Simple and Fun Experiment Demonstrating Accuracy, Precision, and Uncertainty

Charles Baukal
Oral Roberts University (Tulsa, OK)

Abstract

Students sometimes have trouble with the concepts of accuracy, precision, and uncertainty. This paper discusses a simple, fun, and inexpensive experiment using darts to demonstrate these concepts. The lab consists of dropping darts from different heights, a variable number of times, with different darts, and different droppers. Students can see how their accuracy and precision varies as a function of those four variables. The equipment for the lab is very inexpensive and the lab can easily be adapted to different length classes and different numbers of students.

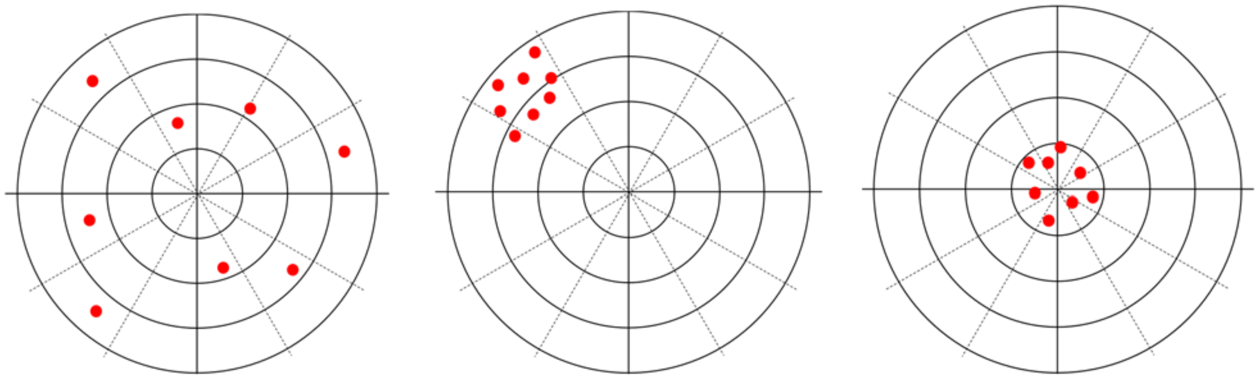
Keywords

Experiment, accuracy, precision, uncertainty

Introduction

Students often believe that accuracy and precision mean the same thing. They also often assume there is no uncertainty in any properties because their textbooks give a single value for each property. All solved problems have a single solution. These imply there is no uncertainty in either property values or solutions, which is clearly not the case in actual practice.

Morris and Langari [1] define (p. 17) the *accuracy* of an instrument as "a measure of how close the output reading of the instrument is to the correct value." They define (p. 17) *inaccuracy* or *measurement uncertainty* as "the extent which a reading might be wrong and is often quoted as a percentage of the full-scale reading of the instrument." *Precision* (p. 18) is defined as "an instrument's degree of freedom from random errors." They note that precision is often confused with accuracy. Figure 1 helps show the difference between the two. Figure 1a shows an instrument that is neither accurate, nor precise. The readings are not very close to the actual value (the center of the target), nor are they very repeatable as they are scattered all over the target. Figure 1b shows an instrument that is precise, but not accurate. The readings are clustered in a tight area, but not very close to the actual value. This is normally easily corrected by calibrating the instrument. Figure 1c shows an instrument that is both accurate and precise, which is obviously the desired condition. The readings are both close to the actual value and clustered together. All three images also show that there is some uncertainty in the readings from the instrument as even Figure 1c shows that in this case, none of the readings are on the actual value. The relatively small deviation from the actual value is the uncertainty for the instrument.



(a) (b) (c)

Figure 1 Accuracy vs. precision: (a) not accurate or precise, (b) precise but not accurate, (c) accurate and precise.

The experiment described here was developed for an Experimental Methods class at Oral Roberts University. The lab is designed to be completed in 75 minutes which is the normal class length for twice-a-week classes. It has been used every spring semester since 2010. It is a favorite among students because it is simple and fun, while clearly demonstrating some important principles.

