TARSIB: Jurnal Program Studi PGMI
Volume 3 Nomor 2, Oktober 2025
e-ISSN 3046-806X p-ISSN 3046-8264
DOI: https://doi.org/10.61181/6vpcvs08

Correlational Study

The Influence of Teaching Situations and Motivational Beliefs on Teaching Quality Ranking

Supeno
Universitas Pangeran Diponegoro Nganjuk, Indonesia
Email: enoazj@gmail.com

Abstract

Ratings conducted by students are usually used to evaluate the class process.Research shows that ratings are not only based on objective situational characteristics, but also reflect the characteristics of the assessor. This study uses data from1.745 undergraduate students of Prince Diponegoro University Nganjuk for the last 5 years.The study used single trait-multistate model analysis to unravel the variants of instructional quality (student support, cognitive activation). The relationship between motivational beliefs and the component that was also estimated to be consistent across the rankings was the effect of people. The ranking of teaching quality over time is consistent in the same lecture. Consistency is lower when teaching quality ratings are assessed in various lectures. Consistency is higher in lectures that are considered difficult by students than lectures that are considered important. Students who reported higher expectations for success and interest in lectures also rated instructional quality more positively in all teaching situations. The findings of the study show that the ranking of teaching quality is located and shaped by the characteristics of lectures.

Keywords: *Teaching quality, Motivational beliefs, Situation, State-trait analysis*

INTRODUCTION

Ranking conducted by students in education is widely used to capture students' experiences and perceptions of the teaching situation over a long period of time (Praetorius, 2014).[1] Previous research has operationalized students' ratings of teaching situations, either as stable individual tendencies or as context-dependent evaluations, thus emphasizing the role of their situation in ranking, with focused on ranking students on aspects of motivation and teaching affect, such as teacher enthusiasm, and; [2] examined the extent to which students' ratings of teaching quality are context-dependent evaluations in settings at primary and secondary levels.[3] Research shows some consistency in students' ratings of teaching situations, but little is known about how these findings translate to higher education.

This research study expands on previous research by examining the location of teaching quality ranking in higher education programs. The ranking carried out by each person on a situation contains variance because the person who assesses, the situation is assessed, and the particular person interacts with the situation.[4] An important question to ask in the educational sciences is what information is captured by students' ratings on the quality of teaching and the extent to which ratings are influenced by the characteristics of the assessor, the situation itself and/or representing people with the effects of situational interactions.[5]

A state-of-the-art statistical model can help delineate the variance of a student's ranking and investigate the situation. This study uses a state-of-state statistical modeling approach to investigate the consistency and specificity of situations in students' ratings of teaching quality. Specifically, this study focused on the differences between the effects of people and the specific effects of situations. The people effect refers to a consistent tendency between individuals to rate the quality of teaching in a variety of teaching situations (e.g., some students consistently tend to rate cognitive activation generally higher than other students across classes). Situation- specific effects capture variability in student rankings in

various teaching situations. These include the effect of the situation (e.g., some lectures are generally considered more supportive by all students than others) and the effect of people's interaction with the situation (e.g., a student finds lectures to be highly structured, while most classmates rate them as disorganized). This study explores whether student ratings vary across different teaching situations. This study also investigated the relationship of motivational beliefs in terms of expectations for success and interest values (Eccles et al., 1983; Eccles & Wigfield, 2020) with a consistent learning quality rating component.[6] The research was conducted in the context of undergraduate education at Prince Diponegoro University Nganjuk.

LITERATURE REVIEW

Quality of teaching

The quality of teaching broadly plays a role for the cognitive and affective-positive development of students across educational contexts and stages.[7] Teaching quality is a general concept that refers to teacher behavior and interaction between teachers and students.[8] The quality of teaching is conceptualized as part of an in-depth instructional structure. Inner structure is distinguished from surface structure – a distinction that helps answer the question of how to facilitate learning. The structure of the surface is explained by the observable strategies and resources used during teaching, i.e. the use of media, social forms, methods and materials. Deep structure represents the quality of interaction between teachers, students and subject matter. Important in this context is that the orchestration of surface and deep structures with subject matter influences students' learning and educational success.

A generic framework of teaching quality that is not specific to a subject and refers to effective practice across subjects, has been suggested by some scholars. This study uses a general three-dimensional model of teaching quality (instructional), namely classroom management, student support and cognitive activation. A generic three-dimensional model of teaching quality was developed using a bottom-up approach, based on data from lecture settings. Some scholars define the three dimensions as not based on theory. The researchers used 21 scales from didactic, teaching research and school climate research for classroom video analysis.[9] The ratings are then examined using factor analysis. It is possible that the generic dimension also applies in the context of higher education, as there is a possibility of overlapping with the model of teaching quality in higher education.[10] Recent studies have shown that these three dimensions have different influences on students' beliefs and motivational achievement.[61]

Classroom management involves organizing activities and using time in the classroom effectively. For example, it results in clear rules, structured lessons and minimal disruption.[11] Student support reflects the extent to which teachers create nurturing and respectful learning environments characterized by emotional and cognitive support, a culture of positive mistakes and ensuring choice in the classroom. Cognitive activation reflects the extent to which teachers present students with challenging tasks, making connections between different concepts clear and connecting new content with previous knowledge, thereby improving students' understanding, engagement with content and the use of reflection.[12]

Teaching quality assessment locations

The quality of teaching can be assessed using several data sources, e.g., student rating, teacher rating, external assessor rating), and; methodology, e.g., surveys and observations. Student surveys are one of the most commonly used methods to measure teaching quality, which is the ranking of students against the quality of teaching that is experienced in a particular subject. Some scholars describe students' ratings of teaching quality as "biased" due to the influence of people's characteristics and people's processing on situations.[13] A student's rating of the quality of teaching is defined as intra-individual and inter-individual variation in the ranking of situations. Ranking as a meaningful part of the situational experience and not as bias. The situational framework and its situation help conceptualize students' ratings of the quality of teaching and the situation. There is no universally agreed definition of "situational" and "situational." Situations in general can be defined as a combination of individually, implicitly and uniquely interpreted understandings, as well as culturally shared, explicit and general understandings of the surrounding environment that generate and limit human behavior. The ranking of situations can thus be broken down into the effect of people, the effect of situations, and the effect of

situational interactions × people + interceptions and errors.[14] Some situations can be so prominent, that there will be very little of a person effect, for example, when someone points a gun at you.[15] Students may have similar perceptions of the quality of teaching. In particular, students tend to view similar situations differently, or different students may interpret the same situation in different ways, or the situation may be completely different between students in the same class due to different treatment.

