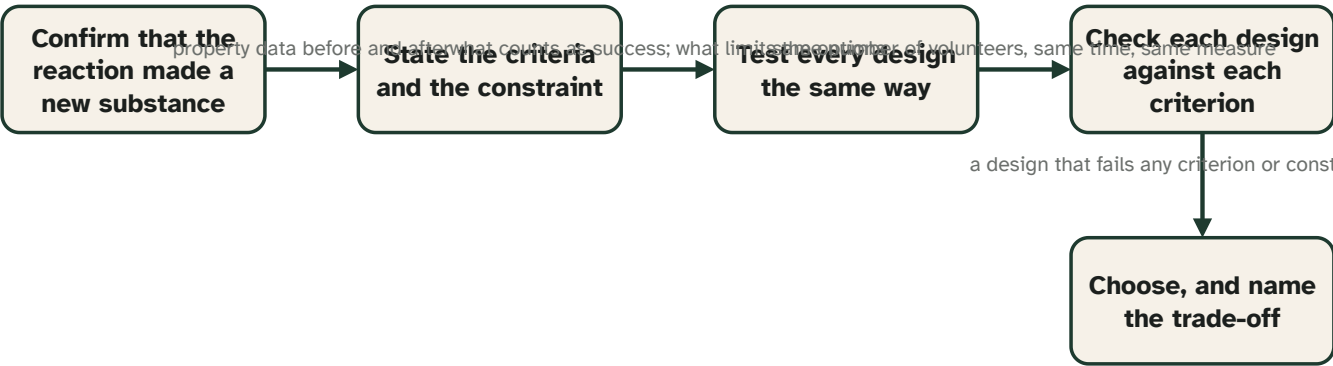
Purpose: extend · **Time:** about 50 minutes · **Format:** pairs

Answers with reasoning

The brief

Lantern Labs sells an anti-itch cream for insect bites. Its active ingredient, Compound R, is a powder taken from the leaves of a mint relative. Customers like the relief but complain that the cream stings for the first minute. The company's chemist ran a reaction on Compound R, hoping to make a new substance, Compound S, that keeps the relief and loses the sting. She handed the review team two sets of results: the property tests she ran on R and on S, and a volunteer trial of three cream designs. The team's job is to decide whether S is really a new substance and, if so, which cream the company should make.

Figure 1. The review team's process



The process is systematic on purpose: the same checks are applied to every design in the same order.

Part A: Is Compound S a new substance?

Table 1. Property tests on Compound R (before the reaction) and Compound S (after), teaching dataset

Simulated teaching dataset: the numbers are invented to be realistic, not measured.

Property	Compound R (before)	Compound S (after)
Appearance	pale yellow powder	white powder
Melting range	121 to 122 °C	121 to 123 °C
Solubility in water at 25 °C (g per 100 mL)	1.8	6.4
Density (g/cm ³)	1.31	1.29
Odor	sharp, minty	none
Burns when held in a flame?	yes	yes

Simulated data for teaching. All tests on pure, dried powders at the same temperature. The density method is good to about ± 0.03 g/cm³; repeat solubility tests agree within 0.2 g per 100 mL.

1. Sort the six properties in Table 1 into two groups: those that clearly differ between R and S, and those that do not (either the same, or different by less than the method's uncertainty). Use the uncertainty statements in the caption.

Answer: Clearly differ: appearance (pale yellow versus white), solubility at 25 °C (1.8 versus 6.4 g per 100 mL, a difference far larger than the 0.2 g repeat spread), odor (sharp and minty versus none). Do not differ: melting range (121-122 versus 121-123 °C, overlapping), density (1.31 versus 1.29 g/cm³, a 0.02 difference smaller than the ± 0.03 uncertainty), and burning (yes for both).

Why: The sort must use the caption's uncertainty statements: the density gap is inside the method's error, so it counts as no difference.

Also acceptable: Accept a student who lists appearance as 'weak difference' as long as solubility and odor are in the differ group.

- Appearance, solubility and odor in the 'differ' group
- Melting range, density and burning in the 'do not differ' group
- Density judged against the ± 0.03 uncertainty

2. The melting ranges overlap. Does that mean R and S are the same substance? What does the narrowness of each range (about 1 to 2 °C) tell you about the two powders?

