# **Similar Density Questions with Very Different Results**

Ashley Lindow\*, Elizabeth Carbone\*, DJ Wagner\*

\*Department of Physics, Grove City College, 100 Campus Drive, Grove City, PA 16127

**Abstract:** While developing a standardized fluids assessment covering buoyancy and pressure, we discovered deficiencies in student understanding of density. In particular, many college students do not recognize that density is a fixed property of a solid substance, such as aluminum or gold. We added questions to our diagnostic exam to probe the extent of student difficulties. In one of our questions, only 50-60% of students (depending upon class) recognize that the density of gold is a fixed value. When similar questions from an existing diagnostic are used, however, 88-100% of students correctly identify the density of a piece of wood and of a diamond as fixed values. In this paper we discuss the differences between these questions and how those differences affect student responses.

<sup>1</sup> Yeend's Density Survey, WWW Document, (http://www.compadre.org/Repository/document/ServeFile.cfm?ID=4313&DocID=1045).

Keywords: Density, Fluids, Education Research, Question Format

PACS: 01.40.-d, 01.40.G-

### INTRODUCTION

While conducting interviews during the development of a fluids statics assessment, we discovered that many of our students do not have a robust understanding of mass density. In particular, only two of our eight interviewees demonstrated reliable knowledge of density as an intensive property of a solid material. To probe the prevalence of this difficulty, we wrote questions concerning density to add to our assessment in spring, 2012. In fall, 2012 (F12) and spring, 2013 (S13), we included questions from Yeend's density assessment [1] to further explore this topic.

One question we wrote requires students to recognize that the density of 24-karat gold is independent of the size or mass of a particular piece. Results were poor, as shown in Tables 1 and 2. However the results were much better for two apparently-similar questions by Yeend. This paper describes our exploration of the effect on student performance due to various "surface" differences between the questions.

### PRIOR WORK

We are not the first to observe this difficulty with density as an intrinsic property of a material. Many studies [2-16] have documented difficulty with density at a variety of ages, and a few of those [2-7] support our observation that the difficulty persists into college.

Students' failure to acknowledge that density is a property of a material can have many sources, one of which is that students often view density as an extensive quality, dependent on size and/or mass. Roach [2] gave a test to a class of 43 college students in general science,

post-instruction. Only 6% of the students demonstrated knowledge that density is independent of size and shape of an object. Yeend et al. [3] also gave a post-instruction test to middle school, high school, and college students. Questions concerning density as an intensive property were answered correctly by only about 30% of the students.

Another facet of students' difficulty with density is overuse of the density equation. Instead of recognizing that two objects of the same material have the same density, students may say that they cannot compare densities without being given both the mass and volume of each object (to use with the "density equation").

A related error is the common tendency for students to inappropriately apply linear relations to problems that cannot be modeled linearly [17]. The form of the density definition may lead some students to apply a linear dependency of density on mass, i.e. four times the mass means four times the density.

Students exhibiting these difficulties are often able to identify the correct memorized formula [8-10], but still might view density as an extensive property of the object. This difficulty highlights how "memorized equations are not sufficient for student understanding or for arriving at the correct answer" [4].

## THE STUDY

The questions from this study were included in assessments taken by students in all three introductory physics classes at Grove City College (GCC). These classes include a calculus-based ("cale") course taken by engineers and physics and computer science majors, a trigonometry-based ("trig") course taken by biology and chemistry majors, and a concept-based ("cnpt") course taken by non-science majors.

In F12, pre-instruction, students performed far worse on our "Rho Gold" question than on the original form of Yeend's similar questions, "Three Piece" and "Diamond." Full text for these questions, and our subsequent modifications, are shown in Fig. 1. Two differences between Yeend's questions and our question

are (1) Rho Gold specifies the ratio of masses of the two pieces whose density are to be compared, and (2) Rho Gold presents the two pieces as initially separate, rather than as a single object which is then cut. The modified questions, given post-instruction F12 and S13, each targeted a single difference.

"Rho Gold": Alberta has several pieces of 24-karat gold, of various shapes and sizes. She measures the density of the first pieces and calls that value d. The second piece of gold she picks up has four times the mass of the first piece. What is its density?