Previous research has shown how student grading is responsive to a variety of teaching situations determined by different times, instructors or subjects.[16] For example, the grading of elementary and secondary school students shows a high level of consistency, i.e. the strong person effect, when students evaluate the same teacher in the same subject over time.[17] Interestingly, noted consistency in students' ratings of teaching quality even when the same elementary and secondary school teachers taught different subjects. This consistency is somewhat decreased when students assess the quality of teaching across situations, such as different lessons or different teachers at the same time or at different times.[18] Similar patterns of consistency and variation across disciplines have been found when examining student rankings in higher education.[19] The findings of the study showed that students' ratings showed a moderate level of consistency, although students rated the quality of teaching in different situations that showed a moderate person effect.

Students' ratings of actual instructional behaviors and their consistency vary according to the dimensions of teaching quality, namely classroom management, student support, and cognitive activation. Some dimensions of teaching quality, such as student support, are more subjective and tied to individual needs of students, while others, such as classroom management, are perceived more universally.[20] Studies in elementary and secondary schools found higher consistency for clarity of teaching, monitoring, and classroom management than ratings of emotional support or student autonomy. The findings provide further support for the ranking location. These findings may on the one hand reflect situational specificity, i.e. dynamic co-construction between teacher and student, while more individualized instructional processes lead to lower consistency.[21] These findings on the other hand can also point to individual effects, reflecting students' individual needs, attitudes, and preferences that are stable across learning environments. There are several studies that have systematically examined the location of student teaching ranking while also considering different dimensions of teaching quality; No studies have been carried out in higher education.[22]

Another factor that is important for a student's ranking of teaching quality is how difficultit is for students to perceive the subject.[23] The workload and difficulty of the lessons perceived by students are important background variables for effective teaching.[24] Situated Expectancy-Value Theory (SEVT) shows that the process of perceived learning environment difficulties can affect how students interpret and behave in their learning environment.[25] This is because students' motivational beliefs are related to their competencies and interests. According to Cognitive Load Theory, high cognitive demands can reduce an individual's capacity for detailed information processing, which leads people to rely more on heuristics or general impressions.[26] People in situations of high cognitive demand, are more likely to revert to stereotypes rather than pay attention to situational details.[27] High cognitive demands in ranking teaching quality, indicate that in lessons or courses that are more cognition-demanding (difficult subjects), students or students may rely more on general impressions than actual situational variations when assessing teaching quality. As a result, student rankings may become more consistent in a variety of situations, more driven by the effect of people. Based on SEVT, it can also be assumed that other factors, such as how important students perceive a lesson, can be considered as the specificity of the situation. This research focuses on perceived difficulties and the importance of courses and how these characteristics determine students' ranking of teaching quality.

Motivational belief in students' rating of teaching quality

Based on SEVT and social perception theory, this study proposes the hypothesis that students' initial motivational beliefs affect the ranking of teaching quality. The level of confidence in students' motivations, for example, interests and expectations for success, affects how students interpret their social environment, for example, lectures on their campuses. Students with stronger motivational beliefs are also more engaged in learning activities, and in turn cause students to receive greater support from lecturers, or perceive social realities to be more supportive.[28] The higher the student's interest and expectation of success, the higher the student's rating of the teacher's teaching.[29] Other researchers also

found that students who were initially interested in the subject rated the quality of teaching in the classroom more positively. This is likely to also be true in higher education.[30] Literature documentation found that there was only one study that did not support the relationship between initial interest in the subject and the rating of the quality of teaching conducted by students.[31] Scholars have found that students with more positive expectations about the value of their subjects also evaluate teaching more positively, but not in their teachers.[32] Interpretations driven by value interests and expectations for syksescan lead to more favorable evaluations in a variety of teaching settings because students are more consistently cognitively engaged across situations.[33] Students with weaker motivational beliefs may not have stable cognitive engagement, making students more responsive to situational cues and thus more located in ranking the quality of accepted teaching.[34]

Social perception theory suggests that individuals form expectations about their social reality. Social reality expectations influence information processing leading to biased perceptions and judgment formation.[35] Social perception theory sees individual beliefs as drivers of information processing, but not individual involvement. Expectations about social reality affect information processing because people may only seek information that matches their expectations or because people unconsciously limit their attention to information that matches their expectations.[36] The impact on the perception of "bias" and the formation of judgments depends on the state of the motivational belief and how much information needs to be processed that is important to the person. People with high cognitive needs are more likely to remember information that is consistent with stereotypes than people with lower cognitive needs. Highly motivated students may be less flexible in their perceptions of their social reality, as perceived evidence of social reality is important to them. Ranking teaching quality in the context of college lectures, it can be assumed that students with high grades and positive expectations for success may be more likely to rely on well-established cognitive schemes, or; more involved in lectures, and resulting in more favorable teaching quality ratings in various teaching situations.

Previous research has provided valuable insights, but there is still relatively little insight into the role of motivational beliefs in shaping the various components of student rating on teaching quality. In particular, it is still unclear to what extent grades and expectations for success affect the ranking of teaching quality in various situations. This research aims to: First, to provide insight into the consistency and specificity of the student ranking situation on the quality of teaching, and; The second is to investigate the relationship between initial motivational beliefs and the consistent components of student ratings of teaching quality.

Substantial findings provide insight into where students rank on teaching quality, but some gaps in the study remain. First, most of the existing empirical evidence comes from primary and secondary education settings, while studies focusing on higher education are still relatively rare.[37] Second, there are not many studies that have systematically examined the location of students' ratings of teaching quality in various lecture characteristics and teaching situations. This research examined variations across different lectures, time points, and course types, i.e., the courses that students considered the most important or difficult. In addition, the role of motivational beliefs in explaining the consistent components of student rankings remains an open question.[38] This study asks the question:

Research Question 1 (RQ1), what proportion of student ratings to teaching quality consists of consistent components (people effect) versus by situation-specific components (situational effects, × effects of people's situational interactions)?