Experimental Setup

In this experiment, different darts are dropped a variable number of times from different heights onto a target, by different students. The objective is to try to hit the center of the target with each drop. The center represents the true or actual value. Students are given detailed instructions on how to set up the experiment and what is required in the lab report. There are typically 4-5 students in each group and the class typically consists of around 15-20 total students so four sets of equipment are usually sufficient. The initial cost of the equipment is very low as it only consists of darts, tape measures, masking tape, and scrap wood. The wood should be at least slightly larger than the targets printed on 8.5” x 11” paper so there is room to tape a target to the wood. The only ongoing “cost” is the paper targets printed on a copy machine. Four setups fit nicely into a small, inexpensive, plastic toolbox, except for the wood and paper targets which are too big to fit into a typical portable toolbox.

The experimental setup for this lab is simple and straightforward. A target (see Appendix) is taped to a piece of wood (see Figure 2). If the target in the Appendix is used, it needs to be scaled appropriately so the distances shown are correct. Scrap wood may be used behind the targets as the appearance of the wood is not important since it will be getting holes in the surface. Particle board works well because it is harder than plywood. For plywood targets, the darts sometimes get stuck in the wood and have to be pried out which can damage the paper targets. The purpose of the wood is to protect the floor from the pointed darts and to protect the dart tips from getting blunted. A tape measure is used to determine the drop height. Students select three different darts, preferably significantly different from each other. Students then complete the trials as shown in Table 1. A new target is used for each trial to make it easier to distinguish which dart holes go with which trial. Students can take measurements (radius and angle) after each drop, but usually just mark each hole (e.g., see the orange circles in Figure 2) and determine the measurements later.

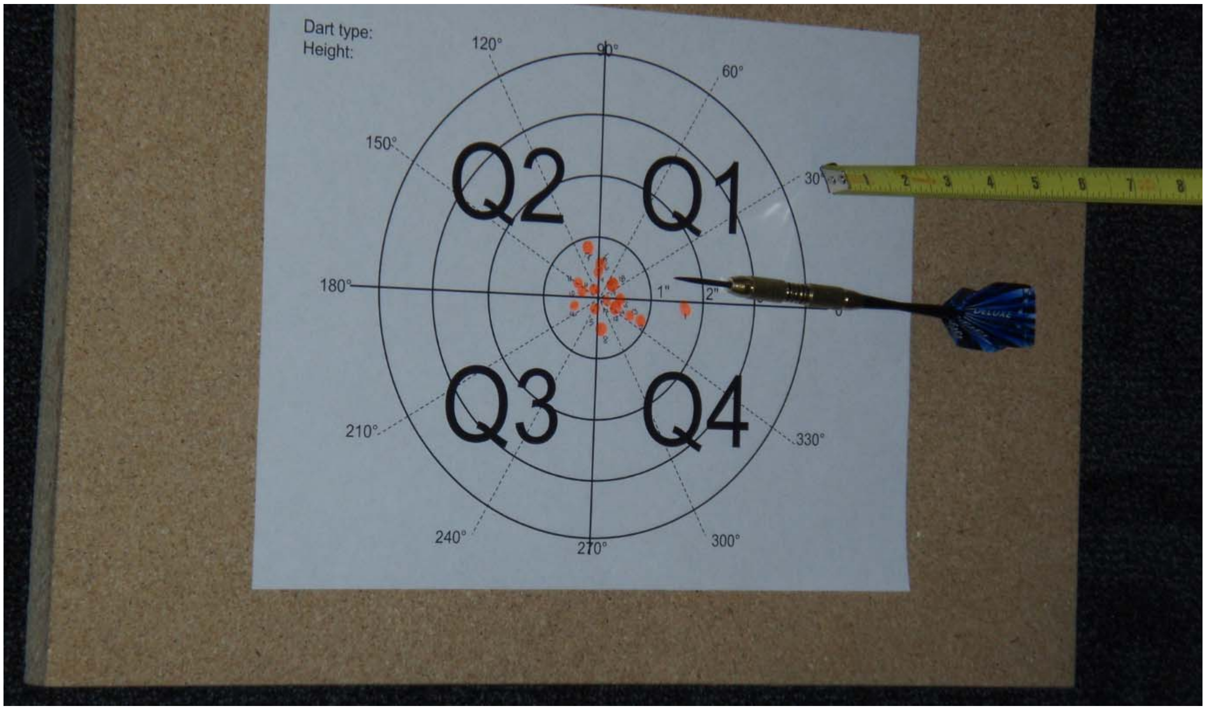


Figure 2 Target after a completed trial.

