

Results Interpretation Guide¹

for the

Student Perceptions of Teaching Effectiveness Inventory (SPTE)

General Information

Audience

This interpretation guide was written for the instructor using SPTE, as well as departmental chairs and other faculty or administrators who may be reviewing SPTE results. This guide is intended to supplement the Results Guide. Throughout the Results Interpretation Guide, numbers are shown in parentheses and in bold [e.g., **(3)**]; these correspond to the locator numbers in the Results Guide. For individuals interested in more information concerning statistical techniques used to generate SPTE results, a technical guide is available.

What is SPTE?

The SPTE is a measure of students' perceptions of teaching and course quality which may be used both for summative (evaluative) and formative (improvement) purposes. When used for summative purposes, it should be used in conjunction with other evaluative materials. The SPTE is a norm-based instrument which includes 39 evaluative items from which six primary and two summary factors were derived using a statistical technique called factor analysis. Results for each of the eight factors are reported relative to two separate norm bases: 1) norms in which the instructor's course ratings are given relative to the ratings for all university courses collected over several years, and 2) norms in which the instructor's course ratings are reported relative to the ratings for all courses in the instructor's College/Division again collected over several years. Each factor is reported using a Scale Score, as well as a Percentile and an Interpretive Profile. In addition, item results are given, with raw scores, Scale Scores, Percentiles, and finally an Interpretive Profile.

Bias Corrections

All scores are corrected for two sources of bias: (a) number of raters (roughly the instructor's class size) and (b) *a priori* student motivation. This correction results in NO correlation between the SPTE Factors/Items and either the number of raters or our measure of student *a priori* motivation which consists of four inventory items.

Validity Items

To minimize random marking, SPTE contains two validity items. If a student fails to respond correctly to either of these items, his/her entire inventory is discarded. We reject about 1% of all completed SPTEs for this reason.

SPTE Factors

SPTE scoring produces six primary and two second order factors which summarize the items on the questionnaire. Because these factors consist of a weighted combination of the items, they are more stable and reliable than any individual item. An instructor's scores on these eight factor-analytically derived dimensions are included in SPTE results.

Primary Factors

- **Course Design:** Measures the extent to which the instructor is perceived as delivering well-prepared lectures, as organized in her/his approach to the class, as having adequately conveyed concepts, and as having utilized appropriate methods of presentation.
- **Rapport with Students:** measures if the instructor was perceived as having created an atmosphere in the classroom that was conducive to learning. Did the instructor come across as a person as well as a teacher? Was he/she generally respectful of students, and concerned and actively helpful regarding students progress in their course? Did the instructor seem enthusiastic about the course? Were students encouraged to express their opinions freely?
- **Grading Quality:** assesses students' opinions about the number and validity of evaluations used in determining grades and the explicitness of the method used to assign grades. It also evaluates the students' perceptions of congruence between their expected grade and self-assessed course performance.

¹ © 1996, 2003, 2005 by the Social Science Research Lab, Wichita State University

- **Course Value:** not only indicates how valuable the students considered the course, but reflects whether the student reports having become more knowledgeable and interested in the subject because of the course and would also recommend the course to others.

These four factors are useful for both summative and formative purposes, providing information about not only the perceptions of the current course, but possible areas that might be adjusted.

- **Course Difficulty:** reflects difficulty as related to rate of coverage, and the level and amount of material.
- **Workload:** measures the number and difficulty of the assignments, especially out-of-class work.

These two factors can be useful in assessing the overall course demands and are primarily formative in nature. In other words, “hard” is neither good nor bad.

Summary Factors

- **Perceived Quality Index (PQI):** is an empirically derived composite of **Course Design, Rapport With Students, Grading Quality,** and **Course Value.** It may be treated as an overall rating of perceived teaching effectiveness.
- **Perceived Course Demands:** includes two factors: **Course Difficulty and Workload.** It is an overall assessment of course demands on the students. It is generally unrelated to quality and tends to vary between courses even with the same instructor.

There is essentially no relationship between the scores on these two second order factors ($r = .04$).

