

The Decision Meeting: Six Hand-Warmer Prototypes, Three Customers

Investigation question: Given test data from six hand-warmer prototypes, which prototype should go to each of three customers with different criteria and constraints, and when none fits, which characteristics from different prototypes can be combined into a better design?

Data-driven decision

Groups of 4

1 session

Setup 5 · Activity 42 · Cleanup 3 min

MS-ETS1-3 · assesses

MS-ETS1-2 · practice

MS-PS1-6 · supports

Scientific purpose

Teams play the engineering review board of a hand-warmer company. They receive a simulated test report on six iron-powder prototypes that differ in iron mass, pouch material and whether salt was added, with peak temperature, time to reach 40 °C, hours above 40 °C and 35 °C, mass and cost for each, plus full temperature-versus-time curves for three of them. Three customers have written briefs with criteria and constraints. Teams build a pass-or-fail decision matrix, match a prototype to each customer, discover that one customer has no passing prototype, trace which design feature controlled which result, and propose a new design that combines the best characteristics with predicted numbers bracketed by the data.

Why this helps students with this chapter

In Lab 2 students tested their own packs; here they sit on the other side of the table and judge someone else's. Six prototypes that each win on one measure and lose on another make it impossible to pick 'the best one' without first asking 'best for whom'. The pairs that differ by a single feature let students see that doubling the iron raised the peak, that the felt pouch traded peak for time, and that without salt the reaction released energy too slowly to warm the pouch at all. The customer with no passing prototype forces the real engineering move: take the fast start from one design and the long life from another and predict what the combination will do.

Standards connections

MS-ETS1-3 (assesses, analyzing test data from several solutions to identify the best characteristics of each and combine them into a new solution): The six prototypes differ by one feature at a time, so students can attribute each result to a feature, name the best characteristic of each prototype for a given criterion, and combine them into a proposed design with predicted performance. That is the PE as written, done on supplied data.

MS-ETS1-2 (practice, evaluating competing designs systematically against criteria and constraints): Students score every prototype against every customer in a pass-or-fail matrix before choosing, which is a systematic process; the criteria and constraints are supplied by the customer briefs, so the role is practice rather than full evidence.

MS-PS1-6 (supports, judging a thermal-energy device by amount, time and temperature): No device is built in this lab; the data are exactly the amount-time-temperature measurements the PE's assessment boundary

names, so the lab supports the design work of Lab 2 and W9 by showing how finished prototypes are compared.

Roles (rotate each session)

- **Meeting chair:** Reads each customer brief aloud, keeps the team working through the matrix one customer at a time, and makes sure every pass or fail has a number beside it.
- **Data analyst:** Plots the three temperature curves, computes the rates, costs per hour and any interpolated values, and checks every number the team writes against the report.
- **Customer advocate:** Speaks for the customers: checks each proposed match against every criterion and every constraint in the brief, and refuses a match that breaks even one constraint.
- **Design recorder:** Writes the decision matrix, the feature-comparison table and the new design proposal, and makes sure every predicted number in the proposal is bracketed by data from two existing prototypes.

Materials

Item	Quantity	Per	Note
Recording sheet with the prototype test report, the temperature-time dataset and the three customer briefs	1	student	
Graph paper or the graph block on the recording sheet	1 sheet	student	
Calculator	1	group	
Colored pencils (three colors) and a ruler	1 set	group	
Index cards, one per prototype, for sorting (optional)	6	group	Copy the six summary rows onto cards so the team can physically sort them by customer.
Sticky notes for the class decision wall	a pad	class	

Estimated cost: 0 to 5 per class (an estimate; Estimate. Printing only; index cards and sticky notes from stock. No chemicals, no equipment.)

Safety

- No chemicals, no heat sources and no building in this lab; it is a data and decision task. Do not try to build iron-powder pouches at home or in class: iron powder is a fine dust that should not be breathed, and the reaction with salt and air is for the dataset only.
- If the class uses scissors to cut index cards, standard scissors rules apply.

