

**Review of PhD Dissertation by mgr Mengyu Cao: “The Role of the Education Quality Assurance System in Shaping Relationships Among University Education Quality, Academic Citizenship Behaviour and Academic Performance”**

Author: mgr Mengyu Cao

Dissertation Program: Interdisciplinary PhD School *Academia Rerum Socialium*, Nicolaus Copernicus University (Faculty of Economic Sciences and Management)

Thesis Supervisors: Prof. Rafał Haffer & Prof. Øivind Strand

Reviewer: prof. dr hab. Łukasz Sułkowski

## Introduction

Mengyu Cao’s doctoral dissertation addresses a timely and significant issue in higher education management: how formal Education Quality Assurance Systems (QAS) influence educational quality and key student outcomes – specifically *student loyalty, academic citizenship behaviour, and academic performance* – and how these relationships may differ across cultural contexts. The dissertation’s central purpose is “*to investigate how the design of the educational quality assurance system and its implemented activities affect the variables university education quality (UEQ), student loyalty (SL), academic citizenship behaviour (ACB), and academic performance (AP), as well as the relationships among them in different cultural contexts*”. In pursuit of this goal, the author employs a mixed-methods case study approach, combining quantitative survey research with qualitative case studies at two European universities: Nicolaus Copernicus University (NCU) in Poland and the Norwegian University of Science and Technology (NTNU) in Norway. These institutions were deliberately chosen for their contrasting national cultures (Poland vs. Norway) according to Hofstede’s cultural dimensions, while both operate under the common European Bologna Process framework, providing structural comparability. This design enables an exploration of cross-cultural differences in QAS efficacy while holding constant certain educational quality standards.

In what follows, this review critically evaluates the dissertation across several dimensions: the clarity and relevance of its research problem formulation (including research gaps, questions, and hypotheses); its originality and contribution to the fields of management and quality sciences (with emphasis on cross-cultural insights); the methodological rigor of its mixed-methods design, data collection, and analytical techniques; the quality of its theoretical framework and literature integration (e.g. use of Hofstede’s model and Social Exchange Theory); the consistency, structure, and clarity of argumentation and writing; the presentation and discussion of findings, including how well the author triangulates evidence and articulates limitations and implications; and finally, *critical and polemical aspects* where the work could be challenged or extended. Three substantive questions are posed to the PhD candidate regarding research design choices, cultural interpretation of results, and generalizability of findings. The review concludes with a recommendation on the acceptance of the dissertation for the PhD degree.

## 1. Clarity and Relevance of the Research Problem, Research Gap, and RQs/Hypotheses

The dissertation establishes a clear and compelling research problem at the intersection of educational quality management and student behavior/outcomes. The author observes that while universities worldwide implement Quality Assurance Systems to enhance teaching and learning, there is limited understanding of *how these systems actually translate into improved educational quality and student success*, or how cultural context might alter their effectiveness. This problem is highly relevant to the field of Management and Quality Sciences, especially as higher education becomes increasingly internationalized. By focusing on *student loyalty* (analogous to student satisfaction and commitment), *academic citizenship behaviour* (an adaptation of the organizational citizenship behavior concept to the academic context), and *academic performance*, the study targets outcomes that are both theoretically and practically important. Notably, the concept of Academic Citizenship Behaviour (ACB) – essentially students' voluntary, constructive contributions to the academic community – is relatively novel in this context and aligns with contemporary views of students as partners or active stakeholders in education, rather than mere consumers. The dissertation frames these constructs within the scope of a QAS, asking in essence: *Does a well-designed quality assurance system foster better educational quality, and if so, does that lead to more loyal, civically engaged students who perform better academically – and do these dynamics differ between different cultural settings?* This line of inquiry addresses a meaningful gap at the nexus of quality management and organizational behavior in education.

Research gaps are explicitly identified and well-grounded in a critical literature review. The author systematically pinpoints several specific gaps in prior research, which serve to justify the dissertation's questions and hypotheses. For example, it is noted that *"the relationship between university education quality (UEQ) and academic citizenship behaviour (ACB) remains largely unexplored"* in higher education settings – past studies of service quality and extra-role behaviors have mostly been in business or general education contexts treating students as customers, with only scant attention to students' role as community members. Likewise, while many studies have examined education quality → student satisfaction/loyalty links, *cross-cultural comparative research* on this relationship is limited. The author also finds a dearth of research on the link between student loyalty and citizenship behavior in academia (only a few studies, and those using a narrow "student as customer" lens), and on loyalty's mediating role in linking quality to citizenship. Further gaps include the need to examine educational quality → academic performance more broadly (beyond single-course contexts), the influence of ACB on academic performance (only limited evidence, often not cross-national), and importantly, the lack of attention to cultural moderators of all the above relationships in higher education. By articulating these gaps, the author demonstrates a strong command of the literature and provides a clear rationale for the dissertation's focus. The identified gaps directly inform the study's objectives and ensure the research is positioned to contribute new knowledge (see Section 2 of this review for more on originality).

