

# STUDENTS' EVALUATION OF TEACHING (SET)

Authored by  
**Mohammed looti**

November 16, 2025

## RECOMMENDED CITATION

Mohammed looti (2025). *STUDENTS' EVALUATION OF TEACHING (SET)*. Encyclopedia of psychology. Retrieved from <https://encyclopedia.arabpsychology.com/?p=18135>

## Introduction to Students' Evaluation of Teaching (SET)

Students' Evaluation of Teaching, commonly abbreviated as **SET**, represents a systematic and formalized process utilized across institutions of higher education and various schooling levels to gather feedback regarding the instructional effectiveness and performance of educators. At its core, SET is a vital practice where students act as primary stakeholders and evaluators, offering structured assessments of the merits and demerits of a teacher's delivery, course design, and overall learning environment. This procedure moves beyond simple anecdotal feedback, relying instead on standardized instruments--often questionnaires or surveys--designed to capture quantitative and qualitative data on multiple facets of the teaching experience. The integration of SET into institutional policy underscores a commitment to continuous quality improvement, emphasizing that effective teaching is a measurable, accountable component of academic success.

The immediate practical application of SET involves providing instructors with concrete, actionable data that highlights both strengths and areas requiring development. As the original definition implies, the fundamental goal is remedial and developmental: the teacher reviews the structured feedback, identifies specific flaws or inefficiencies in their pedagogy, and subsequently implements corrective measures to enhance future instruction. This cyclical process of evaluation, reflection, and modification is crucial for professional growth. Furthermore, the formalized nature of asking students about the relative quality of their learning experience--for example, "To ask students about **merits and demerits** of a teacher is an example of SET"--ensures that the student perspective, often the most direct barometer of classroom efficacy, is systematically incorporated into institutional decisions about teaching quality.

While the term SET is widely used, various institutions may refer to this process using similar nomenclature, such as Course Evaluations, Student Feedback Surveys (SFS), or Course and Teacher Evaluations (CTE). Regardless of the specific title, the underlying methodology remains consistent: collecting student perceptions on instructional skills, clarity of communication, fairness of assessment, accessibility, and enthusiasm for the subject matter. These evaluations serve not only as tools for individual teacher development but also frequently contribute to critical administrative functions, including decisions related to **tenure, promotion**, merit pay increases, and departmental planning. Therefore, SET occupies a pivotal position in the intersection of faculty development, institutional accountability, and student success metrics within the educational landscape.

## Historical Context and Evolution of SET

The formalization of Students' Evaluation of Teaching is a relatively recent phenomenon within the long history of education, gaining significant traction primarily in North American universities during

the late 1960s and early 1970s. This emergence coincided with broader societal movements demanding greater accountability from institutions and a heightened focus on student rights and consumer protection within education. Before this period, teaching assessment was largely informal, relying on peer observation, departmental head judgment, or purely subjective student commentary. The shift towards standardized surveys reflected a desire to introduce **objectivity** and systematic measurement into what had traditionally been considered an artistic or inherently subjective domain. Early pioneers in this field sought to develop psychometrically sound instruments capable of reliably measuring instructional effectiveness across diverse disciplines and course formats.

Initial instruments were often locally developed, varying wildly in their structure and focus, leading to concerns about validity and comparability. However, key research efforts--such as those centered on identifying universal dimensions of effective teaching--helped standardize the field. The development and widespread adoption of validated instruments, such as the Purdue Rating Scale or the Instructional Development and Effectiveness Assessment (IDEA) system, marked a critical turning point, lending **credibility** to the process. These standardized tools allowed for cross-course and cross-departmental comparisons, enabling institutions to establish benchmarks for teaching excellence. The evolution of SET paralleled advancements in educational psychology, particularly research confirming that certain observable teaching behaviors correlate positively with student learning outcomes.