Teacher preparation and setup

1. Print the recording sheet for every student; the test report, the dataset and the customer briefs are all on it. The numbers are a teaching dataset invented for this lab and the sheet says so.
2. Work the decision matrix yourself first (the key gives it) so you can spot a team that passes P2 for a customer with a 50 °C ceiling or P4 for a customer with a mass limit.
3. Draw a blank class decision wall on the board: six prototype rows, three customer columns, space for sticky notes.
4. If using cards, copy the six summary rows onto index cards per group, one prototype per card.
5. Decide whether to run the optional opening: hold up two prototypes 'P2 and P3' (any two objects) and ask which is the better hand warmer; the only correct answer is 'for whom?'.

Investigation design

Independent variable	Design features that differ between prototypes: iron mass (20, 40 or 60 g), pouch material (thin cotton or thick felt) and whether salt was added
Dependent variable	Peak temperature, time to reach 40 °C, hours above 40 °C and above 35 °C, mass and cost of each prototype, as reported in the simulated test data
Controlled	Every prototype tested in the same 22 °C room in still air with the probe inside the pouch; Each test started the moment the sealed wrapper was opened and air reached the iron; Same pouch sizes and the same amount of salt (2 g) where salt was used; Same price list used to cost every prototype
Trials	The report gives one test per prototype, which students are asked to notice and criticize in q9; the lab's own repetition is across teams, who should reach the same matrix from the same data.
Comparison	Pairs of prototypes that differ by exactly one feature (P1 and P2, P2 and P3, P3 and P4, P3 and P5, P2 and P6), and every prototype against each customer's full list of criteria and constraints.

Student procedure

1. Read the test report and the three customer briefs on your recording sheet. The meeting chair reads each brief aloud while the customer advocate underlines every criterion (a goal with a number) and every constraint (a limit with a number).
2. The data analyst plots the temperature-time curves for P1, P2 and P3 on one graph in three colors, draws a horizontal line at 40 °C and one at 35 °C, and marks where each curve crosses them. Check the crossing times against the 'minutes to 40 °C' and 'hours above' columns of the report.
3. Fill in the feature-comparison table: for each listed pair of prototypes that differ by one feature, write the feature that changed and what happened to the peak, the minutes to 40 °C and the hours above 35 °C. These are your cause-and-effect findings.
4. Build the decision matrix one customer at a time. For every prototype write PASS or FAIL for that customer and, beside it, the first criterion or constraint it fails with the number (for example 'FAIL: peak 53 °C, limit 50'). A prototype passes only if it meets every criterion and every constraint.

5. Match a prototype to each customer who has one. If more than one passes, choose the cheapest or the lightest and say which rule you used. If none passes, write 'no prototype passes' and list, for each near-miss prototype, the one thing that kept it out.
6. For the customer with no passing prototype, design a new one. Use the feature-comparison table to pick the characteristic that gives a fast start and the characteristic that gives a long working time, choose an iron mass, and predict the peak, minutes to 40 °C, hours above 40 °C, mass and cost. Every prediction must sit between the values of two existing prototypes, and you must say which two.
7. Post your three matches and your new design on the class decision wall with sticky notes. Compare with other teams: where teams disagree, the numbers settle it.
8. Complete the analysis questions individually, then compare answers within the team.

Expected observations

- Teams reach the same matrix: P3 passes only for Maya; P1 passes only for Mr. Alvarez; no prototype passes for Dani (P2 fails on peak and hours, P3 and P4 on start time, P4 also on mass and cost, P1 and P6 on hours, P5 on everything).
- The graph shows P2 crossing 40 °C at about 7 minutes and falling back through it at about 165 minutes, P1 crossing up at about 8 minutes and down at about 60, and P3 crossing up at about 45 minutes and staying above 40 °C past 300 minutes.
- Feature comparisons come out consistently: iron mass raised peak and hours; felt lowered the peak and slowed the start but lengthened the working time; no salt made P5 useless and P6 slow.
- Dani designs cluster around 45 to 50 g of iron with salt in a felt pouch that has a thin-cotton panel or window, predicting a start of 10 to 15 minutes, a peak of 44 to 47 °C and 4 to 6 hours above 40 °C.

