

FUNDAMENTALS OF AI-AUGMENTED ENGINEERING

CHAPTER 1/LAB-1: Engineering Lab (3 Hours)



Ball Drop Fall Time

Can we trust the model, the computer, or the AI – and how would we know?

Taner Onsay, Ph.D. & Sebnem Onsay
Authors

August 2026 · Revision 1.0

© 2026 AI First Engineering

Problem



Model



Compute



Analyze



Validate



Improve

Today you run the entire workflow, end to end, for real

Ball Drop Fall Time

What is this lab about?

Use a simple falling-object experiment to experience the complete engineering workflow:

Predict → Measure → Compute → Compare → Validate → Improve
(DO › OBSERVE › RECORD › TALK TO AI › EVALUATE › DECIDE › DO ...)

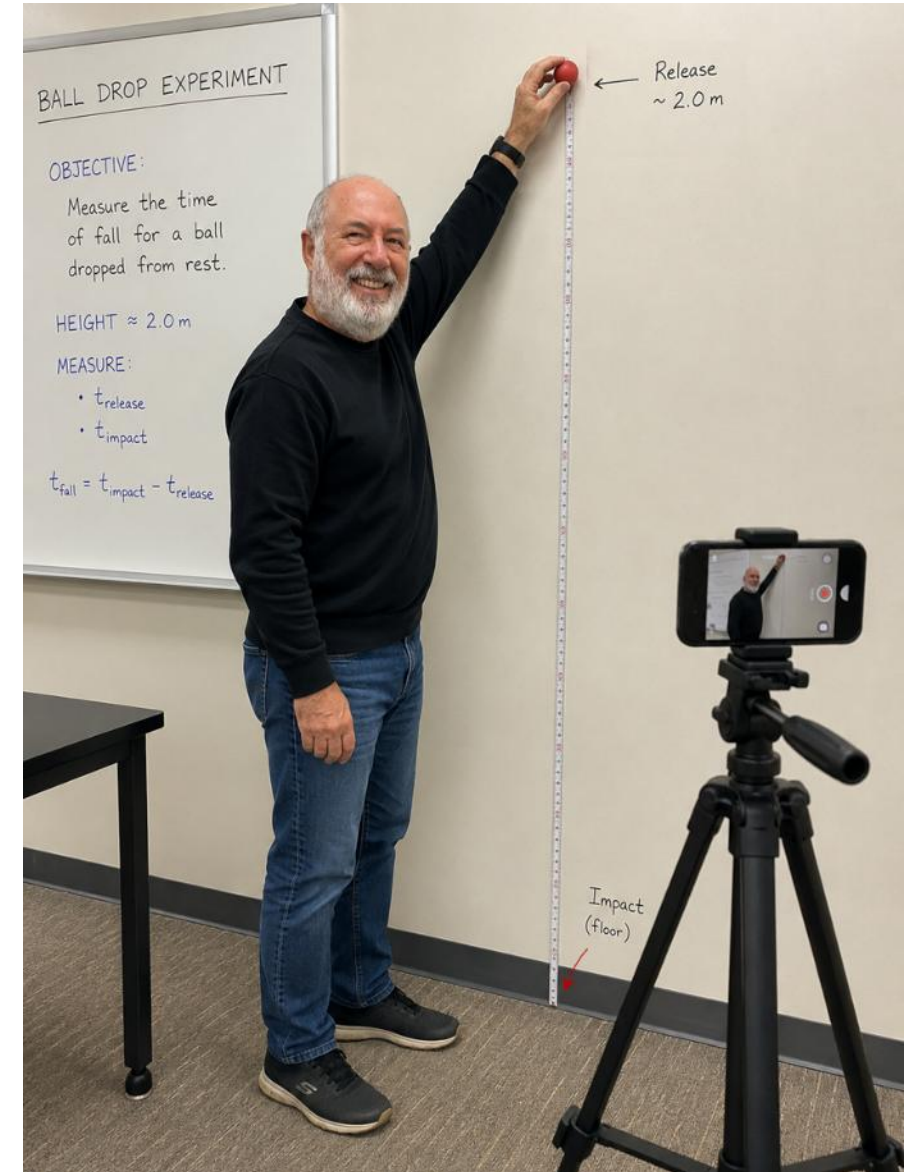
LEARNING OBJECTIVES

By the end of this lab, you will be able to:

- Build and use a simple physical/mathematical model
- Use computation to predict fall time
- Obtain experimental measurements from video
- Compare model predictions with experimental results
- Identify possible sources of disagreement
- Use AI to investigate—independently validate—engineering results