- A.  $d \div 16$ .
- B.  $d \div 4$ .
- C. d.
- D.  $2 \times d$ .
- E.  $4 \times d$ .
- F. Alberta needs more information to determine the density.

"Diamond": A jeweler cut a small chip off a large, uncut diamond. How does the density of the chip compare with the density of the original diamond?

- A. The density of the chip is the same as the density of the original diamond.
- B. The density of the chip is smaller than the density of the original diamond.
- C. The density of the chip is larger than the density of the original diamond.
- D. Impossible to tell unless the volume and mass of each piece are given.

"Diamond Ratio": A jeweler cut a small chip off a large, uncut diamond. The chip has a mass 1/100 as large as the original diamond. How does the density of the chip compare with the density of the original diamond?

- A. The density of the chip is the same as the density of the original diamond.
- The density of the chip is 1/100 of the density of the original diamond.
- The density of the chip is 1/1,000,000 of the density of the original diamond.
- The density of the chip is 100 times larger than the density of the original diamond.

"Diamond Separate": A jeweler is considering various large diamonds and small diamond chips when making a necklace. How does the density of a small diamond chip compare with the density of a large diamond?

- A. The density of the small chip is the same as the density of the large diamond.
- B. The density of the small chip is smaller than the density of the large diamond.
- C. The density of the small chip is larger than the density of the large diamond.
- D. Impossible to tell unless the volume and mass of each piece are given.

"Three Piece": A straight, uniform board is cut into three differently sized pieces. Each piece has identical width and thickness.







Which piece has the greatest density?

- A. Piece A
- B. Piece B
- C. Piece C
- D. They are all the same.

"Three Piece Ratio": A straight, uniform board is cut into three pieces of different length: B is twice as long as A, and C is three times as long as A. Each piece has identical width and thickness.







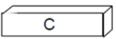
How does the density of piece C compare to the density of piece A?

- A. Piece C has 27 times the density of A.
  - B. Piece C has 3 times the density of A.
- C. Piece C has the same density as A.
- D. Piece C has 1/3 the density of A.
- E. Piece C has 1/27 the density of A.

"Three Piece Separate": Consider three rectangular pieces of aluminum with different length. Each piece has identical width and thickness.







Which piece has the greatest density?

- A. Piece A
- B. Piece B
- C. Piece C
- D. They are all the same.
- Impossible to tell without more information.

"Three Piece Al": A straight, rectangular piece of aluminum is cut into three differently sized pieces. Each piece has identical width and thickness.

Which piece has the greatest density?

- A. Piece A
- B. Piece B
- C. Piece C
- D. They are all the same.
- E. Impossible to tell without more information

FIGURE 1. Original version of the Rho Gold, Diamond, and Three Piece questions, along with the modified versions for Diamond and Three Piece.

TABL	<b>E 1.</b> F12 Post results b	y class. Results for	r question indicate	percent correct and	l standard error.
------	--------------------------------	----------------------	---------------------	---------------------	-------------------

	Sample Size		"Rho Gold"	"Three Piece"		"Diamond"	
	Version A	Version B	Versions A & B	Version A, Original	Version B, Ratio	Version B, Original	Version A, Ratio
Calc	69	72	51± 4.2	$86.8 \pm 4.1$	$87.9 \pm 4.0$	$89.4 \pm 3.8$	$89.7 \pm 3.7$
Trig	25	25	$56 \pm 7.0$	100	$84 \pm 7.5$	$96 \pm 4.0$	$92 \pm 5.5$
Cnpt	37	41	$58 \pm 5.5$	$90 \pm 4.8$	$83.8 \pm 6.1$	$94.6 \pm 3.8$	$92.5 \pm 4.2$

**TABLE 2.** S13 Post results by class. Results for question indicate percent correct and standard error.

	Sample Size		"Rho Gold"	"Three Piece"		"Diamond"	
	Version A	Version B	Versions A & B	Version A, Original (Al)	Version B, Separate	Version B, Original	Version A, Separate
Calc	54	50	$53 \pm 4.8$	$88 \pm 4.6$	$83.3 \pm 5.1$	$87 \pm 4.6$	$82 \pm 5.5$
Trig	17	18	$71 \pm 7.6$	$88.9 \pm 7.6$	$88.2 \pm 8.1$	$100 \pm 0$	$88.9 \pm 7.6$
Cnpt	18	14	$59 \pm 8.7$	$71.4 \pm 12.5$	$66.7 \pm 11.4$	$88.9 \pm 7.6$	$64.3 \pm 13.3$