Likely difficulties

- Teams pass a prototype that meets the criteria but breaks a constraint (P2 for Mr. Alvarez because it is fast and long, forgetting the \$1.00 and 50 °C limits). The customer advocate's job is to catch exactly this.
- Students read 'hours above 35 °C' as the total working time and forget that the clock for Dani's 'hours above 40 °C' starts only when 40 °C is reached; the graph with both lines drawn fixes it.
- Converting 0.9 hours to 54 minutes for Mr. Alvarez's 45-minute criterion trips some teams; model the conversion once.
- Predicted numbers in the Dani design are wishful (felt pouch but a 7-minute start). Insist that every prediction names the two prototypes it sits between.
- Teams treat P5 as a 'bad' prototype to ignore rather than as the clearest evidence of what the salt does.

Structured debrief

1. Fill the class decision wall together: P3 for Maya, P1 for Mr. Alvarez, nothing for Dani. Where a team passed P2 for anyone, point at the 53 °C peak; where a team passed P4 for Maya, point at 80 g and \$2.05. The matrix, not the opinion, settles it.

2. Walk the feature-comparison pairs on the board: doubling the iron (P1 to P2) raised the peak by 9 °C and more than doubled the hours; switching to felt (P2 to P3) cut the peak by 11 °C, slowed the start from 7 to 45 minutes, and nearly doubled the hours above 35 °C; removing the salt (P3 to P5) killed the warm-up entirely. Name the pattern: iron sets how much energy can be released, the pouch and the salt set how fast it is released.
3. Collect the Dani designs. Most will combine a felt body with a thin-cotton window or a two-layer pouch and about 45 to 50 g of iron (60 g in felt breaks Dani's 70 g mass and \$2.00 cost limits). Ask for the predictions and check that each sits between two real prototypes; a predicted 7-minute start with a felt pouch is not bracketed by anything and should be challenged.
4. Energy check: a pouch warms when energy is released by the reaction faster than it leaks to the surroundings, and cools when the leak wins. P5 shows a reaction that never released energy fast enough to get ahead of the leak. No prototype created energy; it was stored in the arrangement of the iron and oxygen atoms before they reacted, and the reaction released it to the pouch and its surroundings.
5. Close by asking which lab the class trusts more, their own three prototypes in Lab 2 or this report of six tested once each, and why one test per design is the weakness q9 targets.

Recording sheet: what students fill in

Lab 3 recording sheet: The decision meeting

You are the engineering review board at a company that makes single-use iron-powder hand warmers. The lab team tested six prototypes and sent the report below. Three customers have sent briefs. Your job: match a prototype to each customer using the data, and where nothing fits, design something that will. The numbers in this report are a teaching dataset invented for this lab; they follow the patterns of real hand warmers but are not measurements of real products.

How the prototypes work and how they were tested — Each pouch holds iron powder; some also hold 2 g of salt, which speeds up the reaction of iron with oxygen. When the wrapper is opened, air reaches the iron through the pouch fabric and the reaction releases energy to the pouch and whatever is touching it. Thin cotton lets air in quickly; thick felt lets it in slowly. Every prototype was tested once in a 22 °C room in still air with a probe inside the pouch. Costs use the company's price list: iron \$0.025 per gram, thin cotton pouch \$0.10, thick felt pouch \$0.45, salt \$0.10.

Test report: six prototypes (simulated teaching data)

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

Prototype	Iron (g)	Pouch	Salt	Peak temp (°C)	Minutes to peak	Minutes to reach 40 °C	Hours above 40 °C	Hours above 35 °C	Mass (g)	Cost (\$)
P1	20	thin cotton	yes	44	15	8	0.9	1.5	30	0.7
P2	40	thin cotton	yes	53	20	7	2.6	3.9	50	1.2
P3	40	thick felt	yes	42	90	45	4.4	7.1	60	1.55
P4	60	thick felt	yes	45	80	35	7.5	10.2	80	2.05
P5	40	thick felt	no	33	120	never	0	0	58	1.45
P6	40	thin cotton	no	41	40	30	0.7	2.9	48	1.1

Simulated data. Room temperature 22 °C. 'Hours above' counts the time the probe read above that temperature.