Research Question 2 (RQ2), to what extent is the belief in student motivation, namely the value of interest and expectation for success related to the ranking consistency component of the three dimensions of teaching quality?

METHOD

Sample

The survey involved the participation of students in all study programs of Prince Diponegoro Nganjuk University for five academic years. The survey has collected data from a total of 1,745 student samples purposively in the 2019/2020 to 2024/2025 academic year. The sample consisted of 26.5% of

students and 73.5 female students. The age of students ranges from 19 to 23 years old (Mean = 20, SD = 3.00). As many as 5.4% of students are low-income and 59.7%.

Students from all study programs are invited to participate in the research. Participating students receive college credit. Students who agree to participate are required in each academic semester to choose one course that is considered the most difficult course and one course that is considered the most important course in that semester. Students choose courses as the "most difficult" courses, when courses are considered challenging because of exceptional material, or high tuition and assignment loads and perceived low instructional quality. Courses that are chosen as "most important" when they are considered necessary for a student's major, and also courses that are considered personally valuable and important for his or her future. The difficult and important courses chosen include Arabic Language, Psychology, and Research Methods courses.

Measuring instruments

Teaching quality measures three dimensions, namely classroom management, student support, and cognitive activation. Each dimension is measured by a single statement. The questionnaire is the result of an adaptation of the teaching quality statement from the OECD 2013. Responses to statements on a scale of 1 to 7, from very high quality to not at all quality. Internal consistency was ≥ .82 in difficult courses and ≥ .85 in important courses.

A student's initial expectation of success is assessed with three items at the beginning of the semester of each academic year for each of the two types of lectures, namely the most difficult course and the most important course. The statement items were developed based on the guidance of the expectation value theory[39] and adapted to the context of high school. The statement items are presented on a scale of 1 = not at all good until 7 = very good. The weight of the item factor varied on a scale of $\lambda \geq .80$ in the lecture of difficult courses and $\lambda \geq .79$ in the lecture of the important course. Internal consistency was ≥ .87 in difficult courses and ≥ .89 in important courses.

Motivational confidence assesses the value of students' initial interest in each type of lecture at the beginning of each academic year. The statement items were developed based on the guidance of expectation value theory (Eccles et al., 2005) and adapted to the context of high school. The statement items presented on a scale of 1 = are not at all expected until 7 = highly expected. The weight of each factor varies on a scale of $\lambda \geq .78$ in the lecture of difficult courses and $\lambda \geq .73$ in the lecture of important courses. Spearman-Brown internal consistency $r \geq .76$ in difficult courses and $r \geq .71$ in important courses. The items of expectations and beliefs of values have been validated.[40]

Statistical analysis

The research data were analyzed using single trait-multistate statistical models.[41] The trait factors in the model capture consistency (the people effect), while the state factor in the model represents cross-situational variation (situational effects, situations × people's interaction effects).

To answer the first research question (RQ1), consistency (Con) and occasion-specificity coefficient (Occ) were estimated. The specificity of opportunity is an indicator of the specificity of the situation. The consistency coefficient measures the proportion of trait variance in student rankings in various situations. The specificity coefficient captures the variance associated with the circumstances and situation factors under investigation (Geiser, 2021). Both coefficients represent the observed proportions [Con (Yit); Occ score (Yit)] or true (error-free) [Con (τ); Variability Occ (τ)]. This study focuses on the actual score.

A complex study design (random situational design) with repeated teaching quality assessments makes it possible to define four models that take into account a wide range of situations. The situation is represented by the type of lecture (difficult vs. important), the various lectures in each academic year and the varying time points (two time points per lecture).

Model A - Same lecture type, same lecture, different time point: The lecture type and lecture are the same, for example, one difficult lecture, while the time varies based on the academic year. These models make it possible to examine the consistency of each dimension of teaching quality in one lecture of a particular course at different points in time. The model is defined separately for difficult and

important lectures in each dimension of quality, i.e. classroom management, student support and cognitive activation.

Model B - Same lecture type, different lectures, different time points: A model of the same lecture type (difficult or important lectures), but three different lectures (in each academic year) and different time points. The model makes it possible to assess consistency across three lectures with the same lecture perception (important or difficult) and situational variation of each dimension of teaching quality (specificity of opportunity). One model is defined for each quality dimension for difficult and important lecture contexts.

Model C - Different lecture types, different lectures, and the same time point: Lectures and lecture types as variables (important and difficult lectures), while keeping the time point constant (one point in a semester in each academic year). The approach makes it possible to investigate the consistency and specificity of opportunities of each dimension of teaching quality at a fixed point in time, so that the consistency of ranking in two lectures, one important and one difficult, is at the same time.

Model D - Different types of lectures, different lectures and different time points: Here lecture types (difficult and important lectures), lectures and time points (six time points throughout the academic year; two per academic semester) are allowed to vary. This approach provides insight into the consistency and specificity of opportunity from each dimension of teaching quality across time, lectures, and lecture types—this can be interpreted as the overall perceived quality of teaching at the university (see Figures 2 and 4). One model is specified for each quality dimension (3 models).

To answer the second research question (RQ2), motivational beliefs (ξEXP' ξIV) were added to the model and examined the relationship (ρ) with the consistent component of teaching quality (ξ, the effect of people). For motivational beliefs, different modeling approaches were used that corresponded to the specifications in Model A–D. Analysis was also carried out on how much variation in ranking consistency could be explained by motivational beliefs.

The analysis was conducted using Mplus 8.11 and a robust maximum-like-lihood estimator (MLR). Loading, interception, and measurement error variance are all set in the model. Invariant loading and interception modeling, consistent people effects are positioned in a variety of situations. The evaluation of model goodness used a comparison of the less constrained model fit with the more constrained one and was guided by the threshold for a sample of ≥ 30. All models of strict measurement invariants match the data well.