In F12 and S13, two versions of our fluids assessment were prepared and randomly distributed post instruction. Version A included the original version of Three Piece and the modified version of Diamond, while Version B included the original version of Diamond and the modified version of Three Piece.

The altered questions given in F12 were Diamond Ratio and Three Piece Ratio, and the altered questions given in S13 were Diamond Separate and Three Piece Separate. These modified questions are shown in Fig. 1. Because different types of wood might have different densities, we changed the substance used in the original Three Piece and Three Piece Separate from wood to aluminum in S13, when we allowed for the pieces to originate from different objects. Additionally, we added a "not enough information" option to Three Piece in S13, to increase similarity between the Three Piece and the other two questions. Since we do not directly compare results on this question across semesters, we believe these changes do not invalidate our conclusions.

The results from both alterations were then analyzed using a three dimensional analysis of variance (ANOVA) with two variables (course and test version) between subject and one (whether the question was modified) within subject.

#### RESULTS

As expected from our initial comparison of Yeend's questions with the Rho Gold question, our students generally performed better on the unmodified questions than on the modified questions. Tables 1 and 2 show

the percent of students answering each question correctly, organized by course. The modifications we made to the questions caused a statistically significant difference in both cases.

In F12, the addition of a ratio had a significant, negative effect on student performance, Wilks' Lambda = .984, F(1, 255) = 4.074, p = .045. Partial Eta Squared was 1.6% for the comparison between the original and ratio questions.<sup>1</sup>

In S13, the modification to include initially separate objects had also had a significant, negative effect on student performance, Wilks' Lambda = .971, F (1,165) = 4.986, p = 0.027. Partial Eta Squared was 2.9% for the comparison between the original and modified questions. While the separate/cut distinction appears to account for a greater percentage of variance in student answers than does the inclusion of a ratio in the question, a z-test<sup>2</sup> shows the difference in eta-squared not to be significant.

Interestingly, the class in which the student was enrolled did not have a significant effect (p = .549) in F12, meaning that the apparent differences between class performance could be due to random factors. The class did have a significant effect, F (2,165) = 3.534, p = .031, in S13. For S13, Partial Eta Squared for variation by course was 4.1%, so the course in which the student was enrolled accounted for 4.1% of the variance. This particular difference between semesters may, in part, be due to differences in population between the fall and spring calc and trig classes. In S13, we gave the assessment to calc and trig students in their second physics class, so students who were unsuccessful in their first semester were not included in this population.

<sup>&</sup>lt;sup>1</sup> Partial Eta Squared is essentially a measure of what percentage of variance in student responses can be attributed to a particular factor and a value over 1% is considered a moderate effect.

<sup>&</sup>lt;sup>2</sup> Fisher's z' transformation and comparisons between Independent r's yielded a z-value of 0.45. [18] For the difference between eta-squared values to be considered significant, z should be at least 1.96.

#### **CONCLUSIONS**

Students' identification of density as an intensive property of a material is dependent upon many features of the question. When a ratio, or a specific number, is given, many students will try to use that information in the answer, or assign linear dependence to the constant of proportionality. Furthermore, students are more likely to treat the density of two pieces of a single object as fixed than to treat the density of any object made of a particular material as fixed. We observed that both of these context changes significantly affect whether or not students answer the question correctly.

The modifications making Yeend's questions more like our Rho Gold question, while significant, still do not appear to account for the entire (30-40%) difference in student performance between questions. Location in the test could be another factor – the Rho Gold question is found early in the assessment without similar questions around it, while the Three Piece and Diamond questions were appended to the end of the assessment along with other density-related questions.