TOOLS YOU WILL USE

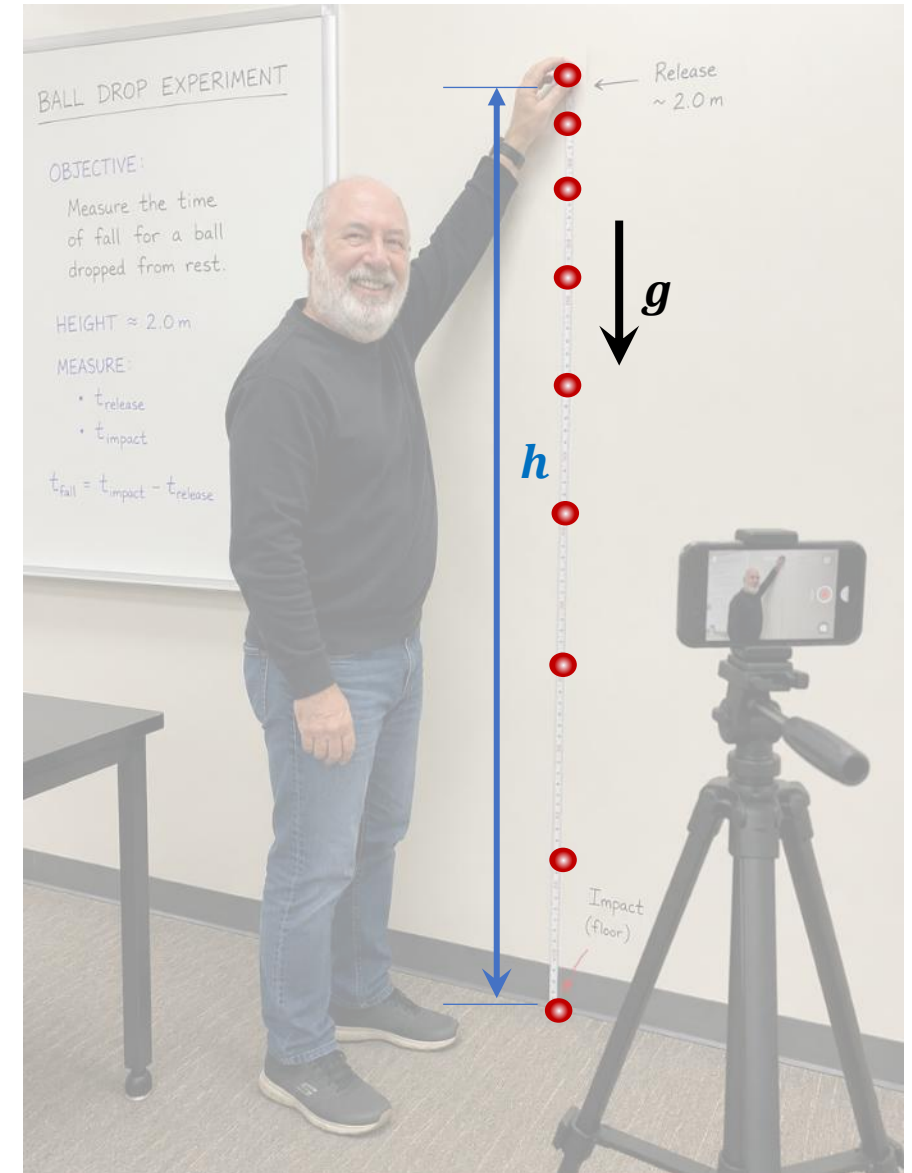
- **Python / Jupyter** — compute the theoretical prediction
- **iPhone/phone video** — measure the physical experiment
- **Tracker** — extract position–time data from video
- **AI assistant** — investigate discrepancies and possible explanations
- **Your engineering judgment** — decide what you trust



What you will do (step-by-step):

- 1 Create your team's AI workspace — the engineering record
- 2 Predict fall times for two heights each — commit first
- 3 Drop and film: 6–8 runs across the team
- 4 Measure each fall time from the video (trim method)
- 5 Analyze at least one video per student with Tracker
- 6 Build the team dataset — every legitimate run
- 7 Plot t vs. h · estimate g from every run
- 8 Converse with AI in stages: evidence › interpretation › explanation
- 9 Challenge the AI — BREAK THE AI — and close the loop
- 10 Assemble the report from the record · review · approve · sign

Would you confidently put your name on your final result? You will — literally — in the report's approval section.



AI joins the team. You lead it.

AI becomes a continuously available member of the engineering team throughout the workflow — while you remain the **lead engineer**.

THE RECURRING LOOP — INSIDE PROBLEM › MODEL › COMPUTE › ANALYZE › VALIDATE › IMPROVE

DO

›

OBSERVE

›

RECORD

›

TALK TO AI

›

EVALUATE

›

DECIDE

... and back to DO — the conversation evolves as the experiment evolves.



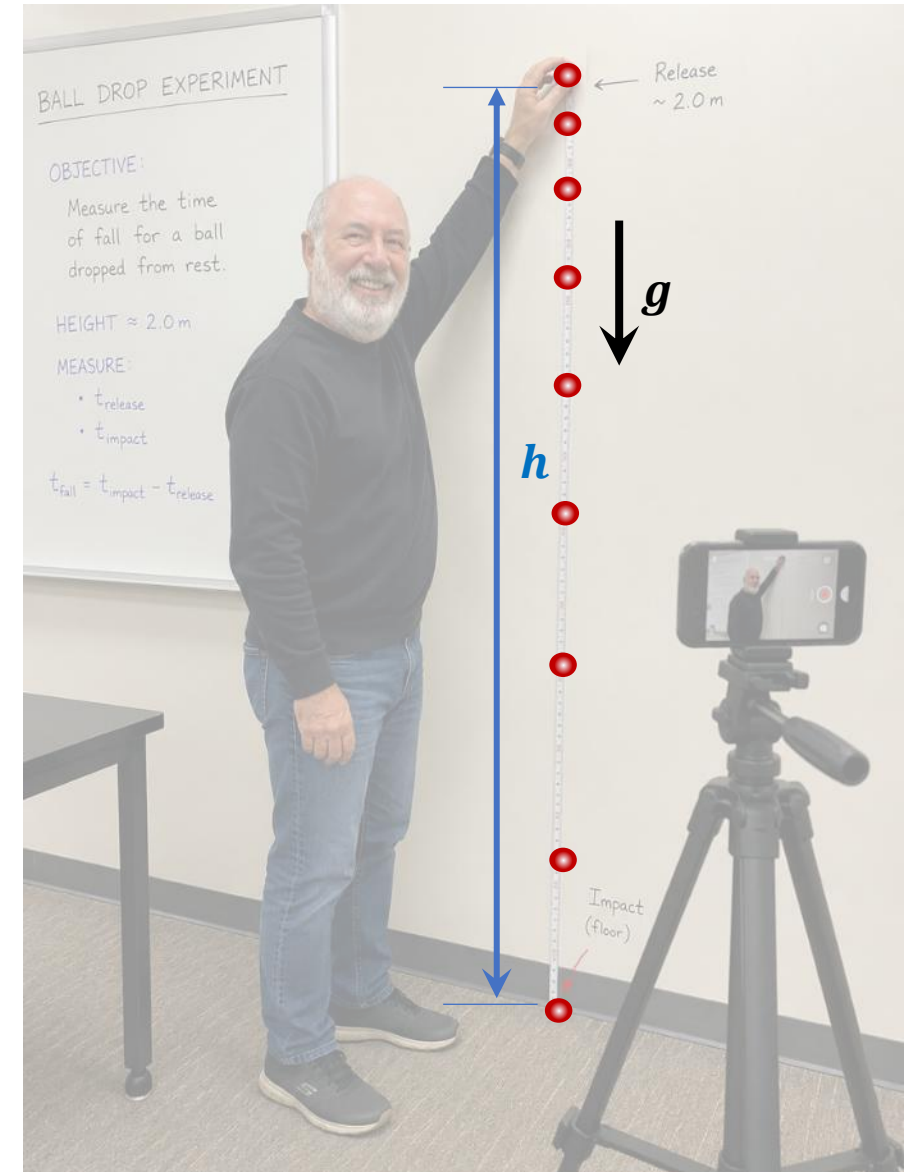
You, not AI: observe, measure, decide what to trust, accept or reject suggestions, approve the result.