Temperature inside the pouch over time for P1, P2 and P3 (simulated teaching data)

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

Time (min)	P1 (°C)	P2 (°C)	P3 (°C)
0	22	22	22
5	34	36	25
10	43	46	28
15	44	51	31
20	44	53	33
30	43	52	37
45	42	50	40
60	40	48	41
90	36	45	42
120	31	43	42
180	24	39	42
240	22	35	41

Time (min)	P1 (°C)	P2 (°C)	P3 (°C)
300	22	31	40

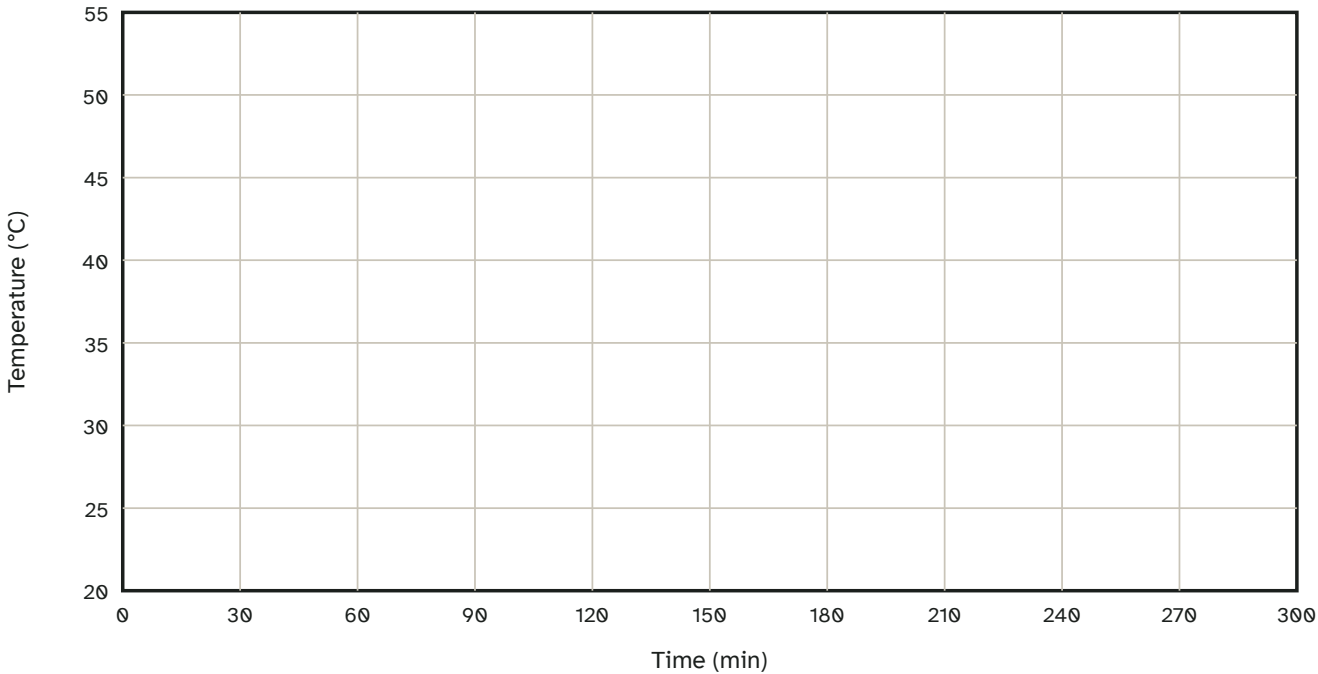
Simulated data. P3 stayed above 40 °C until about 310 minutes and above 35 °C until about 450 minutes (not shown).

Customer briefs — Customer A, Maya, three-day backpacking trip: criteria: at least 6 hours above 35 °C from one pack. Constraints: mass 60 g or less; cost \$1.75 or less; peak never above 50 °C.