More recently, the historical trend has shifted dramatically from paper-and-pencil surveys conducted in the classroom to sophisticated digital platforms. The transition to **online evaluation systems** has introduced new complexities and opportunities. Digitization allows for easier data aggregation, rapid feedback processing, and the integration of evaluation data with institutional databases. Furthermore, contemporary SET systems often incorporate mechanisms for personalized feedback, allowing instructors to receive tailored reports that include comparative data against departmental or institutional averages. This historical trajectory demonstrates a continuous refinement of the methodology, moving from simple feedback mechanisms to complex, **data-driven systems** designed to support evidence-based decision-making regarding teaching quality.

## Core Mechanisms and Methodology of SET

The core mechanism of Students' Evaluation of Teaching involves the deployment of a structured survey instrument, typically administered near the conclusion of an academic term. These instruments are meticulously designed to assess various critical domains of instruction, ensuring comprehensive coverage of the teaching and learning experience. Standardized SET forms often employ **Likert scales**, where students rate specific statements--such as "The instructor clearly explained complex concepts" or "The course materials were relevant"--on a continuum ranging

from "Strongly Disagree" to "Strongly Agree." This quantitative data provides measurable metrics of perceived effectiveness, allowing for statistical analysis of teaching performance. The emphasis on standardized questions ensures that data collected across different classes remains comparable and interpretable by administrators.

In addition to quantitative ratings, most robust SET methodologies integrate open-ended, **qualitative questions**. These questions invite students to provide written commentary, offering depth and nuance that numerical scores alone cannot capture. Typical qualitative prompts might ask students to describe the best aspects of the course, suggest specific improvements, or comment on the instructor's communication style. This dual approach--combining numerical scores with detailed textual feedback--is essential. While quantitative data facilitates administrative decisions and large-scale benchmarking, the qualitative comments are often the most valuable resource for the instructor seeking specific, **actionable guidance** on pedagogical adjustments. Effective utilization of SET requires careful attention to synthesizing both data types.

The methodological rigor surrounding SET administration is paramount to ensuring the validity and reliability of the results. Best practices dictate that evaluations should be administered consistently, usually during class time or through a supervised online environment, to maximize response rates and minimize external bias. Crucially, institutions must guarantee the **anonymity** of student responses. The assurance that feedback cannot be traced back to individual students is fundamental, fostering honest and constructive criticism rather than fear-based or overly generous ratings. Finally, the timing of evaluation is strategic; collecting data toward the end of the semester ensures students have experienced the full scope of the instructor's performance and course requirements, providing a holistic assessment.

## Purposes and Benefits of Utilizing SET Data

The utilization of SET data serves multiple, interconnected purposes within academic institutions, benefiting individuals, departments, and the institution as a whole. Primarily, SET serves a **formative purpose**, acting as a crucial tool for faculty development and self-improvement. When instructors receive detailed feedback, they are empowered to reflect critically on their teaching practices, identify discrepancies between their intended instructional goals and the students' experienced reality, and make necessary pedagogical adjustments. For example, if feedback consistently indicates that students find grading criteria unclear, the instructor can refine their syllabus and communication methods for the subsequent term, directly addressing the identified flaw. This developmental cycle reinforces the professional obligation of **continuous improvement**.

Secondly, SET fulfills an essential **summative purpose**, contributing significantly to institutional personnel decisions. In many higher education systems, SET scores are mandatory components of files used for tenure review, promotion applications, annual performance appraisals, and merit

salary adjustments. While institutions emphasize that SET scores should never be the sole criterion for judging teaching effectiveness, they provide standardized evidence of student perception that complements peer reviews and teaching portfolios. Administrators rely on these metrics to ensure **accountability** and reward faculty who demonstrate consistent excellence in instruction, thereby aligning institutional resources with teaching quality objectives. The formality of this evaluation process ensures that teaching, alongside research and service, is rigorously assessed.