RESEARCH RESULTS

Consistency (Con) and specificity of opportunity (Occ) instructional quality (RQ1)

Variance in students' instructional quality ratings (r teaching quality) was broken down into consistency (Con [τit]) and specificity of opportunity (Occ [τit]). The results of the analysis showed that all models showed good model compatibility.

Model A (Same course type, same course, different time point) shows higher consistency than opportunity specificity. Up to 68% of variances reflect consistent person effects, while less than 54% are residual variances of circumstances. This suggests that the assessment of student teaching quality is relatively consistent over time for the same course and course type (the person effect), with greater consistency than the specificity of the situation in all three dimensions.

Model B (Same course type, different course, different time point) shows lower consistency than opportunity specificity. Less than 34% of variance reflects consistency, while up to 71% reflects residual variance of the state. This suggests that student teaching quality assessments are more varied across different subjects and difficult or important time points, suggesting a stronger influence of contextual characteristics.

Model C (different course types, different courses, same time points) also shows lower consistency than opportunity specificity. Less than 43% of variances reflect consistent people consistency than opportunity specificity. Less than 43% of variances reflect consistent people effects, while up to 77% reflect residual variance of circumstances. This can be interpreted as low consistency in the assessment of teaching quality at different points in time (the people effect) and higher specificity on the opportunity in the assessment of student teaching quality in all three dimensions in various

difficult or important subjects at the same point in time. Thus, even at a single point in time, student assessments differ significantly depending on the course and the type of course.

Model D (different types of courses, different courses, different time points) shows a similar pattern: less than 35% of variance reflects consistent people effects, while up to 70% reflects residual variance of circumstances. This can be interpreted as low consistency in the assessment of teaching quality at various points in time (the people effect) and higher situational specificity in the assessment of student teaching quality in all three dimensions in various subjects that are difficult or important at some point in time.

In summary, student teaching quality assessments show greater consistency over time when referring to the same courses and types of courses. For example, one difficult course is assessed at several points in time (Model A). With increasingly varied situations (Model B–C), for example, student assessments for three subjects that are considered difficult at different times (Model B), the specificity of student assessment situations increases.

Interestingly, when assessments for more situations across different models were integrated, there were some differences in the consistency and specificity of opportunities depending on the dimensions of teaching quality, type of course, and timing of assessment. First, comparing the consistency and specificity of opportunity across different dimensions of teaching quality, there was no clear difference in Model A (same type of course, same course, different time point). In Model B (same course type, different courses, different time points), cognitive activation assessments on average had higher consistency than classroom management and student support assessments. In Model C (different course types, different courses, same time points), the average classroom management assessment had higher consistency than the student support and cognitive activation assessments. In Model D (different types of courses, different courses, and different time points), cognitive activation assessments have somewhat higher consistency in difficult courses compared to important courses.

Second, when comparing difficult and important courses in Model A, the consistency of assessment in all three dimensions tended to be higher in courses that were difficult for students compared to the most important courses. No difference occurred in Model B.Third, the A and C model approaches allow for comparisons of consistency and specificity of events at different points in time. The Model A showed slightly lower grading consistency from fall to spring, while no difference occurred in the Model C.

The relationship between motivational beliefs and learning quality (RQ2)

The relationship between students' expectations of success (Expectation [EXP]) and interest value (IV) for each course with the consistent component of student assessment of teaching quality can be seen in all models.

Model A examines the relationship between students' initial course-specific motivational beliefs and teaching quality in a semester, in both the most important and most difficult courses. For both difficult and important courses, students with higher initial IV and EXP rated overall teaching quality higher ($.22 \leq \rho \leq .65$; $.22 \leq R^2 \leq .44$). In general, IV had a stronger association with teaching quality ($.43 \leq \rho \leq .65$) than EXP ($.22 \leq \rho \leq .49$). Overall, the association was similar in difficult courses ($.22 \leq \rho \leq .65$) and important courses ($.22 \leq \rho \leq .61$) with some indications of a stronger association between EXP and teaching quality in difficult courses ($.29 \leq \rho \leq .49$) compared to important courses ($.33 \leq \rho \leq .43$). The variance described ranges between .22 and .44 for difficult courses and between .22 and .38 for essential courses. Some differences occurred when comparing the dimensions of teaching quality: The association between EXP and student support ratings and cognitive activation ($.35 \leq \rho$ .49) was somewhat stronger than the relationship between EXP and class management ratings ($.29 \leq \rho$ .43) in difficult subjects. In addition, the association of IV with cognitive activation rating ($.52 \leq \rho \leq .65$) was stronger than the association with rating of student support and classroom management in difficult and important courses ($.43 \leq \rho \leq .59$). The variance described was somewhat higher for cognitive activation ($.28 \leq R^2 \leq .44$) and student support ($.22 \leq R^2 \leq .37$) in most models compared to classroom management ($.22 \leq R^2 \leq .24$).

Model B examines the relationship between students' average motivational beliefs each academic year and teaching quality throughout the academic year in both essential and difficult subjects.

For both difficult and essential courses, throughout the academic year, students with higher IV and EXP rated instructional quality items higher ($.58 \leq \rho \leq .80$; $.48 \leq R^2 \leq .69$) than their peers in these same courses. In general, a student's IV score had a stronger relationship with his teaching quality ($.68 \leq \rho \leq .80$) than his EXP ($.58 \leq \rho \leq .63$) and this relationship was somewhat stronger in difficult courses ($.58 \leq \rho \leq .80$) than in important courses ($.61 \rho \leq .80$). The variance described ranges between .55 and .69 for difficult courses and between .48 and .65 for important courses. Several differences occur when comparing the dimensions of teaching quality:

The relationship between motivational beliefs and students' assessment of student support and cognitive activation was stronger ($.62 \leq \rho \leq .80$) compared to students' assessment of classroom management ($.58 \leq \rho \leq .74$). The variance described was slightly higher for cognitive activation ($.65 \leq R^2$ .69) and student support ($.50 \leq R^2$ .59) in each model compared to classroom management ($.48 \leq R^2$ .55).