Another possible factor in the difference in student performance could be that the format of the Rho Gold question closely resembles missing-value format.<sup>3</sup> Because missing-value format is so often used to test linear reasoning, students develop a "superficial association between a linguistic problem formulation and a solution procedure" [17]. Thus, students often apply linear reasoning to any missing-value question – even if it is not valid. Even though we added numbers to Yeend's questions in F12, our modified versions asked for a ratio rather than a "missing value." The fact that our Rho Gold question resembles missing-value format while the original and modified versions of Yeend's questions do not may contribute highly to the difference in the percent of correct answers.

Finally, surface features such as context or answer order can affect responses. When discussing our results with colleagues, the question was raised whether students know that 24-karat gold is a pure substance. Since the unmodified diamond and wood questions of Yeend had similarly-good results, we are not sure that specific material has a huge effect in this case. But the significance of material is worth future investigation.

Based on our findings, we believe that use of Yeend's assessment could lead to an over-estimation of students' mastery of density as an intensive quality of a substance. While it might be desirable to leave out numbers in order to avoid the missing-value difficulty, we do believe there is a significant difference between understanding density being fixed for an object and density being fixed for a substance. It may be beneficial

to add cut/separate modifications to one or both of the questions on Yeend's Density Survey in order to create a more complete assessment.

Additionally, this study should be considered during the formulation of questions to be used for formative assessment. Instructors may want to build redundancy with different contexts into their assessments, to ensure they are measuring what they intend.

#### **ACKNOWLEDGEMENTS**

The authors would like to thank Dr. Gary Welton who provided extensive guidance with the statistical analyses done for this study. They would also like to thank Matthew Goszewski and Zachary Bazan who contributed to this projects as members of the GCC PER group. Partial support for this project was provided by the Grove City College Swezey Scientific Instrumentation and Research Fund.

### REFERENCES

- R. Yeend, M. Loverude and B. Gonzalez (2001), URL (http://www.compadre.org/Repository/document/ServeFi le.cfm?ID=4313&DocID=1045).
- 2. L. Roach, J. Coll. Sci. Teaching, 30 (6), 386 (2001).
- R. Yeend, M. Loverude and B. Gonzalez, in *Physics Education Research Conference 2001*, Rochester, New York, 2001, (unpublished).
- 4. J. McKinnon, J. Geol. Educ., 19 (5), 218 (1971).
- 5. H.F. J. Blanck, Chem. Educ., 54, 628 (1997).
- M.E. Loverude, C.H. Kautz and P.R.L. Heron, Am. J. Phys., 71, 1178–1187 (2003).
- 7. S.J. Hawkes, J. Chem. Educ., 81(1), 14-15 (2004).
- 8. G. Fassoulopoulos, P. Kariotoglou, and P. Koumaras, *Research in Science Education*, 33 (1), 71-87 (2003).
- L.E. Klopfer, A. B. Champagne, S. D. Chaiklin and W. T. Wollman, Sci. Educ., 76 (6), 597-614, (1992).
- J.A. Rowell and C.J. Dawson, Sci. Educ., 61, 245–253 (1977).
- 11. E.D. Gennaro, School Science and Mathematics, 81, 399–404 (1981).
- 12. M.G. Hewson, Sci. Educ., 70, 159-170 (1986).
- 13. L. Peterson-Chin, Science Scope, 27 (7), 16-20 (2004).
- 14. S. Ünal and B. Coştu, *Asia-Pacific Forum on Science Learning and Teaching*, 6 (1), Article 3 (2005).
- 15. M. Chandraskhar and R. Litherland, *Science Scope*, 30 (2), 12-17 (2006).
- 16. H. Stavridou and S. Grammenos, *International Journal of Learning*, 16 (7), 311-323 (2009).
- 17. W. V. Dooren, D. De Brock, D. Janssens and L Verschaffel, *Journal for Research in Mathematics Education*, 39 (3), 331-342, (2008).
- J. Cohen and P. Cohen, Applied Multiple Regression/Correlation Analysis for the Behavioral Science, (Lawrence Erlbaum Associates Inc., 1983).

<sup>&</sup>lt;sup>3</sup> Missing-value format provides students with three values and prompts the student to provide the fourth. For example, "One apple

costs \$0.30, how much do 7 apples cost?" is a question in missing-value format