AI, for you: plans, questions, calculations, organization, hypotheses — on the evidence you provide.

AI assists. Humans observe, verify, decide, and approve.

Never ask AI to invent what you did. Tell AI what you actually did.





Meet your AI engineering teammate



HUMAN TEAM — LEAD ENGINEER

- Perform the physical experiment
- Make the measurements
- Observe what actually happens
- Record the evidence
- Evaluate physical reasonableness
- Decide what to trust
- Approve the final engineering result



AI TEAMMATE

- Help plan the investigation
- Ask useful questions
- Help organize data
- Assist with calculations and code
- Suggest possible explanations
- Help compare evidence
- Help structure the final report

The AI does not observe your experiment. **You do.**



Create your AI workspace — the engineering record

ONE WORKSPACE · “LAB-1 — BALL DROP LAB”

- Report instructions + template
- Setup photos · object + dimensions
- Assumptions · predictions (committed)
- Python code + results
- Every run: video times, Tracker data, plots
- Observations · questions · failed attempts
- AI responses · team decisions · conclusions

The record documents what happened.

The report communicates what matters.

Feed the record as you go — don't reconstruct it at 2:45.

STARTER PROMPTS

PROMPT-1

“Here are the corresponding lecture notes, lab instructions and the lab report template. Just read them for later reference and wait for my command.” (drop or attach PDF versions of all 3 documents right into this chat)

PROMPT-2

“You are our AI engineering teammate for the LAB-1 Ball Drop Lab. We are the lead engineers. Help us work step by step, and ask for measurements or observations when you need them. Never invent experimental data, observations, photographs, calculations, or conclusions. Distinguish clearly between information we provide, calculations you perform, and hypotheses you suggest. Help us identify assumptions and possible errors — but require us to evaluate them. This workspace is the engineering record for our final report.”

Any capable AI platform works — a project/workspace feature (e.g., ChatGPT Projects) is convenient, but one persistent conversation is enough.



Engineering first, software second

0:00–0:15

**Workspace + tool verification**

AI workspace created, starter prompt in · Python and Tracker proven working.

0:15–0:35

**Part A · Predict**

Two heights per student (1.5–2.0 m); every prediction committed before any drop.

0:35–1:15

**Part B · Experiment — 6–8 runs**

Each student drops from both heights; film every run; log everything.

1:15–1:40

**Part C · Measure twice**

Trim-time each video; Tracker on at least one run per student.

1:40–2:05

**Part D · Team dataset**

Table · t vs. h plot · estimated g per run — AI organizes, evidence first.

2:05–2:30

**Part E · Interpret + BREAK THE AI**

Staged conversation — then grade the AI and answer it.

2:30–2:50

**Validate + assemble the report**

Independence debrief; AI drafts from the record; humans verify.

2:50–3:00

**Human review & approval**

Check every number · sign the result · check-out.



Teams & logistics — everything runs in parallel

EACH TEAM'S KIT — NO SHARED EQUIPMENT, NO WAITING



Tape measure

one per team



Ball to drop

Pick a solid-colored ball. It works better with the Tracker software.



Smartphone with video

at least one per team (determine fps: 30, 60, 120, 240 ?)



Laptop with Python + Jupyter & Tracker

at least one working per team

TEAM RULES

- Mix experience where practical; help everyone participate.
- Rotate roles between drops: dropper · camera · timer/analyst · recorder
- Everyone runs the code at least once — the notebook is a team product, the skill is personal
- Claim a drop station and stay clear of other teams' drop zones

EXPERIMENTAL DESIGN: each student, two measured heights in 1.5–2.0 m → a 6–8-run team dataset.



Workspace, Python, Tracker — prove all three



AI WORKSPACE

- ☐ Workspace created + named
- ☐ Starter prompt pasted
- ☐ Lab instructions uploaded



PYTHON

- ☐ Fall-time notebook runs
- ☐ Change `h` › output changes
- ☐ Result lands near 0.6 s



TRACKER

- ☐ Practice video imports
- ☐ Scale + axes set
- ☐ y - t plot appears



If something is broken

- Ask your instructor/lab facilitator for help — don't burn 40 minutes alone
- Paste the exact error into AI and ask what it means (often faster than you'd think)
- Keep your team moving on one working machine meanwhile

*All nine green? Start Part A early — the schedule rewards prepared teams.
Experienced with another IDE? Fine — as long as you can run and modify a notebook and support yourself.*



Commit before you drop

1

Choose & measure TWO heights each

Each student: two heights in 1.5–2.0 m, different enough to matter. Record each value AND how you measured it.

2

Predict every fall time

$t = \sqrt{2h/g}$. Use a calculator, Excel, Python, or AI assistance — your choice of tool.

3

Commit all predictions NOW

All 6–8 predictions go into the engineering record, with inputs, before the first drop. No revising later.



TALK TO YOUR AI TEAMMATE “Here are our heights and committed predictions — log them. Ask us anything we’ve left ambiguous.”