From the stated problem and gaps, the dissertation derives six research questions (RQs) and eleven hypotheses, achieving a good coherence between qualitative inquiry and quantitative testing. The RQs are framed for an in-depth case study examination of each university's QAS

and collectively cover both implementation and outcomes of the quality assurance process. They ask, in sequence:

1. *What measurements are implemented in the educational quality assurance system of this university?* – addressing the tools/indicators used in QAS (e.g. surveys, metrics).
2. *What procedures for improving educational quality, student satisfaction, and the educational quality assurance system are used at this university?* – focusing on the processes for quality improvement and closing the feedback loop.
3. *To what extent does the educational quality assurance system contribute to improving educational quality and student satisfaction in this university?* – probing the perceived effectiveness or impact of QAS on quality and satisfaction.
4. *How does this university handle the educational quality assurance system?* – a broad question on governance/management of QAS (perhaps how responsibilities are allocated, challenges handled, etc.).
5. *How do students perceive the quality assurance system?* – bringing in the student perspective explicitly.
6. *What are the similarities and differences between the QAS of the two universities?* – enabling direct cross-case comparison.

These research questions are clearly stated in the dissertation and align well with the identified gaps and the mixed-methods approach. RQ1–4 focus on each single case (NCU or NTNU) to understand internal QAS workings and effects; RQ5 adds the stakeholder (student) viewpoint; and RQ6 then synthesizes insights across the two cases, directly addressing the cross-cultural dimension. Taken together, the RQs ensure that the qualitative component will richly contextualize how QAS operates and is experienced in each culture, which complements the more generalized hypotheses tested via survey data.

The hypotheses (H1–H11) correspond to a carefully constructed theoretical model linking the study's four primary constructs (UEQ, SL, ACB, AP) and incorporating cultural moderators. Eleven hypotheses may seem extensive, but they flow logically from the literature review's gap analysis and theoretical reasoning. In essence, the author proposes the following (all stated as *positive* relationships or effects):

- Direct effects among main variables: H1: Higher perceived University Education Quality → higher Academic Citizenship Behaviour; H2: Higher University Education Quality → greater Student Loyalty; H3: Greater Student Loyalty → more Academic Citizenship Behaviour; H5: Higher University Education Quality → better Academic Performance; H6: More Academic Citizenship Behaviour → better Academic Performance. These cover all pairwise links in the presumed chain from quality to performance, reflecting theories like service quality → loyalty (from marketing literature) and OCB → performance (from organizational behavior).
- Mediating effects: H4: Student Loyalty mediates the effect of Education Quality on Academic Citizenship Behaviour (i.e., quality fosters loyalty which in turn fosters citizenship); H7: Academic Citizenship Behaviour mediates the effect of Education Quality on Academic Performance (i.e., quality improves performance indirectly by encouraging citizenship behaviors that help learning).

- Moderating (cultural) effects: H8: Power Distance moderates the UEQ→SL relationship; H9: Masculinity–Femininity moderates the ACB→AP relationship; H10: Masculinity–Femininity moderates the UEQ→AP relationship; and H11: Individualism–Collectivism moderates the SL→ACB relationship. These moderation hypotheses stem from Hofstede's cultural dimensions theory, anticipating that differences in hierarchical orientation, gender-role orientation, and collectivist values will alter the strength of certain paths in the model.

All hypotheses are explicitly listed and numbered in the dissertation. Importantly, they are not arbitrary but each is anchored in prior research or theory. For instance, H2 (UEQ→SL) is well-grounded in service quality research (many studies have found that better perceived quality leads to higher student satisfaction and loyalty), whereas H1 (UEQ→ACB) is more exploratory, justified by social exchange theory (students reciprocate good education with extra-role behaviors) and identified as an open gap. The moderation hypotheses are based on documented cultural differences between Poland and Norway – for example, Poland scores much higher on Power Distance and Masculinity, while Norway scores higher on Individualism – and extant findings that culture can influence service evaluations and organizational behaviors. Thus, the hypotheses demonstrate a thorough integration of cross-cultural considerations into the core research model.

Overall, the research questions and hypotheses are coherent and well-aligned with the stated problem and gaps. The mix of qualitative RQs and quantitative hypotheses might have risked disjointedness, but here they clearly complement each other under a unifying purpose. The RQs ensure the study will capture detailed processual and contextual insights about QAS implementation (addressing “how” and “what” aspects), while the hypotheses test the presence and strength of theorized relationships between key constructs (addressing “to what extent” and “why” aspects). The author has communicated these elements with clarity. The relevance is high: universities and quality assurance agencies can benefit from understanding these dynamics, and the cross-cultural angle broadens the applicability of the findings. In sum, the dissertation's front end – the problem statement, gap identification, and formulation of RQs/hypotheses – is a model of clarity and logical consistency. It sets a solid foundation for the research and signals to the reader exactly what will be investigated and why it matters.