Beyond individual and personnel functions, SET data offers significant benefits for **curricular review** and departmental planning. Aggregated evaluation data can highlight systemic issues within a program, such as consistently low ratings for required introductory courses, which might signal the need for curriculum redesign, resource reallocation, or targeted faculty training in specific pedagogical techniques. Furthermore, tracking SET results over time allows departments to monitor the impact of instructional interventions and gauge the success of new teaching methodologies. Thus, SET transforms individual feedback into **institutional intelligence**, driving evidence-based enhancements to the overall quality of education delivered.

## Challenges and Criticisms Associated with SET

Despite its widespread adoption and utility, Students' Evaluation of Teaching is not without significant challenges and persistent criticisms that fuel ongoing debate within the academic community. One of the primary concerns revolves around **response rates and demographic bias**. Low response rates can compromise the representativeness of the data, potentially allowing a small, unrepresentative subset of students (e.g., those who are extremely satisfied or extremely dissatisfied) to skew the overall results. Furthermore, extensive research has explored biases related to student demographics (e.g., gender, race, and physical attractiveness of the instructor) and course characteristics (e.g., course difficulty, grading strictness, and subject matter). These biases raise serious questions regarding whether SET truly measures teaching effectiveness or merely measures student satisfaction, perceived fairness, or even implicit prejudices.

Another major criticism centers on the **misinterpretation and misuse of SET data**, particularly the overreliance on a single aggregated mean score. Reducing complex teaching performance to a numerical average often overlooks the nuances required for effective pedagogy in specialized contexts. For instance, a challenging, upper-level course requiring strict grading might inherently receive lower SET scores than an easy, introductory elective, even if the former represents demonstrably superior teaching. When administrators place disproportionate weight on raw scores for high-stakes decisions like tenure, it can incentivize faculty to prioritize student satisfaction (e.g., "grade inflation," easy assignments) over **rigorous academic standards**, a phenomenon often termed "teaching to the test" of the evaluation form.

Furthermore, the validity of SET instruments themselves is continually scrutinized. Critics argue that student judgment, particularly among less mature or inexperienced students, may not accurately reflect true learning gains or long-term knowledge retention. Students are adept at evaluating observable behaviors, such as clarity, organization, and enthusiasm, but may lack the expertise to assess deeper instructional qualities related to **critical thinking promotion** or disciplinary mastery. Addressing these criticisms requires institutions to adopt sophisticated statistical models to control for known biases, employ qualitative analysis alongside quantitative results, and ensure that SET data is always contextualized by other forms of evaluation.

## Factors Influencing Student Ratings

A comprehensive understanding of SET requires acknowledging the multitude of factors--both internal and external to the teaching process--that significantly influence student ratings. One prominent set of factors relates to **course characteristics**. Course difficulty is strongly correlated with evaluations; courses perceived as easier or those with less demanding workloads often receive higher scores. Similarly, class size plays a role, with smaller classes typically facilitating greater interaction and personalized attention, often translating into higher student satisfaction scores. The subject matter itself can also be a factor, as students often rate instructors in required, highly technical, or less intrinsically interesting courses lower than those teaching popular electives.

Another critical area involves **instructor attributes and student expectations**. Research consistently indicates that instructor clarity, organization, and enthusiasm are among the strongest predictors of positive ratings. Students highly value instructors who appear prepared, communicate expectations effectively, and demonstrate passion for their subject. Conversely, attributes that are unfortunately less related to actual teaching quality, such as physical appearance or shared demographic background with the students, have been shown to exert a measurable, albeit ethically problematic, influence on scores. Students' pre-existing expectations, often derived from peer recommendations or institutional reputation, also frame their evaluation experience; an instructor perceived as being "tough but fair" might receive lower scores on satisfaction but high scores on rigor and fairness.

Finally, **student factors and the evaluation environment** impact the outcomes. Student motivation, prior subject knowledge, and personal interest in the course material can influence their perception of the instructor's effectiveness. A student struggling with the material may attribute their lack of success to the instructor, regardless of instructional quality. The evaluation context--whether the survey is mandatory, anonymous, online, or in-class--also affects response honesty and effort. Recognizing and attempting to mitigate the influence of these extraneous variables through statistical controls or through careful instrument design is a necessary step in validating SET data for high-stakes purposes.