Types of Scores Provided by SPTE

- **Item Means:** The actual means for SPTE items are provided. On the SPTE Inventory Form, the socially desirable choice for each item is randomized to avoid response sets. (One student once suggested to us that if we made the high score on each item “good,” the form would be much easier to fill out.) For presentation, we have adjusted items so that a “high” score is always either “good” or “hard.”
- **Scale Scores:** When the SPTE scores are calculated, each instructor has one score on each of the eight factors. All instructors’ scores are used to create the norm base for all university courses.

Similarly all the scores for each College/Division are used to create the norm base for each College/Division. The individual instructor’s score is converted into two scores, one that represents the relative position of the instructor’s score when compared with all university courses and another that represents the relative position of the instructor’s score when compared with all courses in the instructor’s College/Division. To aid in interpretation, the scores are adjusted to have a mean of 5.5 and a standard deviation of 2. With this information, you can discern how your scores compare to the mean of all scores in the norm base.

Scale Scores allow comparisons among the eight factors. A 5.7 on one factor is directly comparable to Scale Score of 4.9 on another. The instructor’s performance on the first factor was relatively higher than on the other factor. It also permits comparisons from semester to semester.

- **Percentiles:** We provide a Percentile score for each factor and item. The Percentile gives you the percentage of class scores that yours was as high or higher. For example, an instructor who obtains a Percentile of 46.3% on the All University Courses norm base on the Perceived Quality Index (PQI) should interpret that as follows: “The instructor’s PQI score was rated equal to or higher than 46.3% of all the courses in the university.” While Percentiles are the easiest to understand, they are the least useful of the scores provided.
- **Interpretive Profile:** The Social Science Research Lab (SSRL) has completed a study on the relationship of a student’s open ended comments and the numerical factors from SPTE. We (the lab personnel, another faculty member, and 12 doctoral students) read and evaluated comments from 208 classes (14,313 comments in all). Each comment was rated on a 1 through 5 valence scale with 1 being an extremely negative comment and 5 a very positive one. We then computed the mean valence for each class.

The mean valences from these classes were then matched with their SPTE Factor Scores. The correlation between these valences and PQI was $+0.8$ indicating a high relationship between the two.

We also noticed that the SPTE quality factor scores (Course Design, Rapport, Grading Fairness, Course Value, and the Perceived Quality Index) are notably negatively skewed. In other words a lot more classes have scores on the “good” end than on the “bad”

end. Thus, interpretation of the SPTE results has probably been more negative than warranted especially if the Percentile scores were used.

To correct this, we have related Scale Scores on each factor to actual comments from our comment database². We have then produced four ranges based upon comments that we have labeled “Low,” “Good,” “Very Good,” and “High.” We have applied this procedure to all factors and items and report them that way.

Sample comments from classes receiving “Low,” “Good,” “Very Good,” and “High” scores on the PQI follow. Instructors scoring in the High range may occasionally see negative comments and those receiving a “Low” may receive occasional good comments. While the relationship of the PQI scores with comments is quite high ($r=.8$), it isn't perfect.

PQI scores that fall in the “Low” range can expect a set of comments like the following:

Some of us are really trying and it seems like it's never ever good enough.
I am sorry you don't seem to enjoy what you teach.
It must make life rough.
I felt I was robbed from this class and what I could gain if another had taught.
I believe that I would have gotten more out of it if my dog taught the class.
Whenever anything went wrong in the class xxx blamed the computer.
I fear that because of the poor teaching of the instructor, that I am going to struggle through many classes in the future.
The instructor could use some tips on how to work with people.
I felt that the instructor was rude most of the time.

PQI scores that fall in the “Good” range can expect a set of comments like the following:

She/he's nice
Make her lectures more interesting
Very friendly
When they accuse someone of academic dishonesty have proof!
She/he could have graded more effectively.
It challenged me to be one step ahead, instead of one step behind.
Have more than one major paper.

XXXX was very nice and understanding.
She/he was fun
Be clear about what is on the tests!