## 2. Originality and Contribution to Management and Quality Sciences (Cross-Cultural Dimensions)

Mengyu Cao's work makes a strong original contribution to the fields of management (particularly organizational behavior in educational settings) and quality sciences (particularly quality management in higher education). Several facets of originality stand out:

- Integration of Diverse Literature Streams: The dissertation sits at an intersection of multiple research domains that are not often linked. It explicitly overlaps literature on university education quality, student satisfaction/loyalty, organizational citizenship (student citizenship behavior), academic performance, and national culture. The author even provides a Venn diagram illustrating how these streams intersect in the thesis. By doing so, the research moves beyond siloed approaches (e.g. studying quality assurance in isolation, or student behavior in isolation) and

instead examines the *interplay* between institutional quality mechanisms and individual behaviors/outcomes. This interdisciplinary approach is innovative and yields a more holistic understanding of educational quality dynamics.

- Introducing “Academic Citizenship Behaviour” in Quality Assurance Context: The adaptation of the Organizational Citizenship Behaviour (OCB) concept to an academic student context – termed Academic Citizenship Behaviour (ACB) – is a novel element. While the concept of students going “above and beyond” (helping peers, actively participating in academic life, etc.) has been discussed by some scholars, its formal incorporation into a model of educational outcomes linked with quality is new. The dissertation not only uses ACB as a key outcome but measures it in a multidimensional way (applying a five-dimension OCB scale covering altruism, courtesy, civic virtue, sportsmanship, conscientiousness). This comprehensive assessment of students’ extra-role behaviours goes further than prior studies, which often narrowly treated student “citizenship” as participation or cooperation. By demonstrating that ACB can be meaningfully measured and linked to academic success, the study opens a new avenue for research in student engagement and development. It effectively positions students as active contributors to quality, not just passive recipients – a perspective that addresses calls in the literature to redefine students’ roles.
- Novel Hypotheses and Empirical Evidence: Many of the relationships tested were empirically unexamined before, especially in higher education. For example, the hypothesis that *student loyalty mediates the impact of education quality on citizenship behavior* (H4) is, to the author’s knowledge, the first such test in a university context. Similarly, examining whether *ACB mediates the effect of quality on performance* (H7) appears to be a first in the literature – this is an innovative idea linking institutional quality to outcomes via student behavior. The results of these tests, regardless of support, contribute new data points to theory (indeed H7 was not supported in the end, offering a finding that this mediation did not occur, which is valuable information for theory refinement). The cultural moderation hypotheses (H8–H11) are also cutting-edge for the education context. While Hofstede’s dimensions have been widely studied in business settings, their influence on student perceptions and behaviors in university quality assurance has been *largely unexplored*. By successfully executing a cross-cultural comparison, the author provides fresh insights – for instance, the finding that collectivism moderates the link between loyalty and citizenship behavior (supported for Poland vs. Norway) is a novel contribution, suggesting that in more collectivist cultures student loyalty might translate more readily into helping behaviors. In contrast, the lack of effect of power distance (H8 unsupported) is also revealing: it challenges assumptions and suggests that hierarchy might not strongly dampen students’ loyalty response to quality in these contexts. Such nuanced contributions refine our understanding of cultural theories in educational environments.
- Cross-Cultural Design and Findings: The dissertation’s cross-cultural dimension is a major contributor to its originality. By comparing Poland and Norway, the study leverages a natural experiment of sorts: two countries with sharply contrasting cultural profiles (e.g. Poland higher in power distance and masculinity; Norway higher in femininity and individualism), yet both part of a common European higher education area. This design is innovative because it isolates cultural effects in a

controlled way – both universities have mature QAS frameworks aligned with European standards, so differences in outcomes can more credibly be attributed to cultural/contextual differences rather than completely different educational systems. The author's selection of these two cases was also aided by her personal familiarity (being a student at NCU and visiting NTNU), which is a creative leveraging of available opportunities to gain insider insight. The cross-cultural findings themselves add new knowledge: for example, the QAS implementation approaches differ markedly – *"NCU employed a more hierarchical, survey-based system with formal improvement processes, while NTNU utilized a dual approach combining surveys with participatory 'reference groups' allowing real-time feedback"*. This comparative insight is valuable for the management of quality systems, illustrating how cultural values (e.g. Poland's more hierarchical culture vs. Norway's egalitarian ethos) manifest in the way student feedback is gathered and acted upon. Another interesting cross-cultural result is that education quality had a direct positive effect on academic performance only in Norway, not in Poland, whereas academic citizenship behaviour had a positive effect on performance only in Poland. This suggests that Norwegian students may benefit directly from high-quality instruction (perhaps due to a supportive learning culture), while Polish students' academic success is more tightly linked to their own engagement and extra efforts. These are novel interpretations that the author offers, contributing to cross-cultural management theory by highlighting how national culture can channel the pathways through which quality influences outcomes. The dissertation's stance that *"the resolution of the quality assurance problem has an unbreakable connection to cultural context: while quality systems do eventually influence educational results, the pathways through which this occurs vary among cultural settings"* is a fresh and valuable contribution to quality assurance discourse.