In Model C, it examines the extent to which early motivational beliefs in important and difficult lectures relate to the quality of student teaching across important and difficult lectures at the same point in time. Overall, students with higher IV and EXP had higher ratings in the quality of their teaching ($.28 \leq \rho \leq .91$; $.38 \leq R^2 \leq .87$). Again, IV had a stronger relationship with teaching quality ($.56 \leq \rho \leq .91$) than EXP ($.28 \leq \rho \leq .75$). Some differences occurred comparing the dimensions of teaching quality: The association between EXP and student support rating ($.28 \leq \rho \leq .75$) was as strong as the association with classroom management rating ($.49 \leq \rho \leq .62$) and cognitive activation ($.52 \leq \rho \leq .68$). The association between IV and cognitive activation ratings ($.70 \leq \rho \leq .91$) was somewhat stronger than that of class management and student support ratings ($.56 \leq \rho \leq .90$). In addition, ratings collected in one academic year ($.58 \leq \rho \leq .86$) were slightly more strongly associated with motivational confidence than those collected in the following academic year ($.49 \leq \rho \leq .90$). The variance described was somewhat higher for cognitive activation ($.52 \leq R^2 \leq .83$) and student support ($.40 \leq R^2 \leq .87$) in each model compared to classroom management ($.38 \leq R^2 \leq .59$).

Model D examines the extent to which motivational beliefs throughout the academic year in students' important and difficult lectures are related to the quality of teaching throughout the academic year in students' important and difficult lectures, so that cross-dimensional relationships are examined. Overall, students with higher IV and EXP rated higher teaching quality throughout the academic year ($.27 \leq \rho \leq .80$; $.48 \leq R^2 \leq .72$) compared to their peers with lower IV and EXP. Again, IV had a stronger association with teaching quality ($.48 \leq \rho \leq .80$) than EXP ($.27 \leq \rho \leq .63$) and a stronger association in difficult lecture tracks ($.52 \leq \rho \leq .80$) than in important lecture tracks ($.27 \leq \rho \leq .79$). The variance described ranged from .48 to .72 for difficult lectures and between .49 and .69 for important lectures. Some differences occurred when comparing the dimensions of teaching quality: The relationship between IV and EXP was overall stronger for cognitive activation ratings ($.34 \leq \rho \leq .80$) than for ratings of classroom management and student support ($.27 \leq \rho \leq .71$). The variance described was slightly higher for cognitive activation ($.69 \leq R^2 \leq .72$) compared to classroom management ($.48 \leq R^2 \leq .49$) and student support ($.51 \leq R^2 \leq .53$).

After describing the findings of each model, the next step is to provide a comparative summary of the research results: Students' motivational and ranking beliefs about teaching quality were positively correlated, with a stronger association for IV than for EXP. In addition, IV and EXP were more strongly associated with the quality dimensions of cognitive activation teaching than classroom management and student support, with differences of up to 21 percent in Model A, up to 17 percent in Model B and up to 30 percent in Model C. Only in Model D was a stronger association found in difficult lectures than in important lectures; The variance described differs only by 1 to 3%. When comparing these models, the overall relationship was the most homogeneous and strongest in Model B, where students rated their motivational beliefs and teaching quality throughout the academic year for the same type of lecture (most difficult or most important). The variances described, however, overall were highest in Model C. The most heterogeneous relationships were shown in Models C and D. Finally, the findings from Models C and D highlight the relevance of motivational beliefs in a variety of situations to students' ratings of the quality of teaching in higher education. This may indicate that IV and EXP serve as characteristics of stable people that influence how students evaluate instructional quality throughout situations.

DISCUSSION

Students' ranking of teaching quality is assumed to be a combination of the effect of people, the effect of situations, and the effect of situations × people. To understand the importance of teaching quality to students' academic growth, it is necessary to understand which components are captured when asking students about their perceptions of different dimensions of teaching quality. This study discusses the gaps in previous research in understanding the components of student rating on teaching quality that are consistent and specific to the situation by considering various characteristics of lectures and teaching situations (RQ1). This research uses a unique dataset from higher education, where students assess various lectures in each academic year. In the second step (RQ2), the research focused on the components that are consistent in ranking teaching quality (the people effect) and their relationship with students' initial motivational beliefs.

Consistency and specificity of the situation hint at the situation of students' rankingon the quality of teaching

Previous studies have identified that students' ratings of teaching quality vary somewhat throughout lecture time, lecture material, lecturer popularity and/or perceived lecture difficulties.[42] Related to the first research question (RQ1), this study expands on previous research that is largely from primary and secondary education settings.[43] This study found variation in ranking consistency across the investigated situations, i.e., the highest level of consistency in ranking the quality of teaching in the same lecture with the same lecturer at different points in time. As expected, the consistency of teaching quality ratings decreases when considering multiple situations, for example, different lectures, different time points or important and difficult lectures. This research shows that students incorporate their experiences across teaching situations into their rankings, and student evaluations become more nuanced. Nuanced means that student evaluations can reflect more situationally specific experiences when assessing the quality of teaching in a variety of different teaching situations. The joint construction between teachers and students in important teaching situations, and with the variety of teaching situations investigated, also the consistency decreases.[44] But even when teaching situations differ, about one-third of the variation in students' ratings of teaching quality can still be attributed to consistent people effects. This shows that students apply general tendencies in their evaluations, even in different contexts. When students evaluate the quality of teaching, students' general ranking tendencies are crucial.[45] The analysis of this study shows that this conclusion is also supported in the context of higher education. For researchers and evaluators to accurately distinguish between people and situation-specific effects, student evaluations in some lectures must be examined. This is in order to gain a more precise understanding of the effectiveness of teaching. The examination also gained an understanding of the assessment on which aspects of student rating on teaching quality contributed the most to students' academic development.

Considering the three dimensions of teaching quality, some differences were found in the consistency and specificity of situations across the model. Previous studies have shown a higher person effect for clarity of teaching, monitoring, and classroom management than emotional support or autonomy, likely because these strategies are more observable and shared among students within and across lectures, compared to more specific related aspects, such as emotional support or autonomy.[46] In some courses that were considered difficult or important (Model B) or some types of courses that were assessed during an academic year (Model D), it was somewhat in line with previous research showing that student ratings of student support had the highest consistency in the same course (Model A). These results show that the ranking of student support is specifically shaped by the joint construction between students and lecturers in the same lecture and that the relationship between students and lecturers varies across lectures with different lecturers. Classroom management ratings and, in some cases, cognitive activation, on the other hand, are more consistent across some lectures (Models B and C). These findings have significant implications for the measurement and modeling of teaching quality in colleges, namely ensuring the validity of measurement instruments by distinguishing consistent perceptions and specific situations, and improving contextual interpretations of student rankings by taking into account the effects of people.

