Expectations for Statistical Literacy: A Survey of College Faculty

S. David Kriska[1], Mark C. Fulcomer[2], Marcia M. Sass[3]
[1]Management Sciences, The Ohio State University, 3758 Surrey Hill Place, Upper Arlington, OH, 43220
[2]Richard Stockton College of New Jersey
[3]UMDNJ-School of Public Health

1. Introduction

Abstract

This poster presented results from a survey of college faculty regarding expectations for students who have completed introductory students. The sample included faculty who are not statisticians or statistical educators. Nevertheless, their expectations for literacy are consistent with those of several statistical educators.

KEY WORDS: Statistical literacy, business statistics education

Much has been written regarding statistical literacy and we reviewed several perspectives last year. Concise definitions, however, need to be fleshed out to be of use to educators. Moreover, the expectations for literacy are often dependent on the way in which people present themselves. Literacy for a high school graduate has different meaning than literacy for a college graduate. Last year, Kriska, Fulcomer, and Sass (2006) proposed an instrument that could be used to survey various stakeholders to see more specifically what knowledge is important in being statistically literate and what tasks a statistically literate person should be able to perform. In this paper, we present the results of a survey of the Fisher College of Business faculty in which respondents identify their statistical expectations for students entering classes that follow introductory statistics. Specifically, we asked faculty and Ph.D. students who typically have undergraduate class room responsibility:

- What statistical tasks should students be able to do upon entry into their classes?
- What are expectations regarding specific statistics course topics for students entering their classes?

The survey of a business faculty provides a perspective that is unique and important. Most who write about statistical literacy are statistical educators publishing in statistical education journals. The business faculty, however, has as its chief aim educating students with regard to content knowledge about business and providing tools for the students to become effective business decision makers. For the business faculty, teaching of statistics is not a primary objective even though most of the faculty are sophisticated statistical consumers and producers. The importance of this paper is based on the perspective the sample brings to the evaluation task.

2. Method

After obtaining an exemption from the Institutional Review Board at The Ohio State University, two listserves were used to solicit volunteers from the Fisher College of Business Faculty. The first listserve included 105 faculty plus others and the second listserve included approximately 62 Ph.D. students. To obtain the IRB approval, Rao Unnava, the Associate Dean for Undergraduate Education in Fisher was designated as the lead investigator. To solicit volunteers, an e-mail was sent to the faculty and Ph.D. students using the two listserves and a message signed by both Unnava and Kriska. The message to faculty was distributed on June 28, 2007 and the message to Ph.D. students was distributed on June 29, 2007. Follow-up messages, thanking those who had responded and encouraging participation of those who had not yet responded, were sent on July 5, 2007 to the Ph.D. students and on July 6, 2007 to the faculty. The e-mail messages directed the potential respondents to a Qualtrics web site that contained the survey. Data were analyzed based on responses received by July 18, 2007. Anonymity was promised to respondents when making application to the IRB and the Qualtrics software was set up to use anonymous responding. No payment was offered to the respondents except that they could receive a copy of the results. Tables summarizing the results were e-mailed to those who had made this request in August 2007.

After providing a statement that the survey was voluntary and anonymity promised, demographic data concerning departments with which respondents were associated were collected. Then a question was asked about courses on which the responses would be based. Respondents who had received the survey, but who were not associated with a course requiring statistics were asked to respond "none" to this question, and then asked to quit the survey.

Following the demographic items, the statistical survey followed. The instrument, described in Kriska et al. (2006) and modified slightly based on review by several colleagues, contains two parts. The first portion contains 30 statistical tasks in which each task begins with an action verb and the object of the verb contains a specific statistically related activity. Respondents used a 1 to 5 Likert scale with the focus being on level of independence expected of the student in performing the task and including the following anchors: 1 (No Expectation), 2 (Elementary Expectations), 3 (Minimal Guidance), 4 (Independent), and 5 (Leader) in performing the task. The second portion of the survey contains 38 topics that may be covered in introductory statistics and respondents were asked to evaluate the topics using a scale with the following anchors: 1 (No Expectation), 2 (Rote Memory), 3 (Comprehension), 4 (Application), 5 (Synthesis), and 6 (Evaluation.) The scale anchor points for the second portion of the survey are based on Bloom's taxonomy and were designed to elicit expectation in terms of cognitive processing related to the topic.