Customer B, Mr. Alvarez, 25-minute bus wait every winter morning: criteria: reaches 40 °C within 10 minutes; at least 45 minutes above 40 °C. Constraints: cost \$1.00 or less; peak never above 50 °C.

Customer C, Dani, wildlife photographer at dawn in the Sierra: criteria: reaches 40 °C within 15 minutes; at least 3 hours above 40 °C. Constraints: peak never above 48 °C (the pack sits inside a glove against skin); mass 70 g or less; cost \$2.00 or less.

Temperature inside the pouch over time: P1, P2 and P3 (draw lines at 40 °C and 35



Feature comparison: pairs that differ by one feature

Pair	The one feature that changed	Effect on peak (°C)	Effect on minutes to 40 °C	Effect on hours above 35 °C	What this feature controls
P1 → P2					
P2 → P3					
P3 → P4					
P3 → P5					
P2 → P6					

Decision matrix: PASS or FAIL, with the first failing item and its number

Prototype	Customer A (Maya)	Customer B (Mr. Alvarez)	Customer C (Dani)
P1			
P2			
P3			
P4			
P5			
P6			
Our match (or 'none passes')			

New design proposal for the customer with no passing prototype

Design feature or predicted value	Our choice or prediction	Borrowed from or bracketed by which prototypes	Why
Iron mass (g)			
Pouch material or construction			
Salt?			
Predicted peak (°C)			
Predicted minutes to 40 °C			
Predicted hours above 40 °C			
Predicted mass (g)			
Predicted cost (\$)			

1. Customer A gets prototype _____. Customer B gets prototype _____. Customer C gets prototype ____ (or 'none').
Number of prototypes that pass for at least one customer: _____.

Answer: Customer A: P3. Customer B: P1. Customer C: none. Prototypes that pass for at least one customer: 2 (P1 and P3).

Why: P3 is the only prototype with at least 6 h above 35 °C (7.1 h) that also meets Maya's 60 g, \$1.75 and 50 °C limits; P4 lasts longer but weighs 80 g and costs \$2.05. P1 reaches 40 °C in 8 min, holds it 0.9 h (54 min), costs \$0.70 and peaks at 44 °C, so it passes for Mr. Alvarez; P2 is faster and longer but costs \$1.20 and peaks at 53 °C. For Dani, P2 peaks at 53 °C (limit 48) and holds 40 °C only 2.6 h; P3 takes 45 min to reach 40 °C (limit 15); P4 takes 35 min, weighs 80 g and costs \$2.05; P1 and P6 hold 40 °C under 1 h; P5 never reaches 40 °C.

Also acceptable: None; the matrix is determined by the data.

2. For the customer with no passing prototype, which two of their requirements pull against each other in this dataset, and which design feature is responsible for the tension?

Answer: Dani's fast start (40 °C within 15 minutes) and long working time (3 hours above 40 °C) pull against each other, with the 48 °C ceiling closing the easy escape. The pouch material is responsible: thin cotton lets air reach the iron fast, giving a quick start but a short, hot run (P2 starts in 7 min but peaks at 53 °C and lasts 2.6 h), while thick felt slows the air, giving a long run but a slow start (P3 lasts 4.4 h but takes 45 min).

Why: The question checks that students can locate the conflict in the data and attribute it to the single feature that controls the rate of the reaction.

- Names the start-time criterion and the duration criterion as the conflict
- Attributes it to the pouch material (rate of air supply)
- Cites at least one number from P2 or P3

Analysis questions: answer key

3. Calculate the cost per hour above 40 °C for P1, P3 and P4 (cost divided by hours above 40 °C). Which is the best value? Explain why best value alone could not choose a pack for Maya.

Answer: P1: $\$0.70 \div 0.9 \text{ h} \approx \$0.78$ per hour. P3: $\$1.55 \div 4.4 \text{ h} \approx \$0.35$ per hour. P4: $\$2.05 \div 7.5 \text{ h} \approx \$0.27$ per hour. P4 is the best value. Value alone cannot choose for Maya because P4 breaks two of her constraints: it weighs 80 g (limit 60 g) and costs \$2.05 (limit \$1.75). A constraint is a wall; a prototype on the wrong side of it is out no matter how good its value.

