



Using SPSS for item analysis

*More reliable test assessment
using statistics*

Item analysis is a series of procedures that are easily implemented with SPSS software

Instructors frequently assess student learning with multiple-choice question examinations. These examinations are efficient tools because they do not take long to grade. Multiple-choice questions, given their presentation of alternative answers, also called distracters, are also very effective learning assessment tools as well.

Instructors who develop their own examinations can enhance the effectiveness of each test item, or question, if they select and write their items on the basis of item performance data. Such data are available to instructors who submit their test items to item analysis – a series of procedures that are easily easily implemented with SPSS software.

Administration

Administrative tasks such as test composition, administration and answer collection can be accomplished with the SPSS line of products. Here are some examples:

- **Remark Office OMR** enables you to design your own test forms using any word processor, print them on any printer, scan them with your desktop scanner and export the data to an SPSS data format. Remark reads OMR (bubble fields, check boxes, etc.), barcodes and image fields.
- **Teleform 5** makes it fast and easy to design your own test forms. With Teleform Tri-CR technology you get three sophisticated recognition engines, so you get near-perfect data accuracy for hand print recognition, optical character recognition and optical mark reading. You can easily correct mismarked or illegible responses with the Teleform Verifier, and store your data as an SPSS data file to get a quick start on data analysis.
- **SPSS Data Entry 1.0** provides a “drag-and-drop” form design feature and enables either computer-aided test administration (via the Data Entry Station software package) or test administration over the Web. Either way, the data are captured in an SPSS data format so it’s immediately ready for analysis.

Scoring

Once responses to multiple choice questions are entered into SPSS, scoring the results against the answer key is a straight-forward procedure within SPSS. For example, Figure 1a shows a typical test answer file – where each row represents a particular student’s answers to a 10-answer multiple choice exam. The last row is the answer key.

The next figure (Figure 1b) shows the same set of answers, but this time coded “1” (correct) and “0” (wrong). Additionally, each student’s score (percentage of correct answers) has been recorded – as well as how that student compared (top, middle or bottom one-third) with other students in the class.

	name	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10
1	1221	3	2	1	2	3	3	4	5	1	
2	1229	6	3	4	4	3	2	3	1	3	
3	1234	1	3	6	1	2	2	3	1	6	1
4	1345	2	2	1	1	3	4	3	3	2	4
5	1439	2	5	1	1	3	4	2	1	5	5
6	2214	3	1	4	8	9	1	6	1	2	2
7	2245	1	2	1	5	2	4	3	1	2	2
8	3341	5	4	1	2	3	4	3	1	5	1
9	key	2	4	1	1	3	4	3	1	6	1

Figure 1a
A typical test answer file

SPSS is a reliable tool to help you accurately compute Index of Discrimination

	name	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	score	group
1	1221	0	0	1	0	1	0	1	0	1	1	50.00	middle
2	1223	0	0	0	0	1	0	1	1	0	0	30.00	low
3	1234	0	0	0	1	0	0	1	1	1	1	60.00	middle
4	1345	1	0	1	1	1	1	1	0	0	0	60.00	middle
5	1433	1	0	1	1	1	1	0	1	1	0	70.00	high
6	2214	0	0	0	0	0	0	0	1	0	0	10.00	low
7	2215	0	0	1	0	0	1	1	1	0	0	40.00	low
8	3341	0	1	1	0	1	1	1	1	1	1	80.00	high

Figure 1b
A typical test answer file with answers coded correctly (1) and incorrectly (0)

The Index of Discrimination

SPSS offers reliable computation of the Index of Discrimination. One example of a measure of effectiveness for a particular test item is the difference between the percentage of students in the top one-third of the class who correctly answered that test item and the percentage in the bottom one-third who correctly answered the item. Figure 2 shows the result.