Table 1 Trials.

Trial	Dart Type	Drop Height (ft)	Dropper	# Drops
1	1	2	1	8
2	1	4	1	8
3	1	6	1	8
4	1	4	1	4
5	1	4	1	12
6	2	4	1	8
7	3	4	1	8
8	1	4	2	8
9	1	4	3	8

A typical setup is shown in Figure 3. An example of the darts is shown in Figure 4. Over the years, some students have made short (less than 3 minutes) instructional videos for this lab, one of which the instructor shows before the lab. These are very helpful and entertaining for the students. A minimum of three students per team are required as three different droppers are needed for the trials shown inAs shown in Table 2, there are four primary variables considered in this lab: drop height, number of drops, dart type, and the dropper. The experiment is designed so there are three variations of each of these variables so comparisons can be made. In the case of drop height, 4 ft is the standard, with a trial each at 2 ft and 6 ft. While these heights could be modified, 2 ft is about as close as is practical and 6 ft is about as high as practical unless stepladders/ladders are available.

Table 2. Typically, at least two students are active during each trial including the dropper and the one marking where each dart hit the target. The latter is necessary because it is not always clear where a dart hit the target, particularly at the 2 ft drop height, as there may not be enough force to make a clear impression in the target.

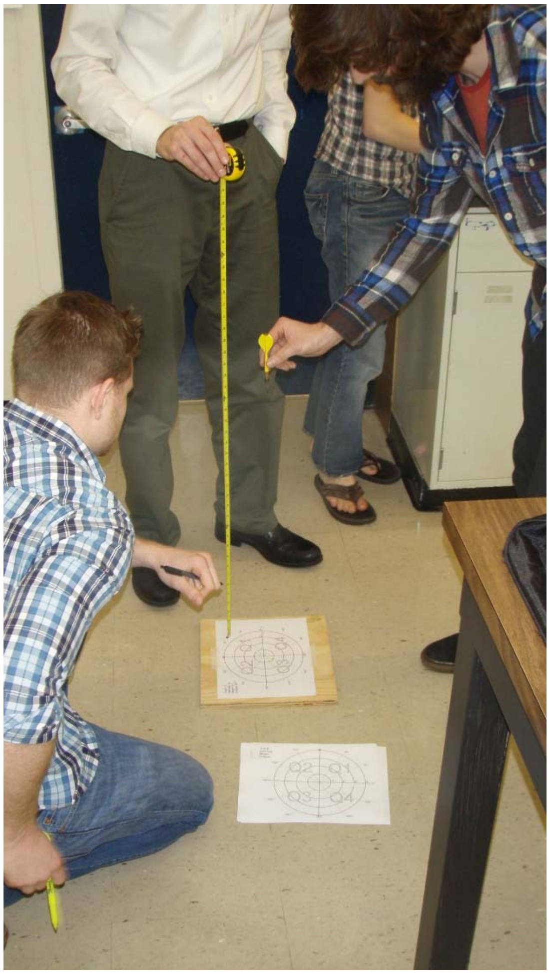


Figure 3 Typical experimental setup.

As shown in Figure 4, the students have 6 darts from which to select 3 for their experiment. There are at least 3 of each of the 6 darts shown and 6 of dart A which is significantly different than the rest. It is recommended every group have dart A for contrast and comparison purposes.

Figure 4 Dart types and their weights: A (8 g), B (20 g), C (25g), D (19g), E (24 g), F (16 g).

Variables

As shown in Table 2, there are four primary variables considered in this lab: drop height, number of drops, dart type, and the dropper. The experiment is designed so there are three variations of each of these variables so comparisons can be made. In the case of drop height, 4 ft is the standard, with a trial each at 2 ft and 6 ft. While these heights could be modified, 2 ft is about as close as is practical and 6 ft is about as high as practical unless stepladders/ladders are available.

Table 2 Variables considered.