Answer: No. Two different substances can happen to melt at nearly the same temperature, so an overlapping melting range is consistent with sameness but does not prove it; it simply gives no information either way. The narrow ranges (1 to 2 °C) show that each powder is a pure substance rather than a mixture, because impurities make a solid start melting lower and over a wider range. So both R and S are pure, and the other properties must decide whether they are different.

Why: This applies the chapter's melting-point ideas in both directions: shared value is not proof of identity; narrow range shows purity.

Also acceptable: Accept 'a melting point can be the same by coincidence' with the purity point.

- Shared melting range does not prove sameness
- Narrow range means each is a pure substance

3. Calculate how many times more soluble S is than R at 25 °C. Then calculate the density difference and compare it with the ± 0.03 g/cm³ uncertainty: is the density difference real? Show your work.

Answer: Solubility: $6.4 \div 1.8 = 3.6$, so S is about 3.6 times as soluble as R at 25 °C. Density: $1.31 - 1.29 = 0.02$ g/cm³. The method is only good to ± 0.03 g/cm³, so a 0.02 difference is smaller than the uncertainty and is not a real difference; the densities should be treated as the same.

Why: The ratio shows the solubility change is large; the density subtraction against the stated uncertainty shows how to decide whether a small difference means anything.

Also acceptable: Accept 3.56 or 'about 3.5 times'; accept 'the density difference could be measurement error'.

4. Should the review team accept that S is a different substance from R? Give your decision with the two strongest pieces of evidence, and name the one test the chemist could have skipped because it could not tell R from S.

Answer: Yes, the team should accept that S is a different substance. Strongest evidence: (1) solubility at 25 °C rose from 1.8 to 6.4 g per 100 mL, a 4.6 g difference, twenty-three times the 0.2 g repeat spread, on tests done at the same temperature; (2) the odor changed from sharp and minty to none (color is a third, weaker piece). The chemist could have skipped the burning test: both powders burn, so it cannot separate them. (The melting range and density also failed to separate them, but those are reasonable tests to try; the flame test was the least informative.)

Why: Students must name solubility as the strongest evidence, support it with a second changed property, and identify an uninformative test.

Also acceptable: Accept the melting range or density as the 'could have skipped' test if the student explains that it gave no difference; the flame test is the best answer because it was certain to be uninformative for two organic powders.

- Decision: new substance
- Solubility with numbers and the same-temperature point
- Odor (or appearance) as the second piece
- Names a test that could not tell R from S

Part B: Evaluate the three designs

- ☐ Criterion 1: At least 70% of volunteers report itch relief within 30 minutes.
- ☐ Criterion 2: No more than 10% of volunteers report stinging.
- ☐ Constraint: The cream must cost no more than \$4.00 per tube to make.

Table 2. Volunteer trial of three cream designs, teaching dataset

Simulated teaching dataset: the numbers are invented to be realistic, not measured.

Design	Active ingredient and amount per tube	Volunteers with relief within 30 min (of 40)	Volunteers reporting stinging (of 40)	Cost to make per tube
Cream 1	Compound R, 2.0 g	33	14	\$2.60
Cream 2	Compound S, 2.0 g	30	2	\$3.90
Cream 3	Compound S, 1.0 g	23	1	\$3.10

Simulated data for teaching. Each cream was tested on 40 adult volunteers under the same conditions; each volunteer used one cream on one bite and reported at 30 minutes.

5. Complete the evaluation table. Convert each count to a percentage of 40 and write yes or no in each check column. Show one percentage calculation below the table.

Answer: Cream 1: relief $33 \div 40 \times 100 = 82.5\%$, meets Criterion 1 (yes); stinging $14 \div 40 \times 100 = 35\%$, fails Criterion 2 (no); \$2.60 meets the constraint (yes). Cream 2: relief $30 \div 40 \times 100 = 75\%$, yes; stinging $2 \div 40 \times 100 = 5\%$, yes; \$3.90, yes. Cream 3: relief $23 \div 40 \times 100 = 57.5\%$, no; stinging $1 \div 40 \times 100 = 2.5\%$, yes; \$3.10, yes.

Why: Each percentage is the count divided by 40 times 100. Only Cream 2 has three yeses.

Also acceptable: Accept 82.5% rounded to 83% and 57.5% rounded to 58%.

- All six percentages correct
- All nine yes/no entries correct
- One calculation shown

6. Which design or designs pass every criterion and the constraint? For each design that fails, name exactly what it fails.