Mempertimbangkan tiga dimensi kualitas pengajaran, ditemukan beberapa perbedaan dalam konsistensi dan spesifisitas situasi di seluruh model. Studi sebelumnya menunjukkan efek orang yang lebih tinggi untuk kejelasan pengajaran, pemantauan, dan manajemen kelas dibanding dukungan emosional atau otonomi, kemungkinan karena strategi ini lebih dapat diamati dan dibagikan di antara mahasiswa di dalam dan di seluruh perkuliahan, dibandingkan dengan aspek yang lebih spesifik terkait, seperti dukungan emosional atau otonomi (Göllner et al., 2018; Wagner et al., 2016). Di beberapa perkuliahan yang dianggap sulit atau penting (Model B) atau beberapa jenis perkuliahan yang dinilai selama satu tahun akademik (Model D), agak sejalan dengan penelitian sebelumnya yang menunjukkan bahwa pemeringkatan mahasiswa atas dukungan mahasiswa memiliki konsistensi tertinggi dalam perkuliahan yang sama (Model A). Hasil ini menunjukkan bahwa pemeringkatan dukungan mahasiswa secara khusus dibentuk oleh konstruksi bersama antara mahasiswa dan dosen dalam perkuliahan yang sama dan bahwa hubungan antara mahasiswa dan dosen bervariasi di berbagai perkuliahan dengan dosen yang berbeda. Pemeringkatan manajemen kelas dan, dalam beberapa kasus, aktivasi kognitif, sebaliknya, lebih konsisten di beberapa perkuliahan (Model B dan C). Temuan ini memiliki implikasi yang signifikan terhadap pengukuran dan pemodelan kualitas pengajaran di perguruan tinggi, yaitu memastikan validitas instrumen pengukuran dengan membedakan persepsi yang konsisten dan situasi spesifik, serta meningkatkan interpretasi kontekstual pemeringkatan mahasiswa dengan memperhitungkan efek orang.

Motivational beliefs are related to the effect of people in student ratings on teaching quality

Researchers still know little about which teaching quality rankings reflect the strongest person effect of student ratings on teaching quality, and which ones reflect the strongest indicators of situation-specific effects.[47] Previous studies using other measures have found that students who are initially interested in a lecture are more engaged in lectures and rate teaching and lecturers in lectures more positively than students who are less interested in lectures at the beginning of lectures.[48] The same pattern exists for the expectation of success and ranking of teaching quality.[49] The findings of this study show that students with higher motivational beliefs (higher expectations of success, greater interest) also give more positive ratings on measures of teaching quality. These findings suggest that a consistent component of student ratings of teaching quality – the people effect – is at least partly explained by students' motivational beliefs.[50] To illustrate this using the results of Model C: if the constant component of student rating averages 37% for classroom management and 33% for cognitive activation, and if motivational beliefs explain an average of 53% and 75% of the variance in these components, this suggests that about 20% to 25% of the total teaching quality ratings can be associated with motivational beliefs. These insights are particularly relevant for interpreting student evaluations, especially when used to draw conclusions about the quality and effectiveness of teaching. For example, for motivational beliefs. The results of this study show that motivational beliefs are not only shaped by students' perceptions of teaching quality, but also significantly influence how students assess the quality of teaching. Higher motivational beliefs lead to information processing that is more consistent with stereotypes or that more motivated students are more consistently engaged throughout lectures, and in turn result in more consistent and more positive teaching quality ratings across situations.

The results of this study also show that lecture characteristics affect the extent to which motivational beliefs encourage the ranking of student teaching quality. For cognitive activation, particularly in Model A, not only were students more motivated reporting more positive ratings of teaching quality, but also that associations appeared stronger in difficult lectures. This shows that motivated students rank more positively in a challenging learning environment compared to low-motivated students. The results of this study are in line with the assumptions of the theory of the suitability of people × environment (Eccles et al., 1993; Rubach et al., 2022) which showed that students who were more interested in or had higher expectations of success in lectures, were able to experience stronger matches with challenging lectures, and resulted in more positive instructional quality rankings.[51]

CONCLUSION

Teaching quality assessments are very complex, as there is consistency and situational specificity in the rankings. This dual nature makes it difficult to interpret lecture evaluations and emphasizes the need to differentiate between what students bring into a learning situation and what happens in it.

The consistency and specificity model of situations helps extract the consistent person effect when accounting for different teaching situations, i.e. different lectures, lecture types, and different time points (Models B, C and D). This study highlights that the ranking of teaching quality depends on the characteristics of the lectures, namely the number of lectures assessed, the time of the lectures, and how important or difficult the lectures are perceived by students. In particular, the consistent component in ranking teaching quality in a moderate to strong manner is related to student motivational beliefs, in particular, the value of students' interests and expectations of success. These insights contribute to a deeper understanding of the psychological basis of student ratings of teaching quality in higher education.

REFERENCES

[1] Anna-Katharina Praetorius, *Messung von Unterrichtsqualität Durch Ratings.*, Pädagogische Psychologie Und Entwicklungspsychologie. 90, Waxmann, 2014, 98.

[2] Sonja Bieg et al., "Teachers' Enthusiasm and Humor and Its' Lagged Relationships with Students' Enjoyment and Boredom - A Latent Trait-State-Approach," *Learning and Instruction* 81 (October 2022): 101579, https://doi.org/10.1016/j.learninstruc.2021.101579.

[3] Gabriela S. Blum et al., "The Nonlinear Interaction of Person and Situation (Nips) Model: Theory and Empirical Evidence," *European Journal of Personality* 32, no. 3 (2018): 286–305, https://doi.org/10.1002/per.2138.