- Practical Contributions: From a practical standpoint, the dissertation provides actionable insights for university administrators and policy-makers. It advocates for culturally sensitive quality assurance systems – rather than one-size-fits-all models – arguing that QAS should be “tuned to represent local attitudes and practices while adhering to fundamental educational quality standards”. This is an important contribution to quality management practice, as international universities or those seeking global accreditation often import QA models without adaptation. The study offers evidence that adapting to local culture (for instance, introducing more participatory mechanisms in high-individualism settings, or clear formal processes in high power-distance settings) can enhance effectiveness. Additionally, by showing the importance of student loyalty and citizenship behaviours, the work suggests that engaging students as partners in quality (e.g. involving them in feedback committees, recognizing their contributions) could improve both the educational environment and student success. These contributions tie management theory to educational practice in a meaningful way.

In summary, the dissertation is original in form and content. It pioneers a complex, interdisciplinary investigation that few have attempted, thereby filling multiple gaps. The combination of mixed methods, cross-cultural comparison, and extension of organizational behavior concepts to education yields a dissertation that is more than the sum of its parts. Its contributions to Management and Quality Sciences are evident: it extends theory (by testing new relationships and introducing cultural moderators) and provides guidance for

quality management practice in higher education. The cross-cultural dimension, in particular, enriches the field's understanding – a contribution that cannot be achieved by single-context studies. The author's work will be of interest not only to scholars of quality in higher education, but also to those in international management, organizational behavior (studying citizenship behaviours in new contexts), and anyone concerned with how to improve universities in an increasingly global educational landscape.

### 3. Methodological Rigor: Mixed-Methods Case Study Design and Analytical Approaches

The dissertation employs a mixed-methods case study design with commendable rigor. The methodology is well-justified and executed in a manner that enhances the credibility of the findings through complementary qualitative and quantitative evidence. Here we evaluate key aspects: research design, data collection procedures, sample and case selection, and the use of analytical tools (PLS-SEM and thematic analysis).

The choice of a mixed-methods case study is well-suited to the research problem, which spans both understanding *processes/practices* (best captured qualitatively) and testing *relationships between variables* (best captured quantitatively). According to the abstract, the study "*employed a mixed-methods-case study approach, combining qualitative case study techniques with quantitative survey methodology*". The research design is clearly depicted in the dissertation (including a figure of the mixed-method framework) and follows a structured approach: first planning and conducting two case studies (NCU and NTNU) with interviews and document analysis, and simultaneously conducting a student survey in those contexts; then integrating the results. The author references established mixed-methods guidelines (Creswell & Plano Clark) to integrate and compare qualitative and quantitative results, indicating a methodologically informed approach. This integration (triangulation) design is a strength: quantitative findings can be interpreted with the rich context from qualitative insights, and any convergences or divergences can be explained, thus strengthening the validity of conclusions.

Using NCU (Poland) and NTNU (Norway) as the two case studies is a strategic design decision that supports the study's aim of examining cultural influences. The dissertation explains that these universities were chosen due to "their contrasting cultural profiles based on Hofstede's cultural dimensions," while both have established quality assurance frameworks under the Bologna Process. This selection ensures a most-different case comparison on cultural variables, while controlling for certain system-level factors (both are public research universities with comparable QA obligations). The author's personal familiarity with these institutions (an insider perspective at NCU and observational knowledge of NTNU) is mentioned as well, which likely facilitated access to data and nuanced understanding – a smart use of available resources to strengthen case study depth. Overall, the case selection is well-justified and appropriate for the research questions. It enables insights not only within each case but also a robust cross-case analysis (addressing RQ6).