PQI scores that fall in the “Very Good” range can expect a set of comments like the following:

Well organized
Can tell he/she wants his students to learn the material, about XXX
Instructor is very well educated with the subject.
The instructor helped the students understand the material through his classroom lectures.
Wrote very good notes.
I couldn't get detailed enough answers to some of my questions.
The teacher was respectful.
That she/he was cool
The instructor covered the topics supposed to be covered very well with what the topics were suppose to be.
The instructor made learning fun and understandable.

PQI scores that fall in the “High” range can expect a set of comments like the following:

I truly appreciated your ideas and real-life experience.
His mastery of the subject matter was superb.
She/he related well with the students.
Get more organized.
(What would you change?) Nothing, she/he did a great job!!
I really enjoyed this class and was upset when I had to miss it.
She/he is exceptionally clear about what she/he expects from her students.
The instructor was wonderful.
He/she is enthusiastic and passionate about this subject which helps in focusing the students on the subject of XXX.
I don't think there is anything he/she could do to improve the course.

How to Interpret SPTE Results

Front Page

- **Identifying Information:** At the very top of the front page, you will find the SPTE title, with the corresponding semester and year immediately below (2). Two lines follow, each with three kinds of identifying information about the instructor's specific course (3). From left to right, the first line

² For each fall and spring semester, we add 25 classes to the comment database.
SPTE II

includes the Department in which the course was taught, the University at which the instructor teaches, and the ID number the instructor's course as assigned by the SSRL. The second line lists from left to right: a) the instructor's Course number (catalog number), b) the College/Division to which the instructor's Department is assigned, and c) the Section Number the course was assigned. (At W.S.U. the section number will be the number listed in the Schedule of Courses. At P.S.U., this number is arbitrary.) **Be sure that the instructor's department is correct. It determines the College/Division norm base used to produce the instructor's scores. If it is not correct, notify the W.S.U. Social Science Research Lab immediately (316-978-3384).**

- **Results Headings:** Immediately below the identifying information [(2) & (3)] are the headings for your results [(1), (6), & (7)]. There are three sections of information provided to you on this side of the page:

- (1) Immediately below the title, Interpretive Profile, you will find the number of students in your class successfully filling out SPTE forms.
- (6) The number of all university classes against which your scores are compared can be found here.
- (7) The number of college (division) classes used for comparison is located here.

- **All University Scores:** Under (6) you will find your scores compared to all University classes. The first column has the Scale Scores, the second, the Percentiles.
- **College/Division Scores:** Under (7) you will find your scores compared to your College/Division. The first column contains the Percentiles, the second Scale Scores.
- **Interpretative Profile for the Quality Scores: (4)** indicates the box which contains the Interpretative Profile for the quality factors of the SPTE. For each of the five factors, you will find either "Low," "Good," "Very Good," or "High.". How to understand these scores may be found above under **Interpretative Profile**.
- **Plotting of the Course Demands Factors:** We have not assigned Low, Good, etc. to the Course

Demands factors for two reasons. 1) Unlike the quality factors, the Course Demands factors are normally distributed. 2) There is little or NO relation between the Course Demands factors and the Perceived Quality Index factors. The difficulty and work load of courses may vary for many reasons unrelated to quality. For example, something would be odd or unusual if an undergraduate Child Psychology course was rated as more difficult than a graduate level research methods course.

For these reasons, the Course Demands factors (5) are plotted as Percentiles. For each of the factors, your Percentile score based on the entire university is plotted with a '■'. The Percentile based on College/Division scores is indicated with a '▼'.

Back Page

The back side of the SPTE results is intended for formative (improvement) uses ONLY. Use for summative purposes may be invalid and is not recommended.

This side of the page provides a summary of your scores on each item. The items are organized according to the factor to which they are related. The items do not appear in the same order as they do on the SPTE itself. In one case, an item appears on more than one factor. The factors and their associated items related to the Perceived Quality Index appear in the upper box (8), while the factors and their related items pertaining to Perceived Course Demands are presented in the lower box (9). Some items have been reversed so that higher raw scores and corresponding Scale Scores will reflect either a socially desirable score or that increased demands have been placed on the student.