Why: Division with units, then the criterion-versus-constraint point. Accept 0.78, 0.35 and 0.27 rounded to two decimals; accept 'about 78 cents an hour'.

Also acceptable: Accept cost per hour above 35 °C if the student states that choice and computes it correctly: P1 \$0.47/h, P3 \$0.22/h, P4 \$0.20/h, same conclusion.

4. Using the temperature-time dataset, calculate the average warming rate in °C per minute over the first 10 minutes for P2 and for P3. These two prototypes hold the same iron and the same salt. Explain what the pouch material changed about the reaction and about how fast energy was released to the pouch.

Answer: P2: $(46 - 22) \text{ °C} \div 10 \text{ min} = 2.4 \text{ °C per minute}$. P3: $(28 - 22) \text{ °C} \div 10 \text{ min} = 0.6 \text{ °C per minute}$. The felt pouch let air, and therefore oxygen, reach the iron much more slowly, so the iron reacted more slowly and released energy to the pouch at about a quarter of the rate. The iron and salt were the same, so the pouch changed the rate of the reaction, not what was reacting.

Why: Rate from the dataset, then the mechanism: the pouch controls the oxygen supply, which controls the reaction rate, which controls how fast energy is released to the surroundings.

Also acceptable: Accept rates computed over 0 to 5 minutes (P2 2.8 °C/min, P3 0.6 °C/min) if the interval is stated.

5. P2 has the highest peak of all six prototypes, yet it fails for every customer. Explain why, using numbers from the briefs. Then compare P2 and P3, which hold the same 40 g of iron: does the felt pouch change how much energy the iron can release in total, or how fast it is released? Use the hours-above data to support your answer.

Answer: P2 fails Maya and Mr. Alvarez on the 50 °C ceiling (peak 53 °C), fails Mr. Alvarez on cost (\$1.20 against \$1.00), fails Maya on hours above 35 °C (3.9 h against 6 h), and fails Dani on the 48 °C ceiling and on hours above 40 °C (2.6 h against 3 h). The hottest pack is useless if it burns and burns out. P2 and P3 hold the same 40 g of iron, so the total energy the iron can release by reacting with oxygen is about the same; the felt changes how fast it is released. P2 released its energy quickly (peak 53 °C, 2.6 h above 40 °C, 3.9 h above 35 °C) while P3 released it slowly (peak 42 °C, 4.4 h above 40 °C, 7.1 h above 35 °C). The felt spread roughly the same energy over a longer time at a lower temperature.

Why: Students should connect peak and duration as two views of one amount of energy released at different rates. Qualitative; no calorimetry. 'About the same total' is the acceptable level of precision.

- Cites specific failures with numbers from the briefs
- States that the same iron means about the same total energy can be released
- Explains that the pouch changes the rate, using peak and hours-above data

6. P5 never reached 40 °C and its peak was only 33 °C, even though it holds as much iron as P3. What does comparing P3 with P5 tell you about the salt? Explain the result with energy language: why does a reaction that releases energy slowly fail to warm the pouch much, even though the total energy released may be similar?

Answer: P3 and P5 differ only in salt, and P5 never reached 40 °C (peak 33 °C after 2 hours) while P3 reached 42 °C and stayed above 40 °C for 4.4 h, so the salt makes the iron react with oxygen much faster. A pouch warms only when the reaction releases energy faster than the pouch loses energy to the surrounding air. In P5 the reaction was so slow that nearly all the energy leaked away to the room as fast as it was released, so the pouch barely rose above room temperature even though, over many hours, a similar total might eventually have been released. Temperature depends on the race between release and loss, not on the total alone.

Why: This is the chapter's rate-versus-amount idea in energy language, and it is why the lesson's engineers cared about pouch material and additives rather than iron alone. Accept 'catalyst' or 'speeds it up' without requiring the word.