3. Results

3.1 Participant Behavior

Table 1 provides results related to the activity of the potential participants in responding to the e-mails inviting them to participate in the survey. Interesting results involve comparisons between the faculty and Ph.D. students. Because the exact number of potential respondents is not known, an upper bound to the response rate is 37%. The Qualtrics survey maintains data on responding to a survey. In particular, completion time is an interesting construct. Qualtrics denotes as respondent as having completed the survey if all items are completed. Note that one faculty completed all items except the last one which asked if the respondent wished to receive a copy of the results. Because the Qualtrics system allows respondents to leave and return to the web site, median times to complete are of greater interest than mean times. Clearly, though, for those who completed the survey, the Ph.D. students appear to have spent less time completing the instrument than faculty who answered all items.

3.2 Statistical Tasks

Table 2 present descriptive statistics for the 17 statistical tasks that were rated as 2.5 or higher, a point half way between the anchors *Elementary Expectations* and *Minimal Guidance* for task performance. Tasks with a mean rating greater than or equal to 3.0 included several indicative of communication tasks, e.g., constructing graphs, making arguments using statistical analysis, and report writing. Making decisions consistent with data analysis is also rated high. Also included in the list are several data management and computation tasks indicative of using a hand calculator and working with spread sheets and data files.

One striking feature of these results is that faculty consistently rated the tasks higher than the Ph.D. students. While no hypothesis tests had been planned regarding these differences, tasks 1, 16, 20, and 24 would have resulted in significant *t*-tests with alpha equal to .05 and application of the Satterthwaite approximation. A second noticeable feature of the results is the relatively large standard deviations.

3.3 Statistical Course Topics

The results for the highest rated topics often taught in introductory statistics appear in Table 3. The scale used in rating the topics focused on the level of cognitive processing the student is expected to exhibit when using the topic. Relevant anchors from the Bloom taxonomy based ratings of the course topics are 2 (*Rote Memory*), 3 (*Comprehension*), and 4 (*Application*).

The highest rated topics, graphing, variation, central tendency, normal distribution, and covariation are almost universally covered in beginning statistics. The relatively low variance in ratings for the variance and central tendency items are of interest and indicative of modest agreement in the importance of these topics.

Several of the specific items deserve comment and will be discussed later. First, the results for hypothesis testing are most interesting. There is a relatively large difference between the mean rating of the Ph.D. students and the faculty and it is a large variance item, especially for the faculty respondents. A second survey item worthy of comment is sampling distributions. The Ph.D. students, on average, saw this topic as one of rote memory whereas the faculty saw it as one in which comprehension is expected.

Table 1: Characteristics of Respondent Behavior

	Faculty	Ph.D. Students	Total
ListServe Counts	105 Faculty + others	About 62	167 + others
Started survey	34	31	62
Completed Survey (As noted by Qualtrics)	13	16	29
Provided usable survey question data	18	13	31
Median completion time for those marked as finished by Qualtrics	7.83 minutes	5.58 min	
Q1 (Finished)	6.32 Min	3.30 min	
Q3 (Finished)	12.17	8.44 min	
Number of starts Immediately following first notice	10 (June 28)	7 (June 29)	17
Number of starts Immediately following reminder notice	11 July 6	20 (July 5)	31
Number of respondents saying “none” to association with courses with stat prerequisite	4	7	11

Table 2: Statistical Tasks with Mean Rating of 2.50 of Greater Sorted by Mean Rating

	Status of Participant								
	Ph.D. Student			Faculty			Total		
	Mean	N	sd	Mean	N	sd	Mean	N	sd
16. Create useful graphs using software	3.08	13	1.038	3.94	17	1.029	3.57	30	1.104
26. Make real-world decisions consistent with data analysis	3.10	10	1.197	3.65	17	.931	3.44	27	1.050
20. Apply a statistical formula using hand calculator	2.69	13	1.109	3.59	17	1.064	3.20	30	1.157
29. Make a persuasive argument based on statistical analysis	2.90	10	1.197	3.29	17	1.213	3.15	27	1.199
12. Create and manage data files using a spread sheet or "flat ASCII" file	2.62	13	1.387	3.47	17	1.231	3.10	30	1.348
1. Formulate a problem that can be addressed with data	2.31	13	1.109	3.47	19	1.073	3.00	32	1.218
8. Comment on the shape of a distribution	2.62	13	1.044	3.24	17	1.091	2.97	30	1.098
2. Identify an appropriate sample statistic	2.54	13	.877	3.21	19	1.134	2.94	32	1.076
14. Identify unusual observations	2.46	13	1.266	3.29	17	1.047	2.93	30	1.202
3. Determine operational definition	2.69	13	.947	3.00	18	1.138	2.87	31	1.056
17. Identify misleading graphs	2.46	13	1.330	3.18	17	1.334	2.87	30	1.358
27. Write report describing statistical procedures	2.50	10	1.269	3.06	17	1.249	2.85	27	1.262