ENGINEER DECIDES the numbers you commit are yours — AI logs them, you own them.

WHY PREDICT FIRST?

A prediction made after seeing the data isn't a prediction — it's a story. Committing first is what makes the comparison in Part C honest, and it's how real engineering validation works.

Also record: object chosen, who measured, and your model assumptions.

Problem ›

Model ›

Compute ›

Analyze ›

Validate ›

Improve



Make reality produce a number



Film the drop

Smartphone video, steady frame, release point and floor both visible. Do several drops.



Log every run as it happens

Run #, student, height, phone + fps, anything odd. Straight into the record — notes are data. Timing comes next; don't bottleneck the drop zone.



6–8 runs — the full matrix

Each student drops from BOTH of their heights. Record every legitimate run — an inconvenient value is still a value.

QUESTIONS TO ARGUE ABOUT

- Keep the phone fixed between a student's two drops
- Release point and floor both in frame · known-length object visible
- Which instant is “release”? Which is “impact”? Agree as a team
- Do repeated drops at the same height agree with each other?

Problem ›

Model ›

Compute ›

Analyze ›

Validate ›

Improve



Manual timing — trim to the truth



Open the run's video in Photos

(or your phone's equivalent editor)



Tap Edit › drag the trim handle to the
RELEASE instant

read the displayed *time* → t_{release}



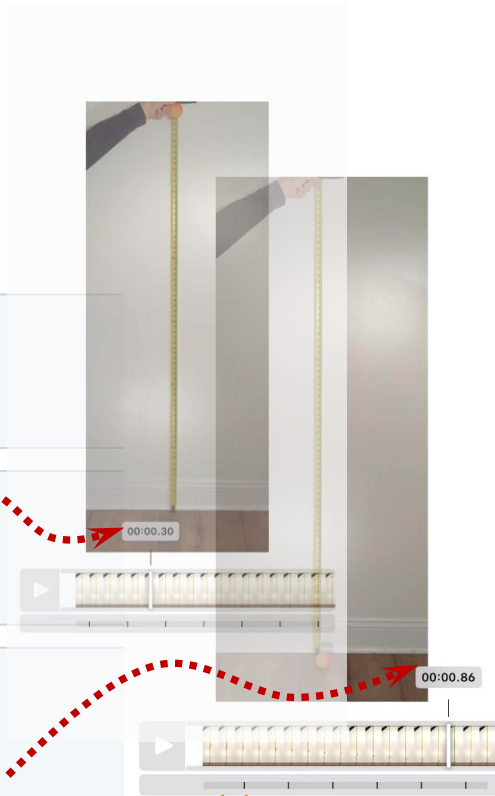
Drag to the **IMPACT** instant

read the displayed *time* → t_{impact}



Record both times in the engineering record

one line per run — all 6–8 runs



HOW SHARP IS THIS KNIFE?

- The editor steps in frames — your resolution is the frame rate.
- ± 1 frame at 30 fps $\approx \pm 0.033$ s · at 60 fps $\approx \pm 0.017$ s · at 240 fps $\approx \pm 0.004$ s
- At both ends of a ~ 0.6 s fall, that is several percent before anything else goes wrong.
- Release/impact ambiguity is real measurement uncertainty — note where the team disagreed.

Keep this yardstick — Part D judges agreement against it.

$$t_{\text{fall}} = t_{\text{impact}} - t_{\text{release}} \quad \text{Example: } t_{\text{fall}} = 00:00.86 - 00:00.30 = 0.5 \text{ (s)}$$

Problem



Model



Compute



Analyze



Validate



Improve



Tracker — same video, second method

1 Set the scale

Use the known-length object in frame.

2 Set the coordinate system

Origin and axes — y along the fall.

3 Track the object

Click it frame by frame through the drop.

4 Get $y(t)$

Open the data table — time, x, y.

5 Estimate the fall time

First motion to impact, from the data.

NOT IN LAB-1

- Formal Curve fitting theory
- Velocity / acceleration extraction
- Uncertainty propagation
- Advanced Tracker features

They arrive when a problem needs them.

Aim: at least one Tracker analysis per student.

LAB-1 competency — “I can turn a video of motion into position/time data.”

Problem



Model



Compute



Analyze



Validate



Improve



Six to eight runs — one table, no discards

Run	Student / dropper	Height h (m)	Predicted t (s)	Manual t (s)	Tracker t (s)	Manual % diff	Tracker % diff	Manual g_est (m/s ²)	Tracker g_est (m/s ²)
1	John	1.55	0.562	0.58	0.57	+3.20%	+1.42%	9.22	9.54
2	John	1.95	0.630	0.65	0.62	+3.17%	-1.59%	9.23	10.15
3	Maya	1.52	0.557	0.57	0.58	+2.33%	+4.13%	9.36	9.04
4	...	...	...	...	...	...	...	...	...

Every legitimate run — two heights per student, 6–8 runs. Do not discard inconvenient measurements; a run may be excluded only for a documented procedural failure, noted in the table.

$$\% \text{ diff} = (t_{\text{measured}} - t_{\text{predicted}}) / t_{\text{predicted}} \times 100$$

$$g_{\text{est}} = 2h / t^2$$

Two measurement methods are used: Manual (iPhone video) and Tracker. Keep the sign of % diff; measured values may fall above or below the prediction.

Record ALL legitimate runs. An inconvenient measurement is still a measurement — deleting it is how bias is born.



TALK TO YOUR AI TEAMMATE “Here is our full dataset. Organize it, compute % differences and g per run, and prepare plots. Do NOT explain any discrepancy yet.”



ENGINEER DECIDES recompute one run by hand — if AI's arithmetic fails there, nothing downstream is trusted.

Evidence first → explanation second. The interpretation stages come after the table and plots exist.