[4] Rauthmann, J. F. (2015). Structuring situational information: A road map of the multiple pathways to different situational taxonomies. *European Psychologist*, 20(3), 176–189. https://doi.org/10.1027/1016-9040/a000225

[5] Richard Göllner et al., "Students' Idiosyncratic Perceptions of Teaching Quality in Mathematics: A Result of Rater Tendency Alone or an Expression of Dyadic Effects between Students and Teachers?," *Journal of Educational Psychology* 110, no. 5 (2018): 709–25, https://doi.org/10.1037/edu0000236.

[6] Jacquelynne S. Eccles and Allan Wigfield, "From Expectancy-Value Theory to Situated Expectancy-Value Theory: A Developmental, Social Cognitive, and Sociocultural Perspective on Motivation," *Contemporary Educational Psychology* 61 (April 2020): 101859, https://doi.org/10.1016/j.cedpsych.2020.101859.

[7] Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). *Generic dimensions of teaching quality*: The German framework of three basic dimensions. ZDM, 50(3), 407–426. https://doi.org/10.1007/s11858-018-0918-4

[8] Decristan, J., Hess, M., Holzberger, D., & Praetorius, A.-K. (2020). Oberflächen- und Tiefenmerkmale: eine Reflexion zweier prominenter Begriffe der Unterrichtsforschung [Surface and deep structures: Reflecting on two key constructs in instructional research]. *Zeitschrift Für Pädagogik*, 66(1), 102–116.

[9] Klieme, E. (2020). Guter Unterricht – auch und besonders unter Einschränkungen der Pandemie? [Good teaching – even and especially under pandemic restrictions?]. In D. Fickermann & B. Edelstein (Eds.), Langsam vermisse ich die Schule … (pp. S.117–S.135). *Waxmann*. https://doi.org/10.31244/9783830992318.07

[10] Helmke, A., & Schrader, F.-W. (2010). Hochschuldidaktik [higher education didactics]. In D. H. Rost (Ed.), *Handwörterbuch pädagogische Psychologie* (4th ed., pp. 273–279). Beltz.

[11] Evertson, C. M., & Poole, I. R. (2008). *Proactive classroom management. In T. Good (Ed.), 21st century education*: A reference hand-book (pp. 131–139). Sage. https://doi.org/10.4135/9781412964012.n14

[12] Leuders, T., & Holzäpfel, L. (2011). Kognitive Aktivierung im Mathematikunterricht [cognitive activation in mathematics]. *Unterrichtswissenschaft*, 39(3), 213–230.

[13] Beran, T., & Violato, C. (2005). Ratings of university teacher instruction: How much do student and course characteristics really matter? *Assessment & Evaluation in Higher Education*, 30(6), 593–601. https://doi.org/10.1080/0260293050 0260688

[14] Rauthmann, J. F. (2015). Structuring situational information: A road map of the multiple pathways to different situational taxonomies. *European Psychologist*, 20(3), 176–189. https://doi.org/10.1027/1016-9040/a000225

[15] Blum, G. S., Rauthmann, J. F., Göllner, R., Lischetzke, T., & Schmitt, M. (2018). The nonlinear interaction of person and situation (Nips) model: Theory and empirical evidence. *European Journal of Personality*, 32(3), 286–305.

[16] Gillmore, G. M., Kane, M. T., & Naccarato, R. W. (1978). The generalizability of student ratings of instruction: Estimation of the teacher and course components. *Journal of Educational Measurement*, 15(1), 1–13. https://doi.org/10.1111/j.1745-3984.1978.tb00051.x

[17] Carpenter, S. K., Witherby, A. E., & Tauber, S. K. (2020). On students' (mis)judgments of learning and teaching effectiveness. *Journal of Applied Research in Memory and Cognition*, 9(2), 137–151. https://doi.org/10.1016/j.jarmac. 2019.12.009

[18] Gaertner, H., & Brunner, M. (2018). Once good teaching, always good teaching? The differential stability of student perceptions of teaching quality. *Educational Assessment, Evaluation and Accountability*, 30(2), 159–182. https://doi.org/10.1007/s11092-018-9277-5

[19] Daumiller, M., Janke, S., Hein, J., Rinas, R., Dickhäuser, O., & Dresel, M. (2023). Teaching quality in higher education. *European Journal of Psychological Assessment*, 39(3), 176–181. https://doi.org/10.1027/1015-5759/a000700

[20] Fauth, B., Göllner, R., Lenske, G., Praetorius, A. K., & Wagner, W. (2020). Who sees what? Conceptual considerations on the measurement of teaching quality from different perspectives. In A.-K. Praetorius, J. Grünkorn, & E. Klieme (Eds.), Empirische Forschung zu Unterrichtsqualität: Theoretische Grundfragen und quantitative Modellierungen (pp. 138–155). *Beltz Juventa.*

[21] Fauth, B., Wagner, W., Bertram, C., Göllner, R., Roloff, J., Lüdtke, O., Polikoff, M. S., Klusmann, U., & Trautwein, U. (2020). Don't blame the teacher? The need to account for classroom characteristics in evaluations of teaching quality. *Journal of Educational Psychology*, 112(6), 1284–1302. https://doi.org/10.1037/edu00 00416

[22] Bieg, S., Dresel, M., Goetz, T., & Nett, U. E. (2022). Teachers' enthusiasm and humor and its' lagged relationships with students' enjoyment and boredom: A latent trait-state-approach. *Learning and Instruction*, 81, 101579. https://doi.org/10.1016/j.learninstruc.2021.101579

[23] Rubach, C., von Keyserlingk, L., Simpkins, S. D., & Eccles, J. S. (2022b). Does instructional quality impact male and female university students differently? Focusing on academic stress, academic satisfaction, and mental health impairment. *Frontiers in Education*, 7, 820321.

[24] Marsh, H. W. (1983). Multidimensional ratings of teaching effectiveness by students from different academic settings and their relation to student/course/instructor characteristics. *Journal of Educational Psychology*, 75, 150–166.