Answer: Only Cream 2 passes everything. Cream 1 fails Criterion 2: 35% of volunteers reported stinging, far above the 10% limit. Cream 3 fails Criterion 1: 57.5% reported relief, below the 70% minimum. Both failing creams meet the cost constraint.

Why: The evaluation table makes the pass/fail pattern explicit; students should name the exact criterion failed.

Also acceptable: None.

- Cream 2 identified as the only full pass
- Cream 1 fails stinging; Cream 3 fails relief

7. A manager says, 'Cream 1 relieved the most people, so it is the best design.' Using Figure 1, explain why that is not how a systematic evaluation works.

Answer: A systematic evaluation checks every design against every criterion and the constraint, and a design that fails any one of them is out, no matter how well it does on another. Cream 1 scores highest on relief but fails the stinging criterion by a wide margin (35% versus 10%), so it is eliminated at the 'check each design against each criterion' step in Figure 1. Picking the winner on one favorite number skips that step and ignores the criteria the company agreed on.

Why: This is the heart of MS-ETS1-2: the process is systematic, and criteria are applied together, not one at a time by preference.

Also acceptable: Accept 'the manager is only looking at one criterion'.

- States that all criteria and the constraint must be met
- Cites Cream 1's stinging failure
- Refers to the Figure 1 step

Part C: Decide and justify

8. Write the review team's recommendation in a paragraph. It must (a) name the cream to make, (b) cite the percentages and cost that show it meets each criterion and the constraint, (c) name the trade-off the company accepts by choosing it over the design with the most relief, and (d) explain in one sentence how Part A's property evidence supports trusting that this cream contains a different substance from the old one.

Answer: Model response: The review team recommends Cream 2, made with 2.0 g of Compound S per tube. In the trial, 75% of volunteers reported relief within 30 minutes, which meets the 70% criterion; 5% reported stinging, which is under the 10% limit; and it costs \$3.90 per tube, inside the \$4.00 constraint. Cream 1 relieved more people (82.5%) but 35% of volunteers felt stinging, so it fails the second criterion, and Cream 3 is gentler still (2.5% stinging) but relieves only 57.5%, so it fails the first. The trade-off the company accepts is slightly less relief (75% instead of 82.5%) and a higher cost (\$3.90 instead of \$2.60) in exchange for almost no stinging. The team is confident Cream 2 contains a different substance from the old cream because Compound S has a solubility 3.6 times that of Compound R at the same temperature and has lost R's sharp odor, which are characteristic properties that cannot both change unless a new substance formed.

Why: Full credit needs the named design, the three numbers tied to the three checks, a stated trade-off against Cream 1, and a link back to the Part A property evidence.

Also acceptable: A student who recommends Cream 1 because relief is the most important criterion has ignored the agreed criteria; that response earns credit only for the trade-off and Part A sentences.

- Names Cream 2
- 75%, 5% and \$3.90 tied to the two criteria and the constraint
- Trade-off: less relief and higher cost than Cream 1, in exchange for no stinging
- Part A evidence (solubility and odor) cited as the reason S is a new substance

Part D: The constraint changes

9. The company lowers the cost constraint to \$3.50 per tube. Which designs still pass? Propose a modified design that might pass everything, say what you would change and why, and name the one test you would run before recommending it.

Answer: With a \$3.50 limit, Cream 2 (\$3.90) now fails the constraint, Cream 1 still fails stinging and Cream 3 still fails relief, so no current design passes. A reasonable modification: a cream with about 1.5 g of Compound S, between Creams 2 and 3. Reasoning: the cost should land between \$3.10 and \$3.90 (roughly \$3.50), stinging should stay low (both S creams were under 5%), and relief should fall between 57.5% and 75%, so it might just reach 70%. The team cannot assume it does; the test to run before recommending it is a new volunteer trial of the 1.5 g cream under the same conditions (40 volunteers, 30 minutes, same reporting), checking the relief percentage against the 70% line.

Why: This is the modify-and-retest move. Full credit needs the realization that nothing passes, a modification with a reason drawn from the data, and a named test.

Also acceptable: Accept a proposal to find a cheaper way to make Compound S while keeping 2.0 g per tube, if the student says the trial does not need repeating because the cream itself is unchanged, but a cost check does.