This index should be viewed in the context of the overall difficulty of the question. For example, the Index of Discrimination might not be useful for a question that only 10 percent of the overall class answered correctly. In most circumstances, however, viewing the correct percentage in each one-third of the class (as in Figure 2) is useful in and of itself. SPSS produces these results accurately and efficiently.

	question	high	middle	low	discrim
1	Q1	.50	.33	.00	.50
2	Q2	.50	.00	.00	.50
3	Q3	1.00	.67	.33	.67
4	Q4	.50	.67	.00	.50
5	Q5	1.00	.67	.33	.67
6	Q6	1.00	.33	.33	.67
7	Q7	.50	1.00	.67	-.17
8	Q8	1.00	.33	1.00	.00
9	Q9	1.00	.67	.00	1.00
10	Q10	.50	.67	.00	.50

Figure 2
Index of Discrimination

Other statistical items of importance

Ideally, one would want the “correctness” of a particular item on a test to be associated with a high overall score (percentage of all items answered correctly). Taken from a different angle, you may prefer “frequently answered correctly” test items from students who did well on the test overall. This association between individual test item and overall test performance is called the point biserial correlation. A more useful correlation is the overall test performance computed excluding the particular test item in question. This measure is called the corrected point biserial correlation of a test item.

SPSS provides a measurement of the internal consistency of the test items

The screenshot shows the SPSS Output window for the Reliability program. It displays a table of Item-total Statistics and Reliability Coefficients. The table includes columns for Item, Scale Mean if Item Deleted, Scale Variance if Item Deleted, Corrected Item-Total Correlation, Squared Multiple Correlation, and Alpha if Item Deleted. The Reliability Coefficients section shows the overall Alpha as .5735 and the Standardized item alpha as .5652.

Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
Q1	4.6250	4.2679	.2614	.	.5460
Q2	4.7500	4.2143	.4429	.	.5148
Q3	4.2500	3.6429	.5423	.	.4632
Q4	4.5000	4.2857	.2000	.	.5625
Q5	4.2500	3.9286	.3930	.	.5114
Q6	4.3750	3.6964	.4845	.	.4783
Q7	4.1250	4.6964	-.0356	.	.5989
Q8	4.1250	5.5536	-.3401	.	.6801
Q9	4.3750	3.6964	.4845	.	.4783
Q10	4.5000	4.2857	.2000	.	.5625

Reliability Coefficients 10 items
Alpha = .5735 Standardized item alpha = .5652

Figure 3
Output for the SPSS Reliability program

SPSS gives test developers a means of measuring consistency. SPSS provides a measurement of internal consistency (reliability) of the test items called Cronbach's Alpha. The higher the correlation among the items, the greater the alpha. High correlations imply that high (or low) scores on one question are associated with high (or low) scores on other questions. Alpha can vary from 0 to 1, with 1 indicating that the test is perfectly reliable. Furthermore, the computation of Cronbach's Alpha when a particular item is removed from consideration is a good measure of that item's contribution to the entire test's assessment performance. In the SPSS output above (Figure 3), these and other statistics are automatically generated using a procedure called Reliability (in the output, the term scale refers to the collection of all test items).

Note the column "Corrected Item-Total Correlation." This column displays the corrected point biserial correlation. You can see that question three seems to correlate well with overall test performance. On the other hand, students who tend to do poorly on the test overall tend to answer question eight correctly. This is not a desired outcome. Notice also the "Alpha" in the lower left-hand corner and the column titled "Alpha if Item Deleted." If question eight is deleted from the scale, the Alpha statistic climbs to .68. With this perspective, question eight needs to be critically examined and perhaps rewritten.

Distracter analysis

In a multiple-choice question, the incorrect choices are called "distracters." When creating test questions, consider the following:

- A percentage of students should select each distracter (in lieu of the correct answer) or the distracter is not effective.
- If too great a percentage of students select a particular distracter, there might be an ambiguity in the wording of the question or the wording of that particular distracter.

- A distracter has value if the percentage of students selecting it differed based on their overall test performance. A good distracter is one that is selected by few students who were in the top third of the class, but chosen by many students in the bottom third.

Building a table that provides this type of analysis for each test item is, again, an easy procedure in SPSS. Figure 4 shows the distracter analysis output for question two of our test.

			Q2				
			1	2	3	4*	5
Percentage	GROUP	1 high 1/3rd				66.7%	50.0%
Answering		2 middle 1/3rd		66.7%	33.3%		
Distractor		3 low 1/3rd	33.3%	33.3%	33.3%		
Group Total			12.5%	37.5%	25.0%	12.5%	12.5%

Figure 4
Distracter Analysis

In the output, the four distracters are displayed across the columns, along with the correct answer to the question (answer 4) marked with an asterisk. Notice the following:

- Overall, 12.5 percent of the students chose distracter five. However, among the top one-third of the performers on the test, 50 percent chose that distracter. In the middle and bottom thirds, no one chose distracter five. This is an indication that there is perhaps an ambiguity in the wording of the question – or the distracter – which was apparent to the better performers.
- Overall, 12.5 percent of the test takers chose distracter one. In the top and middle thirds, no one selected it. And, in the bottom one-third, 33 percent selected distracter one. It seems that this distracter is performing as it should.