Problem



Model



Compute



Analyze



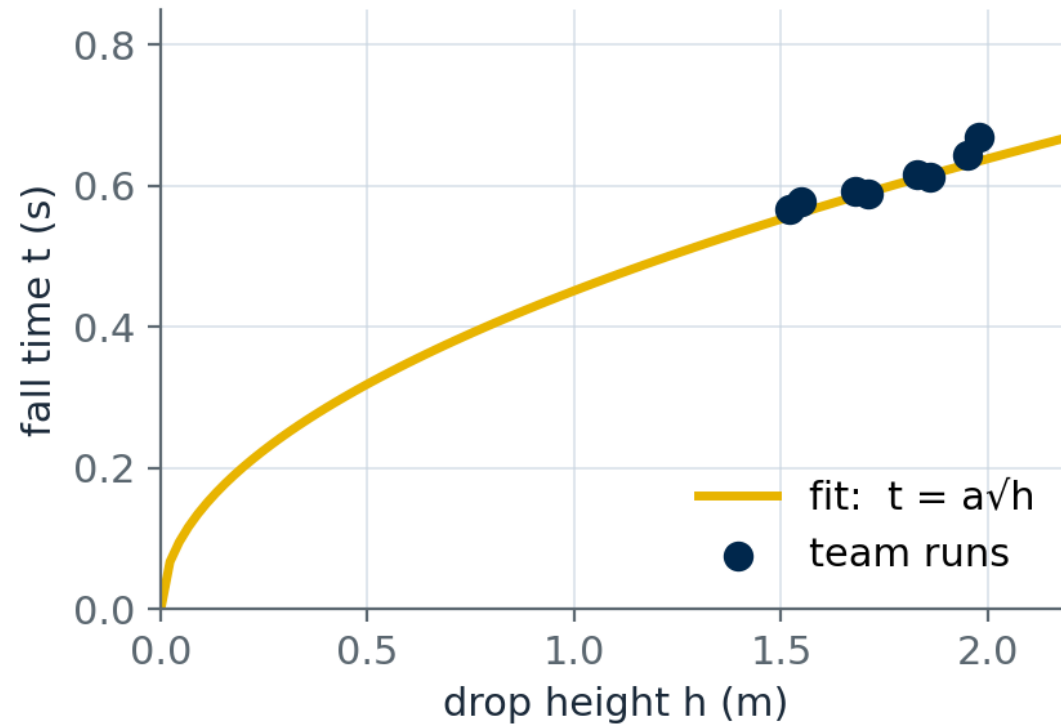
Validate



Improve



Fall time vs. height — does it look linear?



BEFORE any fit — vote as a team:

Does t vs. h look like a straight line? Commit, then let AI fit a curve.

The physics says: $t = \sqrt{2h/g}$ — so t grows like $\sqrt{h}$
Doubling h multiplies t by $\sqrt{2} \approx 1.41$ — not 2.

Does the curve fit make physical sense?



TALK TO YOUR AI TEAMMATE “Plot t vs. h from our table and fit a curve. What shape fits best?”

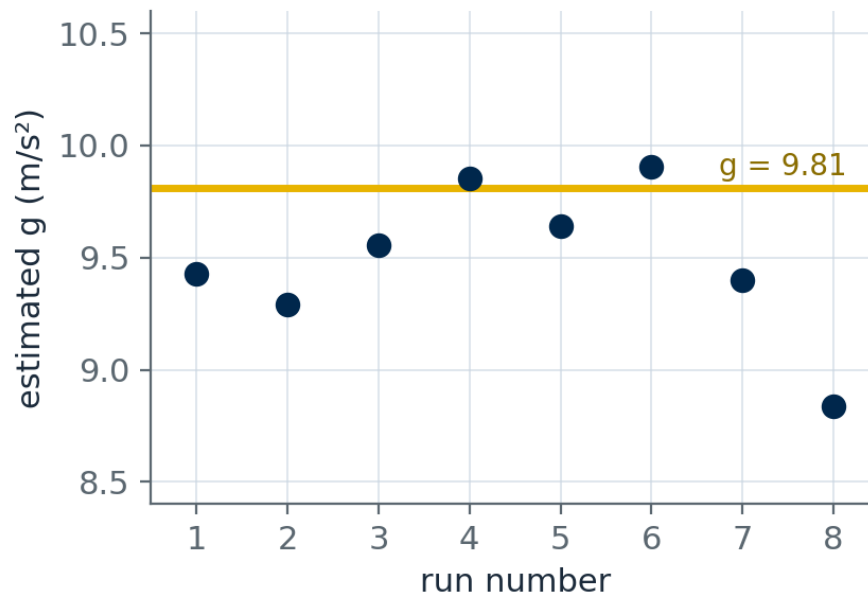


ENGINEER DECIDES compare the fitted shape with $\sqrt{h}$ — the physics, not the fit, is the referee.



Every run votes: what is g?

$$g_{est} = 2h / t^2 \quad \text{for every run}$$



READ YOUR PLOT LIKE AN ENGINEER

- Scatter — how spread out are the eight votes?
- Bias — do they sit mostly above or mostly below 9.81?
- Outliers — does one run disagree with its siblings? Check its notes first.
- Systematic vs. random — same direction every time, or noise?
- What could push every estimate the same way? (late release read, height bias, ...)
- Does the pattern suggest a measurement problem — or a model problem?



TALK TO YOUR AI TEAMMATE “Compute g for each run and plot it against run number with a 9.81 reference line.”



ENGINEER DECIDES the pattern is yours to read — scatter, bias, and outliers are engineering judgments.



Evidence first → explanation second



1 · EVIDENCE

Hand over the dataset. Ask AI to organize, compute, plot, and flag patterns.

“Do not explain the discrepancy yet.”



2 · INTERPRETATION

What agrees? What disagrees?
Scatter or bias? Is t(h) consistent with the model's shape?

Still no causes — just what the data shows.



3 · EXPLANATION

Only now: possible explanations, ranked by likely importance — with the evidence each would need.

Then Part E: grade every suggestion.