- Identifies salt as the feature that speeds the reaction, from the P3-P5 pair
- Explains that temperature rises only when release outpaces loss to the surroundings
- No claim that energy was created or that P5 had less energy available

7. Describe the new design your team proposed for Dani. Name the characteristic you borrowed for a fast start and the prototype it came from, the characteristic you borrowed for a long working time and its prototype, and justify each predicted number by naming the two prototypes it sits between.

Answer: A typical strong proposal: 50 g of iron with 2 g of salt in a thick-felt pouch that has a thin-cotton panel (about a quarter of one face) so part of the iron gets air quickly for the start while the rest reacts slowly for the long run. Fast-start characteristic: thin cotton, from P1 and P2 (start 7 to 8 min). Long-life characteristic: thick felt, from P3 and P4 (4.4 to 7.5 h above 40 °C). Predictions and brackets: start about 10 to 15 min (between P2's 7 and P3's 45, and inside Dani's 15-minute criterion because the cotton panel lets part of the iron start fast); peak about 45 to 47 °C (between P3's 42 and P2's 53, kept under 48 by the mostly felt pouch); 4 to 6 h above 40 °C (between P3's 4.4 and P4's 7.5, since iron is between 40 and 60 g); mass about 70 g (50 g iron + 18 g felt pouch + 2 g salt, between P3's 60 and P4's 80); cost about \$1.85 (50 × \$0.025 = \$1.25, felt \$0.45, salt \$0.10, panel a few cents; between P3's \$1.55 and P4's \$2.05). Every value meets Dani's criteria and stays inside her constraints.

Why: Full credit requires both borrowed characteristics with their source prototypes and a bracket for every prediction. Other designs earn credit (for example 45 g of iron with a larger cotton panel) if the bracketing is honest and every criterion and constraint holds. A 60 g design fails in a felt pouch (60 + 18 + 2 = 80 g and \$2.05), and in an all-cotton pouch it would run hotter than P2's 53 °C, so expect iron at 50 g or less.

- Names a fast-start characteristic and its prototype
- Names a long-life characteristic and its prototype
- Every predicted value bracketed by two named prototypes
- All predictions meet Dani's criteria (40 °C within 15 min, at least 3 h above 40 °C) and constraints (peak 48 °C or less, mass 70 g or less, cost \$2.00 or less)

8. A classmate says the meeting was a waste of time: 'Just give everyone the hottest one that is under budget.' Explain, with one example from the matrix, why a pass-or-fail check against every criterion and constraint gives a more trustworthy decision than picking by a single number.

Answer: Picking by one number ignores every other requirement. Example: the hottest pack under \$1.00 is P1 (44 °C), which would be shipped to Maya and fail her 6-hour criterion by more than 4 hours; or the hottest pack overall, P2, breaks a 50 °C safety ceiling for two customers and the cost limit for one. A pass-or-fail check against every criterion and constraint catches each failure with a number, treats constraints as walls rather than scores to trade away, and gives the same answer whoever runs it. That is what makes the decision trustworthy and defensible to the customer.

Why: This is the ETS1-2 idea of a systematic process. Any correct matrix example earns credit.

- Gives one concrete example where a single-number choice fails a requirement
- Explains that every criterion and constraint is checked and constraints are not traded
- Mentions that the result is repeatable or defensible

9. Every prototype was tested once, in a 22 °C room with still air. Dani will use hers at dawn in the mountains, perhaps near 0 °C with wind. Predict how colder surroundings and moving air would each change a pouch's peak and working time, say which customer's match you are least sure of, and describe the one retest you would order before shipping.

Answer: Colder surroundings make the pouch lose energy faster (a bigger temperature difference between pouch and air), so the peak would be lower and the working time shorter. Moving air brings oxygen to the iron faster and also carries energy away faster, so a thin-cotton pouch might start faster and run hotter but burn out sooner, and a felt pouch would mainly lose energy faster. The least certain match is Dani's (or any mountain use): her design was predicted, not tested, and the test room was nothing like dawn in the Sierra; Maya's P3 is also uncertain for the same reason. The retest to order: build the proposed Dani design and test it, at least three times, in a cold chamber or outdoors near 0 °C with a fan, recording the time to 40 °C, the peak and the hours above 40 °C, and compare with the 22 °C result. Single tests at room temperature should never be the basis for shipping.