[25] Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, 61, 101859. https://doi.org/10.1016/j.cedpsych. 2020.101859

[26] Plass, J. L., & Kalyuga, S. (2019). Four ways of considering emotion in cognitive load theory. Educational Psychology Review, 31, 339–359. https://doi.org/10.1007/s10648-019-09473 -5

[27] Biernat, M., Kobrynowicz, D., & Weber, D. L. (2003). Stereotypes and shifting standards: Some paradoxical effects of cognitiveload. *Journal of Applied Social Psychology*, 33(10), 2060–2079. https://doi.org/10.1111/j.1559-1816. 2003.tb01875.x

[28] Rubach, C., Dicke, A. L., Lazarides, R., Simpkins, S., & Eccles, J. S. (2020). Addressing adolescents' depressive symptoms and risky behavior: The role of perceived parents' and teachers' social support. *Journal of Organizational Psychology*, 20(4), 70–101.

[29] Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Eds.), The scholarship of teaching and learning in higher education: An evidence-based perspective (pp. 319–384). *Springer.*

[30] Dahl, D. W., & Smimou, K. (2011). Does motivation matter? On the relationship between perceived quality of teaching and students' motivational orientations. *Managerial Finance*, 37(7), 582–609. https://doi.org/10.1108/03074351111140243

[31] Olivares, O. J. (2001). Student interest, grading leniency, and teacher ratings: A conceptual analysis. Contemporary Educational Psychology, 26(3), 382–399. https://doi.org/10.1006/ceps. 2000.1070

[32] Patrick, C. L. (2011). Student evaluations of teaching: Effects of the big five personality traits, grades and the validity hypothesis. *Assessment and Evaluation in Higher Education*, 36(2), 239–249. https://doi.org/10.1080/02602930903308258

[33] Hidi, S., & Renninger, K. A. (2006). *The four-phase model of interest development. Educational Psychologist*, 41(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

[34] Steinhart, Y., & Wyer, R. S., Jr. (2009). Motivational correlates of need for cognition. European Journal of Social Psychology, 39(4), 608–621.

[35] McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review*, 90(3), 215–238. https://doi.org/10.1037/0033-295X.90.3. 215

[36] Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology, 2(2), 175–220.

[37] Feistauer, D., & Richter, T. (2018). Validity of students' evaluations of teaching: Biasing effects of likability and prior subject interest. *Studies in Educational Evaluation*, 59, 168–178. https://doi.org/10.1016/j.stueduc. 2018.07.009

[38] Göllner, R., Wagner, W., Eccles, J. S., & Trautwein, U. (2018). Students' idiosyncratic perceptions of teaching quality in mathematics: A result of rater tendency alone or an expression of dyadic effects between students and teachers? *Journal of Educational Psychology*, 110(5), 709–725. https://doi.org/10.1037/edu00 00236

[39] Eccles, J. S., O'Neill, S. A., & Wigfield, A. (2005). Ability self-perceptions and subjective task values in adolescents and children. In K. A. Moore & L. H. Lippman (Eds.), The search institute series on developmentally attentive community and society. *What do children need to flourish? (Vol. 3, pp. 237–249). Springe*r. https://doi.org/10.1007/0-387-23823 -9_15

[40] von Keyserlingk, L., Rubach, C., Lee, H. R., Eccles, J. S., & Heckhausen, J. (2022). College students' motivational beliefs and use of goal- oriented control strategies: Integrating two theories of motivated behavior. *Motivation and Emotion*, 46(5), 601–620.

[41] Geiser, C. (2021). *Longitudinal structural equation modeling with Mplus: A latent state-trait perspective*. Guilford Publications.

[42] Spooren, P., Brockx, B., & Mortelmans, D. (2013). O*n the validity of student evaluation of teaching.* Review of Educational Research, 83(4), 598–642. https://doi.org/10.3102/0034654313496870

[43] Carpenter, S. K., Witherby, A. E., & Tauber, S. K. (2020). On students' (mis)judgments of learning and teaching effectiveness. *Journal of Applied Research in Memory and Cognition*, 9(2), 137–151. https://doi.org/10.1016/j.jarmac. 2019.12.009

[44] Fauth, B., Wagner, W., Bertram, C., Göllner, R., Roloff, J., Lüdtke, O., Polikoff, M. S., Klusmann, U., & Trautwein, U. (2020). Don't blame the teacher? The need to account for classroom characteristics in evaluations of teaching quality. *Journal of Educational Psychology*, 112(6), 1284–1302. https://doi.org/10.1037/edu00 00416

[45] Göllner, R., Wagner, W., Eccles, J. S., & Trautwein, U. (2018). Students' idiosyncratic perceptions of teaching quality in mathematics: A result of rater tendency alone or an expression of dyadic effects between students and teachers? *Journal of Educational Psychology*, 110(5), 709–725. https://doi.org/10.1037/edu00 00236

[46] Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B., & Trautwein, U. (2016). Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. *Journal of Educational Psychology*, 108(5), 705–721. https://doi.org/10.1037/edu00 00075

[47] Göllner, R., Wagner, W., Eccles, J. S., & Trautwein, U. (2018). Students' idiosyncratic perceptions of teaching quality in mathematics: A result of rater tendency alone or an expression of dyadic effects

between students and teachers? *Journal of Educational Psychology*, 110(5), 709–725. https://doi.org/10.1037/edu0000236

[48] Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Eds.), The scholarship of teaching and learning in higher education: An evidence-based perspective (pp. 319–384). *Springer.*

[49] Patrick, C. L. (2011). Student evaluations of teaching: Effects of the big five personality traits, grades and the validity hypothesis. *Assessment and Evaluation in Higher Education*, 36(2), 239–249. https://doi.org/10.1080/02602930903308258

[50] Griffin, B. W. (2004). Grading leniency, grade discrepancy, and student ratings of instruction. *Contemporary Educational Psychology*, 29(4), 410–425. https://doi.org/10.1016/j.cedpsych. 2003.11.001

[51] Rubach, C., von Keyserlingk, L., Simpkins, S. D., & Eccles, J. S. (2022a). Motivational beliefs and positive achievement emotions during COVID-19: A person-environment fit perspective in higher education. In H. Burgsteiner & G. Krammer (Eds.), Impacts of COVID-19 Pandemic's distance learning on students and teachers in schools and in higher education. *International perspectives* (pp. 100–125). Leykam. https://doi.org/10.56560/isbn.978-3-7011-0496-3_ 6