Why the discipline? An AI asked “why is our experiment wrong?” will happily explain a discrepancy that doesn't exist — or invent one. Make it earn the explanation from the evidence.

Same rule for humans, by the way — the data speaks before the theory does.



Why aren't the values identical?

Classify each candidate cause using the lecture's error taxonomy — then rank the top three for YOUR data:

Height measurement

Measurement

Video frame rate limits

Measurement

Release timing / hesitation

Measurement

Object not released cleanly (spin, push)

Assumption

Air resistance

Model / Assumption

Constant-g free-fall model itself

Model

Rounding in the computation

Numerical

A mistake in the code

Programming

Frame-counting / analysis technique

Measurement

A possible source $\neq$ a source large enough to explain YOUR discrepancy. For your top three: estimate the size — don't vibe.

Problem

›

Model

›

Compute

›

Analyze

›

Validate

›

Improve



Part E · 2:05–2:30 · Ask the AI — then grade it



Ask

Stage 3, with everything: heights, predictions, manual + Tracker times, g estimates. Ask for ranked explanations AND the evidence that would support or reject each.



Evaluate every suggestion

Sort each AI-proposed cause into one of the four bins ›



Answer the AI — close the loop

“We evaluated your suggestions — here is what we accept, reject, or remain unsure about.” Your verdicts go in the record.

SORT EACH AI SUGGESTION:



Reasonable



Irrelevant here



Physically insignificant



Potentially incorrect

BREAK THE AI: find one AI suggestion that is possible but too small to explain YOUR discrepancy. Possible ≠ important — and copying the AI's answer is not the assignment; judging it is.



Independence has levels — what do two checks share?



Analytical derivation from $h = \frac{1}{2}gt^2$

INDEPENDENT

Algebra by hand vs. code — independent of the implementation, same physical model.



Physical experiment vs. theoretical model

INDEPENDENT

Reality vs. model — a fundamentally different source of evidence.



Repeated drops · more students, more heights

MORE EVIDENCE

Additional experimental evidence — that's why the dataset has 6–8 runs.



Known physical behavior / benchmark

INDEPENDENT

Trusted reference outside your own computation.



Manual timing vs. Tracker — same video

PARTLY — SAME VIDEO

Different methods, one trial: shared fps, perspective, scale, release ambiguity.



Python vs. Excel — same equation

IMPLEMENTATION CHECK

Two calculators, one model — both can be wrong together.



A second AI confirming the first

EVIDENCE, NOT PROOF

Two AIs can repeat the same model choice — and the same error.

Independence is not binary. Always ask: what assumptions, data, or methods do two checks share?



AI drafts from your evidence — you make it true

FINAL PROMPT

“Using the report template and ONLY the evidence in this workspace, help us prepare our LAB-1 Engineering Report — our actual measurements, calculations, photos, Tracker results, plots, observations, decisions, and conclusions. Do not invent or infer missing experimental information. Flag anything that needs human input or verification.”

Concise evidence + engineering reasoning beats pages of polished prose.

Tables, plots, photos, equations, short observations. AI carries the clerical load so you can spend the effort thinking.

THE DRAFT IS NOT THE SUBMISSION — VERIFY:

- ☐ Every numerical result — against the record
- ☐ Every plot — axes, points, the 9.81 line
- ☐ Every caption — says what the figure shows
- ☐ Photos — actually yours, actually this setup
- ☐ Remove any statement your evidence doesn't support
- ☐ Correct AI errors · add missing human observations
- ☐ Conclusions — match what the data shows, no more

AI may assist in writing the report. AI may not assume engineering responsibility.



Human review & approval

“We have reviewed this report against our original experimental record — measurements, calculations, observations, plots, and supporting evidence. We accept responsibility for the submitted engineering work.”

Human reviewed and approved.

EVERY TEAM MEMBER SIGNS
(Create a digital signature copy paste it)

The recurring question, now literal: ***Would you confidently put your name on this result?***

AI assists. Humans observe. Humans verify. Humans decide. Humans approve.

Problem ›

Model ›

Compute ›

Analyze ›

Validate ›

Improve



One concise report — built from your record



FRAME IT

- 1 Problem / objective
- 2 Physical model
- 3 Assumptions
- 4 Experimental setup + photo
- 5 Prediction / computation



MEASURE IT

- 6 Experimental data —
6-8-run table
- 7 Manual video analysis
- 8 Tracker analysis + $y(t)$



ANALYZE IT

- 9 Fall time vs. height + fit
- 10 Experimental estimate of g
- 11 Model–Exper. comparison
- 12 Error / discrepancy analysis



OWN IT

- 13 AI contribution
- 14 Human verification & judgment
- 15 Conclusions
- 16 Improve
- 17 Human review & approval — signed

A few honest sentences per section. Tables, plots, and photos carry the evidence — prose carries the judgment.

Problem



Model



Compute



Analyze



Validate



Improve

WHAT JUST HAPPENED

You predicted something physical,
measured what actually happened,
used computation and AI,
**found differences — and figured out why — and signed
your name to the result.**

That loop repeats throughout this book. Everything from here — numerical methods, data, simulation, optimization, agentic AI — is this same loop with bigger problems.

Problem

Model

Compute

Analyze

Validate

Improve

RUBRIC for the Lab Report

FUNDAMENTALS OF AI-AUGMENTED ENGINEERING

CHAPTER 1 / LAB-1

Engineering Report Rubric

Ball Drop Fall Time — Main LAB-1 Team Report

Suggested Assessment · 40 points

Used by the instructor and, where applicable, by an AI grading agent. Assessment rewards experimental evidence, quantitative analysis, engineering reasoning, validation, and human responsibility.