	Variables				
Variable	Drop Height (ft)	# drops	Dart type	Dropper	Trials
Drop height	2, 4, 6	8	1	1	1, 2, 3
# drops	4	4, 8, 12	1	1	2, 4, 5
Dart type	4	8	1, 2, 3	1	2, 6, 7
Dropper	4	8	1	1, 2, 3	2, 8, 9

For the number of drops, in the early years of the lab the standard number of drops was 10. However, that number of drops was difficult to complete in the given amount of time and it became tedious for the students so that number was reduced to 8 which seems to work better. The minimum number of drops in a trial is 4 and the maximum in another trial is 12 to see if the number of drops changes the results. One might expect that more drops would improve the accuracy, but that is not always the case.

The third variable is dart type. As shown in Figure 4, a number of darts are available for the students to select the three they will use. It is recommended that they choose three darts that are significantly different from each other so it is easier to determine if the dart type has any impact on the result. Professional dart throwers use heavier darts as they are easier to consistently throw and are not disturbed by, for example, any air currents in the room. It might be expected the heavier darts are even easier to use when dropping them vertically, compared to throwing darts horizontally at a dart board mounted on the wall. However, that is not always the case either.

The last variable considered is the person dropping the darts. The teams select one student to be their main dropper. For the last two trials, two other students are used to see if the dropper has an effect. To make a fairer comparison, the main dropper's second trial is compared against the two other droppers' only trials. While there is a potential learning effect where the first dropper had one trial of practice, it is not believed this has a significant impact.

There are some potential confounding effects for the variables in this lab. Students usually do not get any practice drops so their first trial may not be representative as the droppers are learning what to do to get close to the center of the target. Also, teams rarely pick the same three darts so this is also a potentially confounding variable if comparisons between teams are of interest. Related to that, while dart A is significantly different than the other darts, there is less difference between the rest of the darts which are all fairly similar to each other.

Results

Students are asked to do the following with their results to be included in their lab reports.

- Find the radius and angle for all points on each target (eyeball estimates for angles are fine)
- Make 9 tables (to go in the Appendix), one for each trial
- Order the data in each table from the closest radius (in.) to the farthest radius, and from the smallest angle to the largest angle
- Include the actual targets from each trial in the Appendix of the lab report

They are asked to calculate the following:

- Average radius for each trial
- Radius standard deviation for each trial
- Error bars for the radius measurements using $\pm 1\sigma$ (1 standard deviation)

Table 3 shows some sample results for a single trial. The data are sorted from closest distance to the center to the farthest distance. The data are then put into four radius categories (R1 = 0.00 – 1.00 in., R2 = 1.01 – 2.00 in., R3 = 2.01 – 3.00 in., R4 = 3.01 – 4.00 in.) and four angle categories (Q1 = 0.0° – 89.9°, Q2 = 90.0° – 179.9°, Q3 = 180.0° – 269.9°, Q4 = 270.0° – 359.9°). To date, all data have fallen within the 4 in. radius of the printed targets.

Table 3 Example results for a single trial ordered by distance from target center.

Radius (in.)	Angle (degrees)	Radius Category	Angle Category
0.5	33	R1	Q1
0.7	263	R1	Q3
0.8	125	R1	Q2
0.8	97	R1	Q2
1.5	345	R2	Q4
1.7	220	R2	Q3
2.1	182	R3	Q3
3.2	312	R4	Q4

The following instructions are for result 1a which is the first part of the results related to drop height: make a scatter plot comparing trials 1 – 3: *x*-axis = drop height (ft), *y*-axis = average radius (in.), and 1 line connecting the 3 data points with error bars equal to 1 standard deviation above and below each point. An example plot is shown in Figure 5.