## Best Practices for Implementing and Interpreting SET

To maximize the efficacy and minimize the pitfalls associated with SET, institutions must adhere to several best practices regarding implementation and interpretation. Administratively, the system must be managed with absolute transparency and clear communication regarding the purpose and weight of the evaluations. Institutions should strive for high response rates, utilizing active strategies such as dedicated in-class time for completion or offering minor incentives, while strictly maintaining the **anonymity** of respondents to ensure candid feedback. The design of the instrument itself should prioritize questions that measure specific, observable teaching behaviors known to correlate with learning outcomes, rather than vague statements about general satisfaction.

For faculty receiving the reports, the interpretation must be holistic and context-aware. Instructors should be trained to look beyond the mean scores and delve deeply into the **qualitative comments**, searching for recurring themes and actionable suggestions. When reviewing quantitative scores, instructors should benchmark their results not against an institutional average, but against data from similar courses--those with comparable enrollment, level, and required status--to ensure fair comparison. Furthermore, faculty should be encouraged to utilize SET feedback developmentally, establishing concrete goals for instructional improvement based on identified weaknesses and seeking mentorship or workshops to address those areas.

For administrative use, high-stakes decisions (e.g., tenure and promotion) should mandate that SET data is evaluated alongside other robust evidence of teaching effectiveness. This evidence should include:

**Peer Review:** Structured observation and assessment by departmental colleagues who can evaluate pedagogical methods specific to the discipline.

**Teaching Portfolios:** Collections of instructional materials, reflective essays, and evidence of student learning (e.g., graded assignments, student performance data).

**Student Learning Outcomes (SLOs):** Direct measures of what students have learned, used to corroborate the perception data collected via SET.

By integrating SET scores as one piece of a comprehensive, multi-faceted evaluation system, institutions ensure that decisions regarding teaching quality are based on rich evidence rather than relying solely on potentially biased numerical ratings. This balanced approach upholds academic rigor while valuing student input.

## Future Directions and Research in SET

The field of Students' Evaluation of Teaching is continuously evolving, driven by technological advancements and renewed scholarly focus on optimizing feedback mechanisms. One significant future direction involves the integration of advanced **learning analytics** and Big Data techniques. Future SET systems are likely to move beyond simple end-of-course surveys, incorporating real-time feedback mechanisms linked directly to activities within Learning Management Systems (LMS). This allows for continuous monitoring of student engagement and instantaneous feedback loops, enabling instructors to make mid-course corrections, thereby transforming SET from a retrospective tool into a proactive diagnostic system.

Another key area of research focuses on enhancing the validity of SET by better aligning evaluation questions with measurable student learning outcomes. Researchers are developing instruments that explicitly ask students to rate the degree to which specific instructional techniques (e.g., active learning exercises, problem-based assignments) contributed to their mastery of core competencies, rather than simply rating the instructor's personality. This shift aims to make SET a more direct measure of **pedagogical efficacy** rather than affective satisfaction. Furthermore, the development of sophisticated statistical models is ongoing, focusing on creating more accurate methods for isolating the true instructional effect of the teacher from the noise generated by confounding variables, such as student prior achievement or course selection bias.

Finally, there is a growing emphasis on the ethical implications and **equity concerns** within SET. Future research will increasingly focus on designing and implementing evaluation systems that actively mitigate known demographic biases, ensuring that SET does not inadvertently perpetuate systemic inequalities in faculty assessment. This may involve incorporating bias-reduction training for students completing evaluations or utilizing administrative review panels specifically tasked with interpreting scores through an equity lens. The future of SET lies in its integration into a holistic, technology-enhanced system that provides timely, reliable, and equitable data for the comprehensive assessment and continuous improvement of teaching quality across all educational settings.