Why: Limits of the data: direction of change in cold and wind, which match is least certain, and a specific retest with repetition and realistic conditions.

- Predicts lower peak and shorter time in cold with a reason (faster loss to surroundings)
- Predicts the effect of wind on oxygen supply and on energy loss
- Names the least certain match with a reason
- Describes a retest with repetition under realistic conditions

Individual evidence of learning

Each student submits their own graph with the 40 °C and 35 °C lines, the feature-comparison table, the decision matrix with a number beside every fail, the new design proposal and individual answers to q1 through q9. The feature-comparison table together with q7 is the primary evidence for MS-ETS1-3; the matrix and q8 are the evidence for systematic evaluation.

Assessment criteria

Criterion	Proficient looks like
-----------	-----------------------

Systematic evaluation	Every prototype is marked PASS or FAIL for every customer with the first failing criterion or constraint and its number; matches are P3 for A, P1 for B, none for C.
Cause and effect from one-feature pairs	All five pairs are completed with the correct feature and the direction and size of its effect on peak, start time and hours, and each feature is assigned to what it controls (amount versus rate).
Combining best characteristics	The new design names a fast-start characteristic and a long-life characteristic with the prototypes they came from, and every predicted value is bracketed by two existing prototypes and stays inside Dani's constraints.
Energy reasoning	q5 and q6 explain the results as energy released by the reaction to the pouch at different rates, with the total set by the iron, and never describe energy as created or used up.
Limits of the data	q9 identifies single tests and the 22 °C still-air room as limits, predicts the direction of change in cold and wind, and proposes a specific retest.

Accessibility adaptations

- Provide the graph with axes pre-scaled and the 40 °C and 35 °C lines pre-drawn; allow the data analyst role to be shared so no one student plots all three curves.
- Offer index cards for the six prototypes so the matrix can be built by physically sorting cards into pass and fail piles per customer before writing.
- Supply a bilingual card for criterion, constraint, peak, prototype, rate and release, and a sentence frame for matrix entries: 'FAIL: ___ was ___, limit ___.'
- Accept q5, q7 and q9 as recorded spoken answers or as annotated sketches of the pouch; reduce the task to three prototypes and one customer for students who need it, keeping P2, P3 and the Dani brief so the combination step survives.
- For low vision, print the datasets at large size and provide a tactile or high-contrast graph; no chemicals are involved, so every student can take every role.

Low-cost alternative

The lab already costs nothing beyond printing. With no printer, write the two datasets and the three briefs on the board and have teams copy the matrix onto notebook paper; the graph can be drawn on grid paper.

Model limitations

The test report is an invented teaching dataset that follows the patterns of real iron-powder hand warmers (more iron, higher and longer; slower air supply, cooler and longer; salt needed for a usable rate) but does not describe any real product, and the price list is invented. Each prototype was 'tested' once in still 22 °C air, which hides the variation real tests show and says nothing about cold, wind or a hand drawing energy from the pouch. The cause-and-effect reading relies on pairs that differ by one feature, which real test programs arrange deliberately but which this report hands to students ready-made. The combined design is predicted, not tested; a real board would order the prototype built and tested before deciding. This activity has not been classroom-tested by the authors; the expected matrix and the clusters of student designs are projections from the dataset.

Teacher notes

The matrix is fully determined by the data: P3 for Maya, P1 for Mr. Alvarez, none for Dani. Grade the matrix for the number beside every fail, not just the verdict; the lab's ETS1-3 evidence is the feature-comparison table and the bracketed predictions in q7. Expect the most argument over Mr. Alvarez's 45-minute criterion (P1's 0.9 h is 54 minutes, a pass) and over P4 for Maya (longest-lasting, but 80 g and \$2.05 fail two constraints). Watch for designs that predict a felt pouch starting in under 10 minutes with nothing to bracket it. The energy questions (q5, q6) are where the lesson's content lives: iron sets the amount of energy that can be released, pouch and salt set the rate, and the pouch's temperature is the race between release and loss to the surroundings.