Quantitative Component – Survey Design and PLS-SEM: The quantitative part involved a survey of students at the two universities. The dissertation reports that data were collected from 242 business students (165 from Poland and 77 from Norway) using validated scales for all constructs: UEQ, SL, ACB, AP, as well as cultural dimensions. Targeting business students provided a focused sample within each institution (NCU's and NTNU's business faculties),

which is sensible for consistency; it also aligns with the author's field of management. The sample size of 242, while modest, is adequate for Partial Least Squares Structural Equation Modeling (PLS-SEM), especially given the exploratory nature of some hypotheses and the inclusion of interaction (moderation) effects. The author notes that the sample exceeds minimum requirements for PLS-SEM and follows best-practice guidelines. Indeed, PLS-SEM was a fitting choice of analysis tool here. PLS is well-suited for predictive models and can handle smaller samples better than covariance-based SEM, and it easily incorporates formative constructs or second-order factors if needed (the dissertation's measurement model likely had multiple dimensions for UEQ and ACB). The use of WarpPLS 8.0 software is mentioned, and the author carefully assessed reliability and validity (for instance, reporting composite reliability (CR), Cronbach's alpha, AVE for each construct, and checking for multicollinearity via full collinearity VIF). These details indicate methodological rigor in the quantitative analysis.

Importantly, the author also measured cultural dimensions at the individual level using survey items (rather than assuming all Polish vs. all Norwegian respondents have identical cultural values). This is a rigorous approach because it acknowledges within-country variability. The dissertation even discusses how using validated individual-level cultural scales improves moderation analysis accuracy. By doing so, the study goes beyond a simple country dummy comparison – it empirically tests whether a student's own orientation on, say, Power Distance or Collectivism affects the relationships. This level of detail in cultural measurement enhances the credibility of the moderation findings and addresses a known critique of Hofstede's work (i.e. that national averages may not reflect individual attitudes).

The PLS-SEM analysis was comprehensive: the dissertation includes assessment of the measurement model (with internal consistency and convergent validity stats) and the structural model results for each country, as well as multi-group or moderated analysis. For example, the structural path coefficients and significance for Poland and Norway are reported separately, and the hypothesis support is summarized clearly (which hypotheses held in which country). The quantitative analysis appears rigorous, and the author is transparent about results – even highlighting when certain hypotheses were *not* supported. Additionally, potential bias like common method variance is acknowledged (they cite Podsakoff and note the risk of self-report bias). It's not common to see doctoral students explicitly mention testing for common method variance, so this is a positive sign of methodological thoroughness.

**Qualitative Component – Interviews and Thematic Analysis:** The qualitative component consisted of semi-structured interviews with 15 stakeholders (6 at NCU, 9 at NTNU). These stakeholders included *quality assurance chairs, faculty members, and students* at each university. The inclusion of multiple perspectives (administrators in charge of QA, academic staff, and the student body) is a strong design, as it captures a 360-degree view of how the QAS functions and is perceived. Although the number of student interviewees was relatively small (only 2 at NCU and 3 at NTNU as noted in the limitations), the overall set of interviewees provides rich institutional insight. The interviews were analyzed using thematic analysis and cross-case synthesis. Thematic analysis is an appropriate choice for identifying patterns in qualitative data, and the author likely developed codes corresponding to aspects of the QAS (e.g. "feedback mechanisms," "student involvement," "communication issues,"

etc.) and to the research questions. The use of *cross-case synthesis* means the author did not treat the two cases in isolation, but systematically compared themes across Poland and Norway to draw out similarities and differences. This is vital for answering RQ6 and for triangulating with the quantitative differences observed.

The dissertation's structure indicates that qualitative results are presented question by question for each case, followed by a comparative discussion. For example, the findings detail how each RQ (1 through 6) manifested at NCU and then at NTNU, before synthesizing. This systematic presentation enhances clarity and ensures that the qualitative analysis directly addresses the initial research questions – a sign of methodological discipline. The review of an excerpt shows a clear comparative statement: "*NCU employed a more hierarchical, survey-based system... while NTNU utilized a dual approach combining surveys with participatory 'reference groups'...*". This kind of direct comparison demonstrates effective cross-case analysis rooted in evidence from interviews (indeed, "reference groups" were likely mentioned by NTNU participants as a key practice, and the author highlighted it).

The triangulation of methods is worth special commendation. By integrating quantitative and qualitative findings, the author can validate and enrich the story. For instance, if the survey found high student loyalty at both institutions, the interviews might reveal *why* (perhaps due to certain engagement practices). If a hypothesis was supported in one country but not the other, the qualitative data can offer possible explanations (cultural norms, differences in QAS implementation, etc.). The dissertation explicitly notes that Chapter 4 "*presented integrated findings from both quantitative and qualitative components... and a comprehensive comparative analysis*". This indicates that the author did not keep the two strands separate, but combined them to form a coherent discussion – achieving what mixed-methods design sets out to do. The use of multiple data sources also lends credibility (akin to triangulation for validity). In fact, the methodology chapter emphasizes ensuring rigor through data triangulation.