- States that Cream 2 now fails the constraint and nothing passes
- Proposes a modification between the tested designs with a data-based reason
- Names a repeat trial under the same conditions as the test

10. Explain why the trial gave every cream the same number of volunteers, the same 30-minute limit and the same way of reporting. What would go wrong in the evaluation table if Cream 3 had been tested on 60 people for 45 minutes?

Answer: Keeping the number of volunteers, the time limit and the reporting method the same makes the three results comparable: any difference between creams comes from the cream, not from the test. If Cream 3 had been tested on 60 people for 45 minutes, its relief count could not be compared with the others' 30-minute counts out of 40; more time would raise its relief number for reasons that have nothing to do with the cream, and the percentages in the table would be computed on different bases, so the yes/no checks would be unfair.

Why: Fair testing is the 'test every design the same way' step in Figure 1. Students should name at least one specific way the unequal test would mislead the table.

Also acceptable: Accept 'it would not be a fair test' with one concrete reason.

- States that identical conditions make differences attributable to the cream
- Explains a specific distortion from 60 people or 45 minutes

Scoring rubric

Criterion	4 • Exemplary	3 • Proficient	2 • Developing	1 • Beginning
Confirming the new substance (Part A)	Sorts all six properties correctly using the stated uncertainties, explains the melting-range overlap and purity, and decides 'new substance' on solubility plus a second property.	Decides correctly on solubility and one other property, but misjudges the density difference or the melting-range meaning.	Decides on appearance or odor alone, or treats the shared melting range as proof of sameness.	No decision, or a decision that contradicts the data.
Systematic evaluation (Part B)	All percentages and yes/no checks correct; identifies Cream 2 as the only pass and names exactly what each other design fails; explains why one favorite number cannot decide.	Table correct with one slip; failures named; the manager question answered thinly.	Several table errors, or passes judged on one criterion only.	Table incomplete or evaluation not tied to the criteria.
Decision, justification and trade-off (Part C)	Recommends Cream 2 with the three numbers tied to the checks, states the trade-off against Cream 1, and links the choice to the Part A property evidence.	Recommends Cream 2 with numbers and a trade-off; the link to Part A is missing or vague.	Recommends a cream with numbers but no trade-off, or recommends Cream 1 while acknowledging the stinging failure.	No recommendation, or a recommendation without evidence.
Responding to the changed constraint and fair testing (Part D)	Recognizes that no design passes at \$3.50, proposes a data-based modification, names the retest, and explains fair-test conditions with a specific distortion.	Proposes a modification and a test, but misses that Cream 2 now fails or gives a vague fair-test answer.	Proposes a change without data reasoning, or omits the test.	Part D blank or off topic.

Teacher notes

Everything in this sheet is fictional and all numbers are simulated: Lantern Labs, Compounds R and S, the creams and the volunteer trial. The design mirrors the textbook's salicylic acid to aspirin story (a reaction makes a new substance that keeps the benefit and loses a side effect, judged against criteria) without using a real medicine, so no dosing or health claims arise; if students raise aspirin, keep to 'medicines are judged against criteria and none is perfect'. Table 1 is built so that two tests fail to separate R and S (overlapping melting range, density inside the ± 0.03 uncertainty) and the flame test is uninformative, which is why students must rank evidence rather than count differences. The evaluation table in q5 is the systematic process of MS-ETS1-2; insist on the percentage calculations and the yes/no columns before the recommendation. Part D's \$3.50 constraint is the modify-and-retest step; the 1.5 g proposal is an estimate, and the key point is that a new trial, not a guess, decides it.

Scaffold offered to students: Percentages out of 40 and a pass/fail habit

To turn a count out of 40 into a percentage, divide by 40 and multiply by 100: $30 \div 40 \times 100 = 75\%$. A shortcut: each volunteer is 2.5%, so multiply the count by 2.5. For the check columns, write the rule next to the number before deciding: '75% $\geq$ 70%? yes.' For Part A, start by crossing out any property whose two values are identical or closer than the uncertainty in the caption; what is left is your evidence.

Extension: Design the next trial

Cream 2 passed, but 25% of volunteers got no relief within 30 minutes. Propose a fourth design (for example, a different amount of Compound S, or S mixed with a small amount of R) and predict its three results in Table 2's format, with a reason for each prediction based on Creams 1 to 3. Then write the one new criterion you would add to the company's list, explain why the current three criteria let a cream that smells bad or separates in the tube pass, and describe how you would measure your new criterion fairly.