THE VARYING DIFFICULTY ACROSS TOPICS (I.E., CHAPTERS) IN SELECTED MARKETING TEXTS

John R. Dickinson
University of Windsor
MExperiences@bell.net

ABSTRACT

In the realm of educational measurement considerable research has focused on item analysis, most fundamentally item difficulty. This over a period of decades. Little research has focused on the difficulty of the object of measurement. In a practical context, the present research investigates the variability of difficulty across topics (i.e., chapters) in selected marketing texts. That variability is found to be considerable with implications for professors and authors.

INTRODUCTION AND PURPOSE

With experience, instructors may come to judge the difficulty of the various topics comprising their courses. Personal assessment, student queries, performance on exam questions plus, no doubt, additional inputs might inform those judgements. The present study puts forth a straightforward, simple even, approach to more formally, systematically, and quantitatively making those judgements. The obvious implication of the outcomes is for instructors to give more attention, in whatever forms, to the more difficult topics and less attention to the less difficult topics. Authors, likewise, might heed the results in revising their textbooks.

For the present research, "topic" is operationalized as a chapter in the relevant textbook. The "topic" or "chapter" basis for the analyses here does not qualify as item analysis. Considerable research has focused on the difficulty of individual items or questions. Item analysis refers to the evaluation of items, i.e., questions, comprising tests. Its purpose is, "...toward the determination of the best possible items for inclusion in a test." (Rogers 1995, p. 388) The most elemental property of an item is its difficulty. "The classical approach [to item analysis]...begins by computing difficulty..." (Millman & Greene 1989, p. 358) Though seemingly an obvious and easy exercise, little research has analyzed the difficulty of course topics.

More clearly stated, the present research measures the difficulty across chapters in selected marketing textbooks.

DATA

Samples of questions from banks accompanying six texts were drawn. Among the six were two editions of a consumer behavior text plus a second consumer behavior text and three editions of a retailing management text. The texts, the total number of multiple-choice questions in the respective banks, and the number of questions sampled from each question bank are reported in Table 1.

TABLE 1
BANK AND SAMPLE QUESTION COUNTS

Text	Total Questions	Sample Questions (percent of total)
Levy, Weitz, & Grewal (2014, LWG), *Retailing Management*, Ninth Edition	1210	612 (50.6)
Levy & Weitz (2012, LW), *Retailing Management*, Eighth Edition	1190	624 (52.4)
Solomon, Zaichkowsky, & Polegato (2011, SZP), *Consumer Behaviour*, Fifth Canadian Edition	1148	671 (58.4)
Levy & Weitz (2009, LW), *Retailing Management*, Seventh Edition	1332	736 (55.3)
Solomon, Zaichkowsky, & Polegato (2008, SZP), *Consumer Behaviour*, Fourth Canadian Edition	1019	674 (66.1)
Hawkins, Mothersbaugh, & Best (2007, HMB), *Consumer Behavior*, Tenth Edition	1624	958 (59.0)

EXAMINATIONS

Providing data for the present analyses were undergraduate courses typically taken in the third year of a student's university program, the courses having as prerequisites two semester-long principles of marketing courses. For each class the first midterm exam covered about the first third of the chapters, the second midterm exam covered about the middle third of the chapters, and the noncumulative final exam covered about the last third of the chapters. Each of the exams counted for 20 percent of the students' final course grades. Exams were scored as the percent of questions answered correctly; no penalty was deducted for incorrect answers. Questions not answered were deemed incorrect for calculating exam scores and for the present research.

SAMPLING METHOD

Multiple-choice questions are arranged in the test question bank according to the order in which the question content appears in the textbook. For each examination, specific multiple-choice questions were selected on a systematic sampling basis. This systematic sampling approach was an attempt to ensure that:

- a cross section of each chapter content was included among the examination questions,
- all respective midterm and final examinations were of comparable composition, and
- a representative sample of the test bank questions was obtained.

Table 2 summarizes the overall data. All questions analyzed had five options: the correct answer plus four distractors.

TABLE 2
QUESTIONS, ANSWERS, CORRECT ANSWERS

Text	Total Questions	Total Answers [a]	Mean Answers per Question [b]	Percent Correct Answers [c]
LWG (2014), 9th	612	25259	41.27	76.38
LW (2012), 8th	624	23692	37.97	69.62
SZP (2011), 5th	671	28172	41.99	57.82
LW (2009), 7th	736	26615	36.16	66.89
SZP (2008), 4th	674	26947	39.98	60.99
HMB (2007), 10th	958	31309	32.68	62.63

a   Potential answers, i.e., including omitted answers.
b   Essentially the number of students or class size.
c   Percent of potential answers, i.e., omitted answers are counted as incorrect.