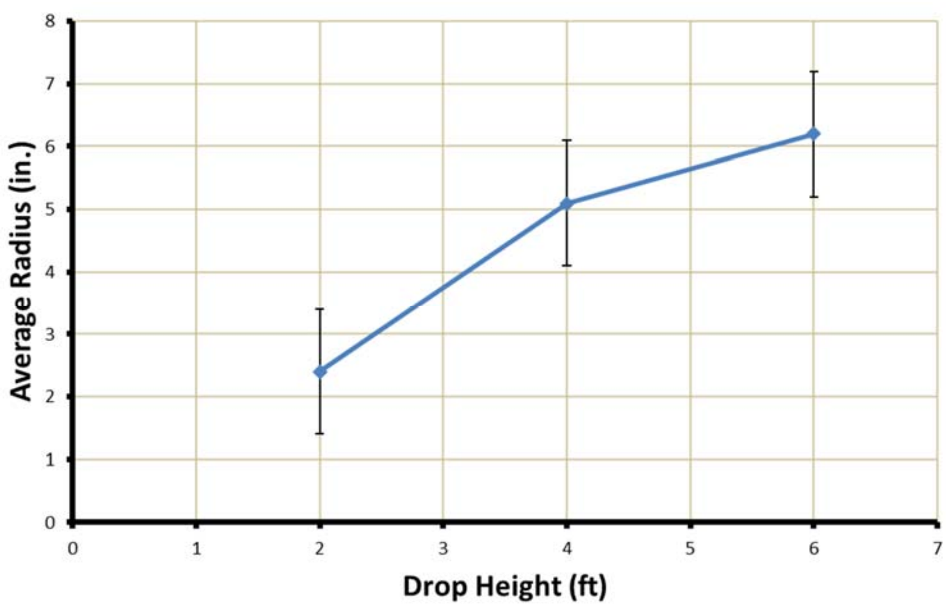


Figure 5 Example result for drop height (trials 1, 2, and 3).

The instructions for result 1b are to put the trial 1 data in radius and angle categories as described above and shown in Table 3. Then the students are asked to determine the frequency of data in each radius and angle category. Table 4 shows some example results. Figure 6 shows the data in Table 4 plotted as bar graphs. Figure 6a gives an idea of the accuracy of the drops for the first trial, where the closer to the target center the more accurate the drops. That particular plot shows the drops were not particularly accurate as most of the drops were in the farther radius categories

(R3 and R4). Figure 6b gives an idea of precision with respect to angle. Most of the drops were in angle categories Q1 and Q2, so the dropper consistently dropped on one side of the target.

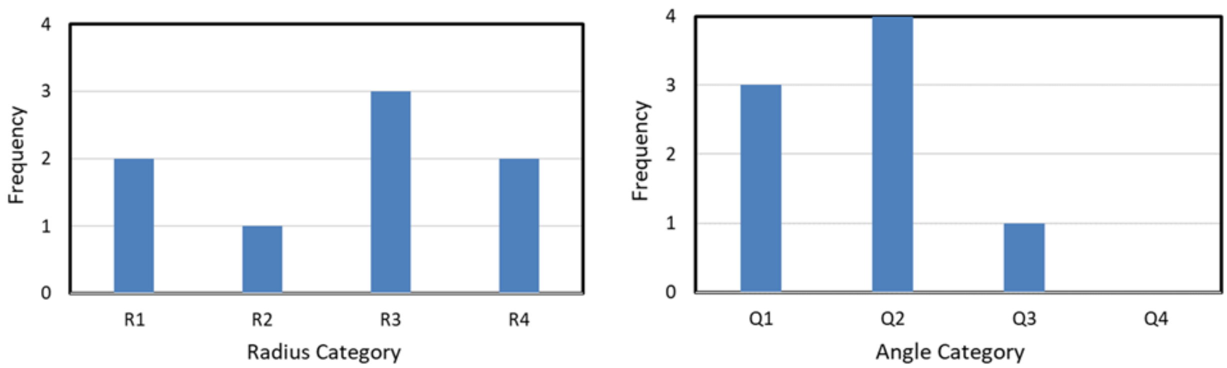
Table 4 Example results for a single trial by radius and angle category.

(a)

Radius Category	Frequency
R1	2
R2	1
R3	3
R4	2

(b)

Angle Category	Frequency
Q1	3
Q2	4
Q3	1
Q4	0



(a) (b)

Figure 6 Example results for drop height by: (a) radius category, (b) angle category.

For the second major result which is related to the number of drops, students are asked to make a scatter plot comparing trials 2, 4, and 5: *x*-axis = number of drops, *y*-axis = average radius (in.), and one line connecting the three data points with error bars equal to one standard deviation above and below each point. Figure 7 shows an example result for the number of drops. In that particular example, the accuracy improved with the number of drops.