**Rigor and Ethical Considerations:** Although not explicitly asked, it's worth noting if the dissertation mentioned any steps for quality and ethics. We see evidence that best practices were followed (validated instruments, pilot testing perhaps, etc., though not quoted here). The sample demographics are detailed (age, gender, year of study, etc.), which is good for transparency. The author also acknowledges limitations (see Section 6 of this review) which shows a self-critical rigor. There's no explicit mention of ethics procedures in the excerpts, but presumably interviews had consent, anonymity etc., and given the nature of the study (non-sensitive educational context), there's minimal ethical risk.

**Minor Methodological Limitations:** No study design is without constraints. In this case, the *unequal sample sizes* (165 vs 77) for the survey is one limitation – it can complicate comparisons. The author recognizes that a larger Norwegian sample would improve statistical balance. Additionally, using self-reported academic performance can be considered a limitation due to potential inaccuracy or social desirability bias (also acknowledged as common method variance risk). The qualitative sample, as mentioned, had relatively few student voices; expanding that would yield even more insights. However, these limitations are minor in the face of the study's overall rigor and were, on the whole, handled transparently by the author.

In conclusion, the methodology of this dissertation is a strong point. The mixed-methods design is *well conceived and rigorously implemented*. The quantitative survey with PLS-SEM provides robust statistical testing of the theoretical model, and the qualitative case studies give depth and context, with careful analysis procedures. The combination provides confidence in the findings: where the two methods agree, results are corroborated; where they diverge, the author offers plausible explanations. The methodological choices are justified and aligned with the research questions, reflecting a high level of research design competency by the PhD candidate.

#### 4. Quality of Theoretical Framework and Integration of Literature

The dissertation is grounded in a comprehensive theoretical framework that draws on relevant literature from quality management, organizational behavior, higher education, and cross-cultural psychology. The integration of these literatures is one of the work's notable strengths, as it demonstrates the author's ability to synthesize concepts from different fields into a coherent model for analysis.

**Foundational Theories and Concepts:** Early in the dissertation (Chapter 1), the author reviews the concept of "quality" in higher education, tracing its evolution from basic inspection to modern quality management systems. This likely includes discussions of classic quality frameworks such as the SERVQUAL gaps model (Parasuraman et al., 1985) and Total Quality Management (TQM) principles. Indeed, excerpts show that the text covers SERVQUAL (including its five dimensions and gaps concept) and references to quality gurus like Crosby (Cost of Quality) and Deming, etc. This provides a solid background, situating the university quality assurance system (QAS) within established quality science knowledge. By doing so, the author ensures that the study is anchored in Management and Quality Science theory – not treating QAS as a black box, but as a construct that can be understood via prior models and definitions. For example, understanding the "gaps" in service quality might inform what aspects of educational quality the QAS should monitor (gap between student expectations and perceptions, etc. as in SERVQUAL). This level of theoretical grounding is appropriate for a PhD and shows scholarly depth.

**University Education Quality (UEQ):** The literature review delineates what constitutes *university education quality*. Likely it includes both academic factors (curriculum, teaching, learning outcomes) and service factors (administration, facilities). The author appears to have considered various assessment frameworks and accreditation standards (the mention of AACSB accreditation in context suggests awareness of external quality benchmarks). By reviewing these, the dissertation sets a definition for UEQ and possibly identifies dimensions or indicators used in the survey (for instance, teaching quality, course content relevance, etc.). Integrating literature here ensures clarity on how "quality" is measured – a critical step since quality is a multifaceted and sometimes contested concept in education. It's worth noting the study's innovative stance of treating students not just as consumers but also as partners; literature acknowledging this dual role (e.g. Svensson & Wood, 2007) is cited. This indicates the theoretical framework for UEQ is enriched by stakeholder theory (students as key stakeholders) and by emerging paradigms of co-creation in education.

**Student Loyalty (SL):** For student loyalty, the author draws on both higher education literature and analogies to customer loyalty in marketing. Several references (Helgesen &

Nesset, 2007; Annamdevula & Bellamkonda, 2016, etc.) are mentioned as establishing the quality → loyalty link. Expectancy-Disconfirmation Theory (EDT), which underpins customer satisfaction leading to loyalty (Oliver, 1980; Fornell et al., 1996), is explicitly mentioned as a basis. By using EDT, the author explains that if students' expectations are met or exceeded by the university (high perceived quality), they will be satisfied and more loyal – a well-founded theoretical expectation. The framework likely includes loyalty constructs like willingness to recommend the university, intention to continue to further studies there, etc. Integrating those concepts from marketing into the education context is handled carefully with literature support, which strengthens the hypotheses development for H2 and the like.