ANALYSIS

Describing the difficulty of textbook chapters is straightforward. For the questions sampled randomly from a given chapter the students' scores, as the percent correct, based on those questions is calculated. That is the measure of chapter difficulty. Of relevance to the purpose of this study is the variability of that measure across the chapters comprising the textbook. In Table 3 are reported the minimum, maximum, range (=maximum-minimum), and standard deviation of those chapter difficulties.

RESULTS

TABLE 3
RANGE OF DIFFICULTY ACROSS TEXT CHAPTERS

Text	Chapters	Mean % Correct*	Minimum % Correct	Maximum % Correct	Range	Standard Deviation
LWG (2014), 9th	18	76.51	63.69	83.47	19.78	5.39
LW (2012), 8th	18	69.81	56.01	80.84	24.83	5.73
SZP (2011), 5th	17	57.86	50.43	66.03	15.60	3.79
LW (2009), 7th	19	66.87	53.61	75.59	21.97	4.97
SZP (2008), 4th	17	60.98	55.13	68.20	13.08	3.64
HMB (2007), 10th	20	62.60	55.21	73.35	18.15	4.64

*   These values differ slightly from "Percent Correct Answers" in Table 2. For Table 2 all questions were analyzed together as a single whole. Here questions were analyzed on a chapter basis and then the mean taken.

From Table 3, the difference in chapter difficulty can be as much as 24.83 percentage points, with four of the six texts being greater than 18 percentage points. Five of the six minimum percents correct are less than 60 percent, i.e., below a minimum standard for student advancement at some universities. The topics of those chapters are presumably candidates for greater attention in the classroom and for revision in subsequent editions of the textbooks.

It may also be seen in Table 3 that the ranges for the two editions of SZP (2008, 2011) are considerably smaller than the ranges for the three editions of LW(G) (2009, 2012, 2014). The former text is for consumer behavior and the latter is for retailing. The two textbooks are not substitutable, of course. By analogy, though, the difference in difficulty variability for the two texts suggests a criterion for choosing between two texts in the same subject area.

Results presented in Table 3 are based on single chapters. With this, it is possible that outlying single chapters may exaggerate the range of difficulty.

A second basis for describing difficulty, then, is to examine the three easiest chapters together and the three most difficult chapters together. These results, perhaps better characterizing the textbooks, are reported in Table 4.

TABLE 4
PERCENT CORRECT FOR THREE MOST EXTREME CHAPTERS

Text	Three Highest % Correct Chapters	Three Lowest % Correct Chapters	Difference
LWG (2014), 9th	82.16	66,34	15.82
LW (2012), 8th	78.46	60.72	17.74
SZP (2011), 5th	62.57	51.84	10.73
LW (2009), 7th	73.86	58.63	15.23
SZP (2008), 4th	66.74	56.63	10.11
HMB (2007), 10th	71.25	56.37	14.88

These results are for the three chapters analyzed as a single whole, rather than the mean of the three chapters analyzed separately.

As expected, the differences in Table 4 are less than the ranges in Table 3. Still, three of the six differences are greater than 15 percentage points with a fourth being nearly so. And four of the six lowest percents correct are less than 60.

DISCUSSION

Reported here are substantial differences in topic/chapter difficulty for a selection of marketing textbooks. With experience, of course, instructors may well come to adjudge the difficulty of specific topics. Here that judgement is complemented with a simple systematic approach to quantifying variations in difficulty. Such systematic quantification can promote several useful implications.

The most direct implication of the difficulty of a topic/chapter is to guide the emphasis, in whatever form, given to the topics by instructors. Another implication pertains to planning tests. Overall test scores can presumably be influenced by the proportions of exam questions based on the various topics/chapters. Consistent difficulty should make it easier for the instructor to select questions without concern for differences in difficulty. In contrast, if the chapters vary widely in difficulty then this might be taken into account by the instructor in selecting different numbers of questions so as to make the overall exam more or less difficult.

Another implication may be in the reviewing of courses. Differences in variability of test scores across classes may be attributable to the variability of the topics/chapters comprising the courses.

Topic difficulty in the present research is measured as the percent of questions answered correctly. Topics *per se* may be inherently difficult (e.g., rocket science) or easy (e.g., not rocket science). Here, though, topics *per se* are not studied. Rather, topics as presented in textbook chapters are studied. Of course factors other then the topic *per se* may affect difficulty as analyzed in this study. One obvious factor is the treatment of the topic in class by the professor. (For the present research, this is not a factor. The courses here are project courses with there being no lectures or other planned addressing.) Akin to treatment by the professor, a second factor is the manner in which the topic is presented in the textbook. At the time of use, the textbook may be seen as a given. In future editions authors might give consideration to difficulty when revising chapters (or in the case of new textbooks composing the chapters originally). There is no absolute imperative that chapters be of equal difficulty. The consideration would be whether unequal difficulty is attributable to the topic *per se* or to artifacts attending the topic presentation.