Criterion	What full credit looks like	Pts
Problem, model, assumptions & predictions	Clear objective; physical model and assumptions stated; measured heights and committed predictions documented before the drops; calculations are reasonable.	5
Experimental setup, procedure & evidence	Actual setup/photo and procedure documented; height and video conditions identifiable; observations and team participation evident.	5
6–8 run dataset & manual video measurements	All legitimate runs included; table contains required run information, predicted and manual fall times; release/impact timing method documented.	6
Tracker analysis & position–time evidence	Required Tracker work is shown; scale/axes and tracking are appropriate; $y(t)$ evidence and Tracker fall time are included and compared with manual timing.	5

Continued

Problem

>

Model

>

Compute

>

Analyze

>

Validate

>

Improve

RUBRIC for the Lab Report

Quantitative analysis: t vs. h, fit & g	Calculations complete; t vs. h plot labeled and interpreted; g estimated from runs and compared with 9.81 m/s ² ; numerical work is consistent.	7
Model–experiment comparison & discrepancy analysis	Predictions and measurements compared quantitatively; scatter, bias, outliers, resolution, and likely error sources evaluated using evidence.	5
AI use & critical evaluation	AI contribution summarized; team evaluates suggestions and independently verifies what is reasonable, irrelevant, insignificant, or potentially incorrect.	3
Engineering conclusions, improvement & judgment	Conclusions follow from evidence; team states what it trusts and why; improvements address observed limitations appropriately.	3
Completeness, clarity & signed approval	Report is concise, organized, complete, and readable; figures/tables are labeled; required human review and approval <u>is</u> signed.	1
TOTAL		40

Problem

›

Model

›

Compute

›

Analyze

›

Validate

›

Improve

CHAPTER - 1

Lab 1 · ≈ 25 MIN

Lab Extension Challenge:

(Individual Assignment)

Students can do it at their convenience (individually) but remind them it is part of the Lab-1 assessment.

Problem



Model



Compute



Analyze



Validate



Improve



Four Balls — One Height

4 balls dropped from the SAME height,
 $h = 73 \text{ in} = 1854 \text{ mm}$.

Which will hit the ground first?
Why?

An extension of the Chapter1-LAB-1 Ball Drop Lab. You already built the model, the program, and the measurement method — now four objects put them to the test.

Small Bounce Ball (SBB)	Small Steel Ball (SSB)	Large Bounce Ball (LBB)	Standard Tennis Ball (STB)
			
Mass $m=15$ grams Diameter $D=29$ mm	Mass $m=67$ grams Diameter $D=29$ mm	Mass $m=90$ grams Diameter $D=57$ mm	Mass $m=57$ grams Diameter $D=66$ mm
Drop Height $h = 73 \text{ in} = 1854 \text{ mm}$			

Problem

>

Model

>

Compute

>

Analyze

>

Validate

>

Improve



An assessment mini-challenge — not another lab report

1

Designed as part of the Lab-1 assessment

The rubric is on the last slide of this section.

2

During lab or on your own

Complete the analysis in the laboratory period or outside it — your choice. The videos and this deck are all you need.

3

Submit a separate short report

Download and use [Chp-1-Lab1-LabExtension-Challenge-FourBallDrop-Report-Template.doc](#)

THE ENGINEERING QUESTION

Can our measurements distinguish the four cases — and does the experimental evidence justify improving our model?

Same workflow as the whole lab:

**Predict → Observe → Measure →
Compare → Question → Ask AI →
Validate → Decide**



ENGINEER DECIDES *the question is not which ball wins — it is what your evidence can support. AI assists. Humans observe, verify, decide, and approve.*

Problem

>

Model

>

Compute

>

Analyze

>

Validate

>

Improve



Predict before you peek

1

Same time — or measurably different?

Do you expect all four balls to reach the ground at essentially the same time, or do you expect measurable differences?

2

If one arrives first, which one?

Name your ball: SBB, SSB, LBB, or STB.

3

Why?

Briefly explain the physical reasoning behind your prediction.

4

What does the simple model predict?

Compute the ideal fall time before watching anything.

THE IDEAL MODEL

$$t = \sqrt{(2h / g)}$$

$$h = 1.854 \text{ m} \cdot g = 9.81 \text{ m/s}^2$$

Look closely: **the model contains neither mass nor diameter.**

So what — if anything — should make the four balls differ?



ENGINEER DECIDES *record your prediction in the engineering record NOW — no revising after you watch the videos (correctness is not graded).*

Problem

**Model**

Compute



Analyze



Validate



Improve



Watch first — measure nothing yet



Small Bounce Ball (SBB) — Regular Speed

[video separately provided]



Small Steel Ball (SSB) — Regular Speed

[video separately provided]



Large Bounce Ball (LBB) — Regular Speed

[video separately provided]



Standard Tennis Ball (STB) — Regular Speed

[video separately provided]

QUESTION 1

Based only on the regular-speed videos, can you confidently determine whether the four fall times are different?

QUESTION 2

Is visual observation alone sufficient evidence for an engineering conclusion? Why or why not?



TALK TO YOUR AI TEAMMATE *“Here is what each of us thinks we saw at regular speed — log our observations, and note where we disagree.”*

Problem



Model



Compute



Analyze



Validate



Improve



Now measure — one fall time per ball



Small Bounce Ball (SBB) — Slow Motion

[video separately provided]



Small Steel Ball (SSB) — Slow Motion

[video separately provided]



Large Bounce Ball (LBB) — Slow Motion

[video separately provided]



Standard Tennis Ball (STB) — Slow Motion

[video separately provided]

MEASURE LIKE YOU DID BEFORE

Use the same video-analysis method you used earlier in this lab — frame trimming in your phone's player, or Tracker.

Determine one measured fall time for each ball, and record HOW you extracted it (frames, frame rate, release and impact instants).

No new methods, no advanced analysis — the LAB-1 toolkit is enough.



TALK TO YOUR AI TEAMMATE “Here are our four measured fall times and how we extracted them — log them in the engineering record.”