**Academic Citizenship Behaviour (ACB):** The concept of ACB is built upon Social Exchange Theory (SET) and Organizational Citizenship Behaviour (OCB) literature. Social Exchange Theory (Blau, 1964) is used to argue that when students receive something valuable (quality education), they may feel obliged to reciprocate with positive behaviors beyond minimum requirements. This is an excellent application of SET to the student context, treating favorable educational experiences as the “benefit” and voluntary student contributions as the “reciprocal act.” The dissertation also invokes Cognitive Consistency Theory (Heider’s balance theory, 1946), which suggests students will align their attitudes and behaviors – for instance, a student who values and is proud of their high-quality university might act in ways (OCB) consistent with that positive attitude, such as helping peers or engaging in campus activities. This dual theoretical underpinning (exchange and consistency) provides a strong logical basis for hypotheses H1, H3, and H4 (in particular H4, loyalty mediating quality → ACB, resonates with consistency theory: satisfied students internalize and then exhibit citizenship behavior). The literature review identifies prior studies of OCB in education – e.g., the author notes only one study (Sharif & Lemine, 2021) looked at quality and OCB in higher ed but treated students purely as customers, and another (Nagy & Marzouk, 2018) looked at loyalty and student citizenship but in a limited way. By discussing these, the author shows mastery of the niche literature on student extra-role behavior, and then justifies using an established five-dimension OCB scale to capture ACB. The integration here is commendable: classic OCB dimensions (altruism, civic virtue, etc.) are reinterpreted in the student context with support from literature (e.g., Allison et al., 2001 on student OCB and academic achievement). Additionally, the author discusses theoretical links like OCB’s known positive impact on performance in workplaces and then argues this should similarly apply in classrooms (H6 hypothesis) – indeed citing evidence that certain OCB dimensions (civic virtue, altruism) positively correlate with academic performance. This shows a tight integration of theory and prior findings to build each piece of the model.

**Academic Performance (AP):** The dissertation treats academic performance as a key outcome. Literature on factors affecting academic performance is referenced (e.g., Ibietan et al., 2016 on how poor quality lowers performance, or Alyahyan & Düşteğör, 2020 on learning environment impacts). This indicates the author didn’t assume a simplistic link; instead they acknowledge multiple influences on performance and carve out where educational quality fits among them. They also noted that prior studies on quality → performance often focused on specific courses or used diverse methods, highlighting the need for a broader investigation (which their study provides). By grounding the expectation that quality can influence grades in past research, the author lends credibility to H5. Moreover, discussing performance literature allows them to consider what *else* might drive performance – in the

limitations, they indeed suggest incorporating variables like *self-efficacy, learning effort, engagement* in future, showing they understand performance is multifactorial.

**Cultural Dimensions (Hofstede's Framework):** The integration of Hofstede's cultural dimensions theory is another highlight. The author doesn't simply mention culture in passing; rather, an entire subsection (Chapter 2.6) is devoted to analyzing how specific cultural dimensions could moderate each relationship. They focus on Power Distance (PD), Masculinity vs. Femininity (MAS/FEM), and Individualism vs. Collectivism (IDV/COLL) – the three dimensions where Poland and Norway differ most significantly. The literature review likely explains each dimension (e.g., PD: acceptance of hierarchy; MAS: emphasis on achievement vs. quality of life; IDV: focus on individual goals vs. group cohesion) and then theorizes their impact. For example, the author hypothesized PD would moderate quality→loyalty (H8) because in a high-PD culture (Poland) students might be less inclined to feel "loyal" or voice feedback due to deference, whereas in low-PD (Norway) they might respond more strongly. Although H8 turned out non-significant, the attempt was grounded in reasonable theory (and the non-finding was interpreted, possibly with the insight that individual-level PD scores didn't align or that modern students are more homogeneous in this regard). For MAS/FEM, the theory was that competitive, achievement-oriented environments (masculine) might make the link between going the extra mile (ACB) and performance outcomes more salient (hence H9, which was supported in Poland), whereas in a feminine culture like Norway, academic performance might be less tied to such behaviors. The author supports these ideas with references: e.g., culture's impact on loyalty formation, or on OCB in workplaces (noting that cultural effects on OCB outcomes have been studied in work settings but not in education). This shows a very thorough integration – bridging macro-cultural theory with micro-educational processes.

The literature integration is exhaustive and up-to-date. The reference list (spanning dozens of recent studies from 2014–2023 as seen in excerpts) indicates the author included contemporary findings. For instance, references to 2022 and 2023 works in service quality and education demonstrate currency. Classic theories (Blau, Heider, Hofstede, Parasuraman) are balanced with modern studies, giving the framework both depth and relevance.