Item analysis enables instructors to improve their classroom practices and allows item writers to enhance their tests

Productivity-enhancing tools

SPSS is as flexible as it is powerful and can be adapted to your teaching environment. Consider the following:

- SPSS makes test administration tasks more productive. SPSS can be used in the interactive mode wherein the kinds of tables and data transformations illustrated in this paper can be accomplished on the fly. This method is useful for researchers who are interested in more in-depth analyses. An intuitive Windows interface makes these tasks simple – even for first-time SPSS users.
- SPSS can be customized for increased flexibility. When appropriate, each of these procedures can be easily added to the menu bar.
- SPSS can also run in production mode. This is useful, for example, if you are interested in running the same item analysis procedures every time you administer a test.

Summary

Item analysis is an extremely useful set of procedures available to teaching professionals. SPSS is a powerful statistical tool for measuring item analysis and an ideal way for educators to create – and evaluate – valuable, insightful classroom testing tools. Item analysis enables instructors to improve classroom practices and enables test writers to enhance their examinations. SPSS is an ideal tool in accurate item analysis.

About SPSS

SPSS Inc. is a multinational software company that delivers “Statistical Product and Service Solutions.” Offering the world’s best-selling desktop software for in-depth statistical analysis and data mining, SPSS also leads the markets for data collection and tabulation.

Customers use SPSS products in corporate, academic and government settings for all types of research and data analysis. The company’s primary businesses are: SPSS (for business analysis, including market research and data mining, academic and government research); SPSS Science for scientific research; and SPSS Quality for quality improvement.

Based in Chicago, SPSS has offices, distributors and partners worldwide. Products run on leading computer platforms, and many are translated into Catalan, English, French, German, Italian, Japanese, Korean, Russian, Spanish and traditional Chinese. In 1997, the company employed nearly 800 people worldwide and generated net revenues of approximately \$110 million.

Contacting SPSS

To place an order or to get more information, call your nearest SPSS office or visit our World Wide Web site at www.spss.com

SPSS Inc.	+1.312.651.3000	SPSS Israel Ltd.	+972.9.9526700
	Toll-free: +1.800.543.2185	SPSS Italia srl	+39.51.252573
SPSS Argentina srl.	+541.814.5030	SPSS Japan Inc.	+81.3.5466.5511
SPSS Asia Pacific Pte. Ltd.	+65.245.9110	SPSS Kenya Ltd.	+254.2.577.262
SPSS Australasia Pty. Ltd.	+61.2.9954.5660	SPSS Korea KIC Co., Ltd.	+82.2.3446.7651
	Toll-free: +1800.024.836	SPSS Latin America	+1.312.651.3226
SPSS Belgium	+32.162.389.82	SPSS Malaysia Sdn Bhd	+60.3.704.5877
SPSS Benelux	+31.183.636711	SPSS Mexico Sa de CV	+52.5.682.87.68
SPSS Central and Eastern Europe	+44.(0)1483.719200	SPSS Middle East & South Asia	+91.80.545.0582
SPSS Czech Republic	+420.2.24813839	SPSS Polska	+48.12.6369680
SPSS East Mediterranean & Africa	+972.9.9526701	SPSS Russia	+7.095.125.0069
SPSS Federal Systems	+1.703.527.6777	SPSS Scandinavia AB	+46.8.506.105.50
SPSS Finland Oy	+358.9.524.801	SPSS Schweiz AG	+41.1.266.90.30
SPSS France SARL	+33.1.5535.2700	SPSS Singapore Pte. Ltd.	+65.533.3190
SPSS Germany	+49.89.4890740	SPSS South Africa	+27.11.706.7015
SPSS Hellas SA	+30.1.7251925	SPSS Taiwan	+886.2.25771100
SPSS Hispanoportuguesa S.L.	+34.91.447.37.00	SPSS UK Ltd.	+44.1483.719200
SPSS Hong Kong Ltd.	+852.2.811.9662		
SPSS Ireland	+353.1.496.9007		