Problem



Model



Compute



Analyze



Validate



Improve



Four numbers against one model

Ball	Mass (g)	Diameter (mm)	Measured Fall Time (s)	Ideal Fall Time (s)	Difference (s)	Percent Difference
SBB	15	29				
SSB	67	29				
LBB	90	57				
STB	57	66				

IDEAL TIME

$$t = \sqrt{(2h/g)}$$

$$h = 73 \text{ in} \\ = 1.854 \text{ m} \\ \text{same for all four}$$

THE QUESTION THAT DECIDES EVERYTHING — TIMING RESOLUTION

Estimate the timing resolution of your video measurement (roughly one frame: $\approx 1 / \text{frame rate}$). Are the measured differences among the four balls large enough to distinguish confidently — or could some differences simply be within the measurement / video-resolution uncertainty?

✓ **ENGINEER DECIDES** *a difference you cannot resolve is not evidence. Compare every difference against your timing resolution.*

Problem



Model



Compute

**Analyze**

Validate



Improve



Interrogate the pattern — don't assume one

MASS

Do the measured fall times show any apparent relationship with mass?

*15 g · 57 g · 67 g · 90 g — four masses, four measured times.
Line them up.*

DIAMETER

Do the measured fall times show any apparent relationship with diameter?

29 mm · 29 mm · 57 mm · 66 mm — same exercise, other variable.

SAME DIAMETER, DIFFERENT MASS

SBB (15 g) and SSB (67 g) share $D = 29$ mm.

One variable held fixed. What does this comparison suggest?

THE LARGER BALLS

LBB (57 mm) and STB (66 mm) give another useful comparison.

What do the measurements suggest when diameter grows?

✓ **ENGINEER DECIDES** *use your measured evidence — not what you expected to see. The data decide what the pattern is.*

Problem

›

Model

›

Compute

›

Analyze

›

Validate

›

Improve



Ask AI for hypotheses — then grade them

SUGGESTED PROMPT

*“Four balls of different masses and diameters were dropped individually from the same height. The simple free-fall model predicts the same fall time for all four. What physical effects **could** cause small differences in their measured fall times?”*

Any capable AI assistant (Engineering Companion). Record a concise summary of the most relevant hypotheses. Do NOT copy the full answer.

NOW CRITIQUE THE AI RESPONSE

- 1 Which suggestions are physically plausible for THIS experiment?
- 2 Which are actually supported by your measurements?
- 3 Any effects that are real — but probably too small to explain what you measured?
- 4 Any claims you cannot verify from the available evidence?
- 5 What additional measurement would test one of the AI hypotheses?

 **THE KEY DISTINCTION** *a physically possible explanation is not necessarily an explanation large enough to account for the measured discrepancy.*

 **ENGINEER DECIDES** *AI generates possibilities. Engineers use evidence to decide which possibilities are credible.*

Problem

›

Model

›

Compute

›

Analyze

›

Validate

›

Improve



Is the simple model good enough?

1 MODEL ADEQUACY

Based on your measurements AND your measurement resolution:
is $t = \sqrt{2h/g}$ adequate for all four balls?

Support your answer with quantitative evidence. Either conclusion is acceptable — “adequate at the resolution of our experiment” or “there is evidence to question it” — if your reasoning is defensible.

2 MODEL IMPROVEMENT & THE PROGRAM

If your evidence says the model should improve: what physical effect would you consider adding?

Conceptually, how would improving the physical model change the computer program you wrote earlier in this lab?

You are NOT expected to derive or program an air-drag model.

Better physical model → different computational model → different computer program

Problem

›

Model

›

Compute

›

Analyze

›

Validate

›

Improve



One compact report — 10 points

LAB EXTENSION CHALLENGE REPORT

separate from your main Lab-1 report · ½–1 page

- 1 Initial prediction and brief reasoning
- 2 Completed four-ball quantitative results table
- 3 Brief measurement-resolution / uncertainty assessment
- 4 Brief mass / diameter comparison
- 5 Concise AI-generated hypotheses
- 6 Your critique / validation of the AI response
- 7 Final judgment: is the simple model adequate?
- 8 How an improved model would change the program

No formal Intro / Methods / Results / Discussion structure required.

Recommended Challenge Rubric	Points
Prediction and physical reasoning	1
Quantitative analysis / results table	3
Measurement resolution / uncertainty reasoning	2
Mass / diameter comparison	1
AI use AND critical evaluation	2
Final engineering judgment / model improvement	1
TOTAL :	10

Serious enough to take seriously — small enough that LAB-1 still belongs to the main lab.

Problem



Model



Compute



Analyze



Validate



Improve

RUBRIC for the Lab Extension Report

© 2026 AI First Engineering - Fundamentals of AI-Augmented Engineering - Chapter 1 / LAB-1
FUNDAMENTALS OF AI-AUGMENTED ENGINEERING
CHAPTER 1 / LAB-1

Lab Extension Challenge Rubric

Four Balls — One Height - Individual Assignment

Suggested Assessment - 10 points

Separate from the main LAB-1 team report. One compact individual challenge report. Assessment rewards evidence, measurement resolution, critical evaluation of AI, and engineering judgment.

Criterion	What full credit looks like	Pts
Prediction and physical reasoning	Initial prediction is recorded before viewing the slow-motion evidence and includes brief defensible physical reasoning. Correctness of the prediction itself is not graded.	1
Quantitative analysis / results table	Four-ball table is complete with measured and ideal fall times, differences, and percent differences; calculations are reasonable and traceable to the video analysis.	3
Measurement resolution / uncertainty reasoning	Video frame rate or equivalent timing resolution is identified and used to judge whether measured differences are <u>actually distinguishable</u> .	2
Mass / diameter comparison	Measured results are compared thoughtfully with ball mass and diameter, including the useful same-diameter comparison where appropriate.	1
AI use AND critical evaluation	Relevant AI hypotheses are summarized concisely and critically evaluated against the measurements; unsupported or insignificant explanations are identified.	2
Final engineering judgment / model improvement	A defensible conclusion is given on model adequacy, supported by evidence; any proposed model improvement is connected conceptually to how the computational program would change.	1
TOTAL		10

Problem



Model



Compute



Analyze



Validate



Improve