Within each box, from left to right, four separate kinds of information about each item are reported.

- **Raw Score:** These are the item means which the instructor obtained on the items (10). It has a range of one to five, with higher numbers representing the more desirable or more demanding end of the item.
- **Scale Scores:** The second piece of information for each item is the Scale Score (11). Like the factors, each item has been compared to its norm base and reported as a Scale Score. The distribution of each normed item has been adjusted to have a mean of 5.5 and a standard deviation of 2. The norm base used is the College/Division norms. Keep in mind that each item is adjusted according to its own distribution (mean and standard deviation). Consequently, getting different Scale Scores from

the same raw scores is not only possible but likely.

It is recommended that an instructor use the Interpretive Profile or the Scale Scores rather than raw scores for making decisions about their own performance and how it might be improved in the students' eyes. Raw scores are essentially meaningless on their own. They are unstable over time, and despite having a one to five range, a large number of students don't utilize the full range such that practically, they have an even smaller range. Because of this, comparisons from semester to semester or between items are not useful. Scale Scores for items, while less stable than factor scores, do provide some standard range for the purpose of comparisons as does the Interpretive Profile.

- **Percentiles:** The third kind of information is the Percentile for each item (12). The Percentile is based directly upon the Scale Scores and can be interpreted in the same fashion as the factor Percentiles. For example, an instructor who receives a 57% can interpret that he/she did as well or better on that item than 57% of the instructors in the norm base.
- **Interpretive Profile:** We have related the items to comments in the same way we related the factors. In (13) you will find "Low," "Good," "Very Good," or "High" for each item.
- **Item Description:** The item text (14) is given along the far right. Note that the actual text has been reworded slightly from the questions on the SPTE itself. This is primarily due to limited space and the goal of reflecting the good/hard end of the item.
- **Item Cautions:** Again, be very careful in interpreting any scores related to items. Anomalies often occur. It is possible for two different classes with exactly the same item means to have different scale and Percentile scores. We have seen one case where the same instructor in two different classes averaged a perfect five on the same item in both classes, yet the Scale Score and Percentile were different. To understand how this could happen, one must remember that the Scale Scores are corrected for both class size and especially *a priori* motivation. If in the first class, there were few students who really wanted to take the course and the second had many students who dreaded taking the course, and if the item mean are the same for both courses, the first course will get a lower Scale Score (and thus a lower

Percentile score) than will the second. Again, item data should NOT be used for summative purposes.

What to Use for Tenure & Promotion, Salary Evaluations, etc.

We believe strongly that SPTE or any other measure of students perceptions of teaching effectiveness should not be the ONLY data used for teaching evaluation. While student perceptions of teaching quality should play a significant role in teaching evaluation, the use of other data such as course materials, peer evaluation, syllabi, graded exams and papers, etc. is essential.

When using SPTE for evaluation, the Scale Scores for each of the factors and the Interpretive Profile are appropriate. While the Percentile is provided, it is especially sensitive to the negative skewness of the distributions of the items and factors related to quality resulting in an overly negative interpretation. The back page (the item summary) IS NOT appropriate for Tenure, Promotion or Salary decisions. Items have less stability, less variance and are more susceptible to chance variation. The item summary is intended only for an instructor's formative use.

Departmental chairs and university administrators should be most concerned with the Perceived Quality Index and its four factors, as a measure of student perceptions of course and teaching quality. On those factors, instructors producing "Low" on the Interpretive Profile, especially on the PQI factor should be of primary concern.

A Few Hints on How to LOWER Your SPTE Scores

Over the years that we have administered SPTEs, we have noticed several things which instructors have done that are guaranteed to lower an instructor's SPTE scores.

1. Insult your class. Tell them regularly how stupid they are.
2. Give an exam during the same class period when your SPTEs are being administered.
3. "Chew out" a student in front of class.
4. Go out of your way to make your class angry with you. The more justifiable the anger the better!