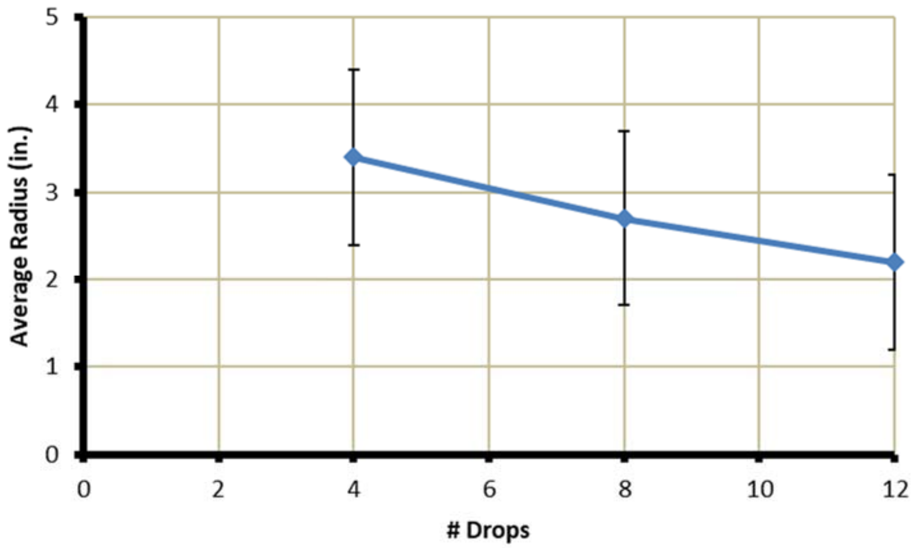


Figure 7 Example result for number of drops (trials 2, 4, and 5).

The third major variable is the dart type. Students are instructed to make a line graph comparing trials 2, 6, and 7: *x*-axis = dart type, *y*-axis = average radius, and one line connecting the three data points with error bars equal to one standard deviation above and below each point. Figure 8 shows an example result for dart type comparisons. In this example, dart D was more accurate compared to darts A and F.

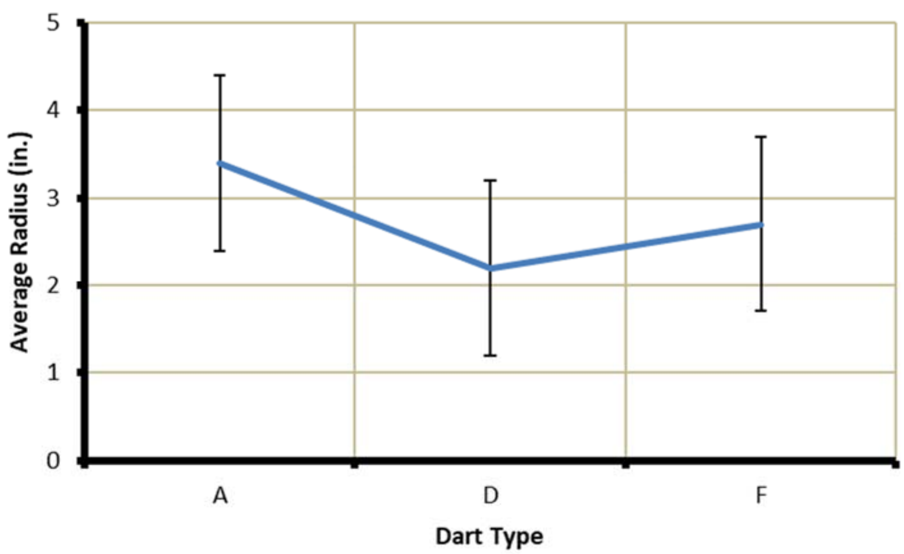


Figure 8 Example result for dart type (trials 2, 6, and 7).

The last major variable is the dropper. Students are instructed to make a line graph comparing trials 2, 8, and 9: *x*-axis = dropper (1, 2, 3), *y*-axis = average radius, and one line connecting the three data points with error bars equal to one standard deviation above and below each point. Figure 9 shows an example result for the dropper. In that example, dropper 1 was the most accurate.