The second component of topic difficulty is "difficulty" and the measurement thereof. Thus, yet a third factor here may lie in the multiple-choice questions themselves. With wording of the question stems and question answer alternatives, the questions may be composed, intentionally or not, as being of varying difficulty. As to this third factor, the multiple-choice questions were selected from published question banks accompanying the texts. All those published banks classify questions as Easy, Medium, or Hard. It may be that difficulty as measured here is affected by the mix of Easy, Medium, and Hard questions randomly selected to comprise the exams. More specifically, an apparently difficult topic/chapter here may have, say, a greater proportion of questions classified as Hard than an apparently easy topic/chapter. Research into this possibility is presently underway.

A final implication lies in research in course development and administration. Consider a simple experimental design investigating, say, three course formats: lectures, lectures with tutorials, and self-study. Differences in mean scores on multiple-choice exams common to the three formats might be analyzed with a basic one-way analysis of variance (ANOVA). For, say, the first midterm exam the test might comprise six chapters. Within each treatment condition, varying chapter difficulties could

contribute to unexplained error resulting in a finding of statistical insignificance. It does not matter that the same chapters appeared under all three conditions. The unexplained error would remain.

An example of this possibility is two recent studies that examined the effect of three different multiple-choice question orderings on exam scores. One approach (Dickinson 2018a) used a one-way ANOVA with score on the total exam being the dependent variable. That study found no significant effect of the orderings.

As a precaution against the effect of unexplained error just noted, another analysis of the same data (Dickinson 2018b) took a "closer look" by analyzing scores on individual chapters, thereby eliminating the varying difficulty, i.e., unexplained error, of the chapters comprising the exam. No significant effect on exam-chapter scores was found.

The findings of the total-exam-score and chapter-score analyses are consistent. This is not due to there being little or no unexplained error due to varying chapter difficulty in the former analysis. As attested to in the present research, such unexplained error *is* present. Rather, the chapter-score analysis which precluded that unexplained error affirms that the question orderings do not have a significant effect.

REFERENCES

Dickinson, John R. (2018a). The effect of jumbling multiple-choice questions on exam scores (abstract). *Proceedings*, 2018 Annual Meeting of the Decision Sciences Institute Conference, 1709.

Dickinson, John R. (2018b). A closer look at the effect of jumbling multiple-choice questions on exam scores (extended abstract). In L. Lindgren, L. Samii, and U. Sullivan (Eds.), *Fall 2018 Educators' Conference Proceedings*, Marketing Management Association, 51-52.

Hawkins, Del I., Mothersbaugh, David L., & Best, Roger J. (2007). *Consumer Behavior*, Tenth Edition. Boston: McGraw-Hill Irwin.

Levy, Michael, Weitz, Barton A., & Grewal, Dhruv (2014). *Retailing Management*, Ninth Edition. New York: McGraw-Hill Education. ISBN 978-0-07-802899-1, MHID 0-07-802899-1

Levy, Michael & Weitz, Barton A. (2012). *Retailing Management*, Eighth Edition,. New York: McGraw-Hill Irwin.

Levy, Michael & Weitz, Barton A. (2009). *Retailing Management*, Seventh Edition. New York: McGraw-Hill Irwin. ISBN-13: 978-0-07-338104-6, ISBN-10: 0-07-338104-7

Rogers, Tim B. (1995). *The Psychological Testing Enterprise: An Introduction*. Pacific Grove, CA: Brooks/Cole Publishing Company. ISBN: 0-534-21648-X

Millman, Jason & Greene, Jennifer (1989). The specification and development of tests of

achievement and ability. In Linn, Robert L. (Ed.), *Educational Measurement*, Third Edition. New York: American Council on Education and Macmillan Publishing Company, 335-366. ISBN: 0-02-922400-4

Solomon, Michael R., Zaichkowsky, Judith L., & Polegato, Rosemary (2011). *Consumer Behaviour*, Fifth Canadian Edition. Toronto: Pearson Prentice Hall. ISBN: 978-0-137-01828-4

Solomon, Michael R., Zaichkowsky, Judith L., & Polegato, Rosemary (2008). *Consumer Behaviour*, Fourth Canadian Edition. Toronto: Pearson Prentice Hall. ISBN-13: 978-0-13-174040-2, ISBN-10: 0-13-174040-7