24. Make statistical inferences based on data and a model	2.17	12	1.193	3.13	16	1.147	2.71	28	1.243
30. Use statistical results to aid in developing a strategic plan	2.30	10	1.160	2.94	17	1.391	2.70	27	1.325
7. Identify extreme observations	2.38	13	1.193	2.88	17	1.166	2.67	30	1.184
19. Perform calculations using statistical software package	2.08	13	1.115	2.88	17	1.317	2.53	30	1.279
18. Identify an appropriate statistical model to evaluate chance as a possible explanation for an outcome	2.15	13	.987	2.82	17	1.237	2.53	30	1.167

Table 3: Statistical Course Topics with Mean Rating of 2.50 or Greater Sorted by Mean Rating

	Status of Participant								
	Ph.D. Student			Faculty			Total		
	Mean	N	Std. Deviation	Mean	N	Std. Deviation	Mean	N	Std. Deviation
32. Graphing	4.00	10	1.491	4.47	15	1.125	4.28	25	1.275
34. Measures of variability	4.00	11	.894	4.27	15	1.100	4.15	26	1.008
33. Measures of central tendency	3.82	11	1.079	4.20	15	1.082	4.04	26	1.076
63. Normal distribution	3.36	11	1.120	4.00	15	1.512	3.73	26	1.373
35. Measures of correlation, covariation, and proportion of variance explained	3.09	11	.831	4.00	15	1.414	3.62	26	1.267
50. Simple Regression Analysis	3.18	11	.982	3.53	15	1.552	3.38	26	1.329
51. Correlation analysis	2.55	11	1.128	3.80	15	1.568	3.27	26	1.511
41. Confidence intervals	2.73	11	1.009	3.47	15	1.727	3.15	26	1.488
36. Sample proportions.	2.18	11	.874	3.60	15	1.724	3.00	26	1.575
52. Multiple Regression analysis	2.55	11	.820	3.00	15	1.464	2.81	26	1.234
31. Scales of measurement	2.36	11	1.286	3.14	14	1.748	2.80	25	1.581
42. Hypothesis testing	2.18	11	1.079	3.20	15	1.859	2.77	26	1.632
43. t-tests	2.27	11	1.009	3.13	15	1.685	2.77	26	1.478
64. t- distribution	2.18	11	.874	3.13	15	1.727	2.73	26	1.485
39. Sampling Distributions	2.09	11	1.044	3.00	13	1.414	2.58	24	1.316

4. Discussion

4.1 Respondent Behavior

The results for the response to the survey invitation are of interest. No doubt, having the Associate Dean co-author the request worked to encourage accessing the *Qualtrics* web site. On the other hand, completion of the anonymous survey can be viewed as showing interest in statistical task performance and statistical reasoning. Of the 62 individuals who started the survey, the usable data provided by 31 respondents and 11 individuals who indicated they were not involved in teaching courses that had a statistics prerequisite are encouraging. Based on this result, over two-thirds of the respondents appear to have been seriously engaged in the survey. The results also show the value of a friendly reminder since half of the survey starts occurred immediately following the reminder.

The time spent on the survey shows that the demands were reasonable. The time spent can be used when distributing future surveys to indicate the amount of time an individual can expect to spend on the survey completion task. The lesser amount of time spent for the survey by the Ph.D. students compared to the faculty may show that the life of a graduate student is indeed filled with time pressures. Interestingly, a comment received from a colleague after the completion of data collection is that the relatively large number of items on the instrument discouraged some participation.

4.2 Statistical Tasks

A primary reason for conducting this survey was to flesh out what is meant by statistical literacy. Clearly, business students who have taken introductory statistics are expected to require only *Minimal Guidance* when they construct graphs, make decisions based on statistical results, develop arguments using statistics, and formulate problems that can be addressed with data. A testable hypothesis is whether they are expected to perform such tasks upon graduation with more than minimal guidance. In addition, students should have some skill in using a hand calculator and be able to work with data files. The data support our view, Kriska, Fulcomer, and Sass (2006), that understanding statistical calculation has value in introductory statistics, but statistical reasoning involves more than calculation.