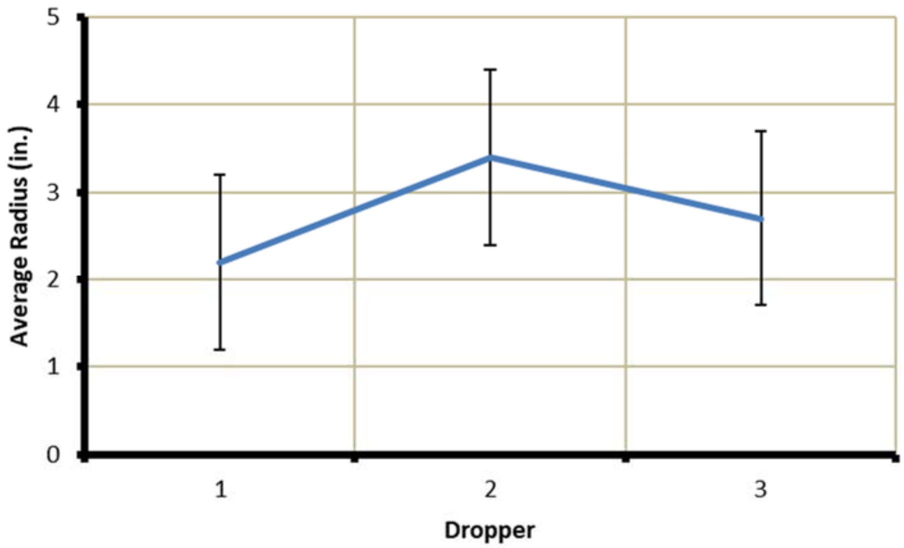


Figure 9 Example result for the dropper (trials 2, 8, and 9).

In their lab reports, students are asked to answer the following questions:

- Is the accuracy (deviation from the center point) a function of the drop height?
- Is the experimental uncertainty (standard deviation) a function of the number of drops?
- Are there any differences in accuracy between the different dart types?
- Are there any differences in accuracy between the different droppers (people dropping the darts)?

Discussion

The purpose of putting the data into radius and angle categories is to visualize the accuracy (the closer the radius is to the center the more accurate) and the precision (e.g., the more data in one angle category the more precise).

Students can also see if there are any outliers. This is more likely to happen during the first drops by a dropper until they get used to the process. It is also more likely at the highest drop height as it gets more difficult to hit close to the center from higher heights. The author is not aware of an instance where students completely missed the target.

There are some things that might be done to improve the validity of this experiment, but which are not usually possible due to time constraints. Note that high validity is not a specific objective of this lab. One thing that could be done is to have each dropper take the same number of practice drops and compare the trials right after their practice drops. Another is that one trial could have a significantly greater number of drops compared to the others to see if the standard deviation is reduced as might be expected.

Conclusions

The lab described here is simple, inexpensive, and fun. There are many possible variations of the lab including changing the variables (e.g., different heights than shown here) and the number of variations for each variable (e.g., fewer or more droppers than 3). This may be desired if the time for the lab is significantly shorter or longer than 75 minutes.

Lab teams are recommended to be no fewer than 3 and no more than 6. Fewer than 3 means there will be fewer different droppers and more work per student for analyzing the data and writing the report. More than 6 means there will be too many students not actively engaged in conducting the lab, which will be de-motivating.

This lab is very portable and can be done in almost any location, including outside if desired. There is not much equipment needed and it can easily be carried anywhere. It does not require a permanent setup like some labs require.

There are some caveats, particularly related to safety. While it has not happened after 12 years of conducting this lab, it is possible a student gets hurt because the darts are sharp. Horseplay is not permitted because of the danger of using darts. The classes doing this lab are relatively small and can effectively be monitored by a single instructor; larger classes should have more supervision to make sure no one gets hurt.

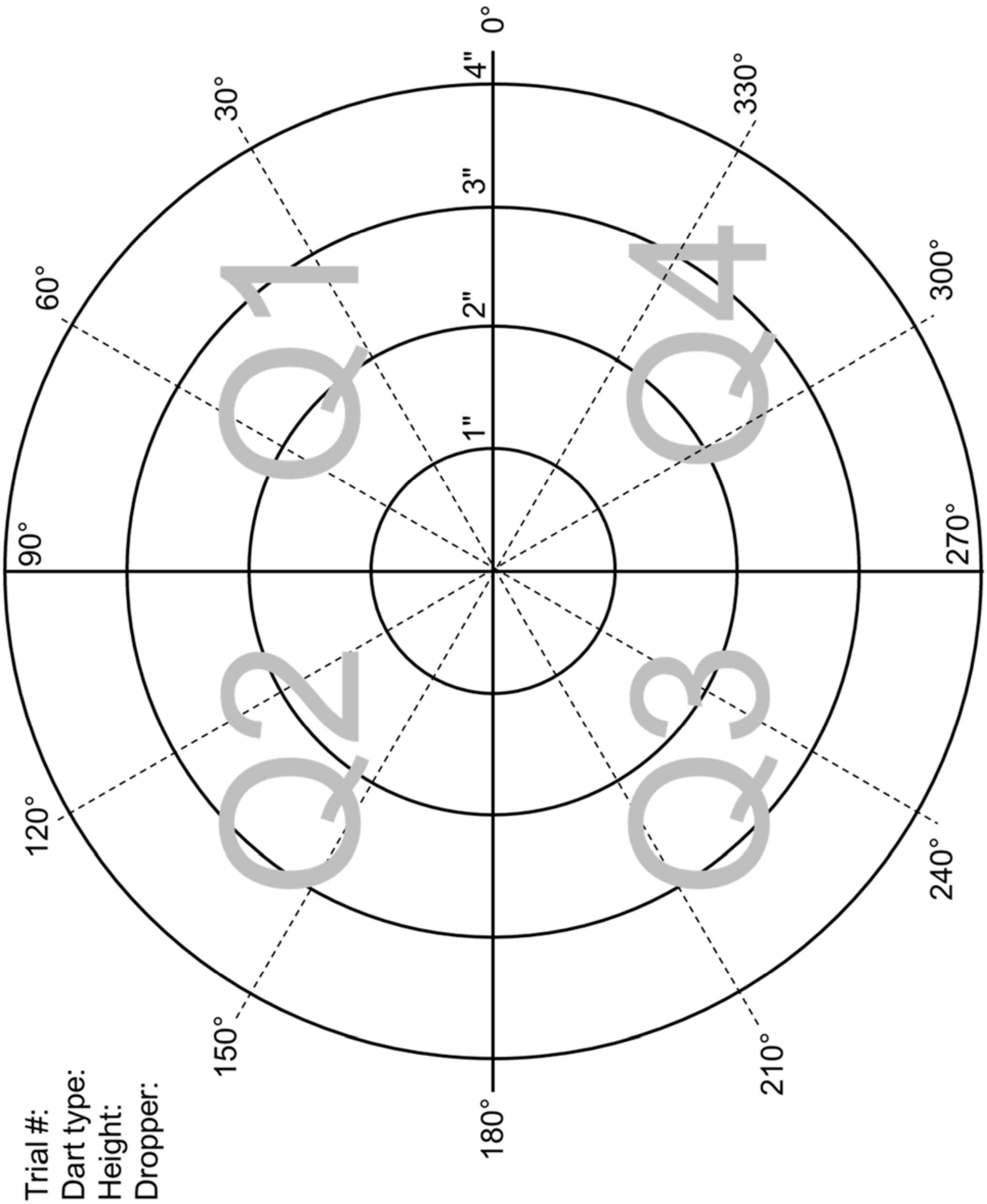
References

[1] A. Morris and R. Langari, *Measurement and Instrumentation*, 3rd ed., Academic Press, London, 2021.

Charles Baukal

Charles Baukal is an adjunct instructor in Mechanical Engineering at Oral Roberts University where he has been since 2008. He works full time as the Director of the John Zink Institute which is part of John Zink Hamworthy Combustion headquartered in Tulsa, OK where he has been since 1998. He has a Ph.D. in Mechanical Engineering from the University of Pennsylvania and an Ed.D. in Applied Educational Studies from Oklahoma State University. He is a registered Professional Engineer and is an inventor on 11 U.S. patents. He has authored / edited 16 books on the subjects of industrial combustion and engineering education. He has over 250 publications and presentations and serves on many advisory boards.

Appendix



Note: the target should be scaled so the distance dimensions are correct.