The results for statistical tasks are also consistent with definitions of statistical literacy that have been proposed by statistical educators. Li's (2005) emphasis on appreciating the role of statistics in decision making is in harmony with the high rating given to the making real world decisions consistent with data analysis. Walman (1993) defines statistical literacy to include an appreciation for the use of statistics in decision making and the same survey item appears consistent with Walman's definition. Similar conclusions for these items follow when considering Rumsey's (2002) emphasis on thinking critically and making good decisions based on statistical information, and Gal's (2002) definition that includes interpreting and critically evaluating statistical information.

Kriska et al. (2006) note that several statistics educators emphasize the importance of statistical consumerism. The idea is that to function in today's world an individual must understand statistical reasoning. The mean ratings above 2.5 suggest performing with minimal guidance shows a number of expectations, e.g., graphing, decision making, making arguments, and report writing, can be interpreted as requiring students to be sophisticated consumers of statistical information without an expectation of a high skill level. The lack of mean ratings greater than 4.0 (anchored by the word *independent*) suggests there are no expectations for students leaving introductory statistics to be skilled statisticians. Because the focus of the survey was on undergraduate business students, the results seem reasonable. On the other hand, the need to make decisions consistent with data analytic results suggests that the consumer be sophisticated.

4.3 Statistical Course Topics

One of the aims of the survey was to identify topics most important to students when they take other courses. Clearly, the results show there is an emphasis on basics. The highest cognitive expectations were related to topics often covered early in the introductory statistics class, topics such as, graphing, central tendency, variation, correlation, and regression. The use of a cognitive processing type of scale helps to illustrate that the expectations in these areas are more than memorizing than the concepts, but aimed at comprehending the topics.

Two of the specific topics attract attention. For hypothesis testing, often a major focus in introductory statistics, there is much variability in the ratings and substantial differences between the Ph.D students' ratings and faculty's ratings. As with all of the results, the course the instructors teach will affect the pattern of results. For example, some academic areas may be more concerned with estimation of parameters rather than demonstrating statistical significance; and those instructors who primarily use statistical estimation may have no expectations for hypothesis testing. Hence, the large degree of variability in the ratings obtained here. In addition, the faculty may be associated with higher level course in which the expectations for statistical analysis, including hypothesis testing, are greater and therefore, the difference between the two groups.

The second topic of great interest is sampling distributions. We see sampling distributions as a challenging topic for students, but also as the foundation for statistical inference. The Ph.D. students rated the topic quite close to the *Rote Memory* anchor while the faculty rated it close to the *Comprehension* anchor. We certainly believe that students need to understand sampling distribution to understand statistical reasoning. While the Ph.D. students' ratings are not reflective of our expectations for students leaving introductory statistics, the results suggest that an appreciation for the sampling distribution concept develops while one is in graduate school.

4.4 Future Research and Final Comment

The aim of this survey was to identify task expectations for students taking courses following introductory statistics along with identifying statistics course topics with the greatest cognitive demands in the courses following introductory statistics. The mean ratings provided here provide empirical evidence about these expectations.

Perhaps the most important point to make about the survey is that the sample is not composed of statistical educators. Clearly business faculty and Ph.D. students have taken statistics courses and can be presumed to be users of statistical procedures and information, but their primary interest is not teaching statistics. Nevertheless, the results reported here support the views of statistical educators who have put much effort into defining and studying statistical literacy.

Opportunities for further investigations include examining statistical demands in the world of work, comparing business school expectations to other kinds of disciplines, and surveying students near or after graduation about the utility of their statistical training. Surveys could easily be designed using the instrument used here or a variation of it.

Acknowledgments

We sincerely appreciate the efforts of , Rao Unnava, the Associate Dean for Undergraduate Education in Fisher in helping us obtain an IRB exemption.

References

Gal, I. (2002). Adults' statistical literacy: meanings, components, responsibilities. *International Statistical Review, 70*(1), 1-51.

Kriska, S. D. Fulcomer,. M. C. Sass. M. M. (2006). Expectations for statistical literacy: a comparison among psychology, business, and public health professions. *Proceedings of the American Statistical Association, Statistics Education Section (CD-ROM).* Alexandria, VA: American Statistical Association.

Li, J. (2005). Statistics education for junior high schools in china. Proceedings of the International Association for Statistical Education (IASE) Roundtable on "Curricular Development in Statistics Education", 2004 *Voorburg, the Netherlands*: *International Statistical Institute*.

Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education, 10*(3).