**Coherence of Theoretical Framework:** What's impressive is how the author weaves all these pieces into one model. At the end of the literature review, the dissertation explicitly presents a theoretical framework figure and summary, linking UEQ, SL, ACB, AP with arrows and placing culture as a moderator. The narrative explains how each hypothesis fits into the bigger picture. The result is that by the time the reader reaches the end of Chapter 2, the rationale for every hypothesis is clearly understood and supported by literature. For example, the argument could be summarized as: *High service quality (UEQ) increases student satisfaction/loyalty (EDT theory) and perhaps directly incites extra effort (SET theory); satisfied students are more likely to exhibit citizenship behaviors (consistency and exchange theories); those behaviors can contribute to better performance (OCB theory); also, high quality may directly help performance (educational effectiveness research); and loyalty may intermediate some of these effects (customer loyalty theory). Meanwhile, these processes might look different in Poland vs Norway due to differences in how hierarchy, competition, and group orientation shape student attitudes (cultural theory).* Each clause in

that summary is backed by multiple sources in the dissertation, reflecting *integrative thinking*.

**Critical Note on Theoretical Choices:** One could argue whether alternative theoretical frameworks might also apply – for instance, other cultural models like Schwartz's values or the GLOBE study could have been used, or whether Self-Determination Theory might explain student motivation and citizenship. However, the choices made (Hofstede, SET, EDT, OCB) are the most relevant and widely accepted for the questions at hand, and the author provided justification for them. The discussion on using individual-level culture scores shows theoretical astuteness (addressing the ecological fallacy issue often raised with Hofstede). The reliance on Hofstede might be seen as traditional, but given the stark differences between Poland and Norway in that framework, it was a pragmatic and effective choice – and the author did so thoughtfully.

In conclusion, the theoretical framework and literature review in this dissertation are comprehensive, well-structured, and well-integrated. The candidate demonstrates a deep understanding of each relevant body of knowledge and connects them logically to formulate the research model. Concepts are defined with nuance, theories are used appropriately (not just name-dropping but actually informing hypothesis development), and literature gaps are identified to position the study. This strong theoretical grounding not only lends credibility to the hypotheses, but also means the findings can be meaningfully discussed in light of existing theory (which the author indeed does in the discussion, referencing back to these frameworks). The work's foundation in theory is thus a major strength, elevating it from a purely empirical study to one that significantly engages with and contributes to scholarly discourse.

## 5. Consistency, Structure, and Clarity of Argumentation and Writing Style

The dissertation is written in a formal academic style and is well-organized, making it relatively easy to follow despite its breadth and length (over 300 pages). The structure is classic and effective: an introduction, literature review (theoretical framework), methodology, results, and conclusions/discussion, with clear subdivisions throughout. This conventional structure is executed with logical consistency, allowing the argument to develop step by step.

**Consistency of Argumentation:** From the introduction of the research problem through to the conclusions, the dissertation maintains a clear line of argument. Each chapter builds on the previous one in a cohesive manner. For example, Chapter 1 lays the contextual groundwork about educational quality and assurance; Chapter 2 builds the argument for specific relationships and gaps, leading to hypotheses; Chapter 3 describes how the research will address those hypotheses and questions; Chapter 4 presents the evidence; and Chapter 5 (or the latter part of Chapter 4, depending on how it's structured) ties it all together with interpretation. This progression is coherent – at no point does the reader have to wonder “why is this being discussed?” because each section's purpose is signposted. The dissertation even provides summaries of each chapter's content and role. For instance, near the end of the introduction or lit review, the candidate outlines how Chapter 3 will cover the mixed-methods approach and how Chapter 4 will present integrated findings. These previews and transitions help reinforce the logical flow.

The argumentation is also consistent in that claims made early on (like the importance of culture, or the expected link between quality and outcomes) are systematically examined with evidence later. There are no major logical leaps; hypotheses follow logically from theory, and conclusions follow logically from results. For instance, if H6 (ACB→AP) was only supported in Poland but not Norway, the discussion consistently refers back to the cultural context to explain this difference, which ties back to the original argument that culture matters. The text we have suggests the author consistently connects findings to the overarching thesis about QAS's role and cultural moderation, rather than presenting a laundry list of results without interpretation.

**Structure and Organization:** The dissertation's internal structure uses headings and subheadings effectively. Each research question and hypothesis is treated systematically. The qualitative results section is organized by case and by question, which mirrors the RQs – a very reader-friendly approach since the reader can see exactly how each question was answered. The quantitative results are structured by first evaluating measurement models, then structural models for each country, then tests of moderation, which is the logical order for SEM reporting. The use of numbered lists (e.g., listing the RQs 1–6, listing the hypotheses H1–H11) in the text helps in clarity and emphasis. There is also evidence of summary tables/figures: e.g., “Figure 1. Quantitative research model” and the mention of tables for reliability and hypothesis results. These visual elements likely aid understanding and break up the text.

The writing style is formal and academic, as appropriate for a dissertation. The author generally writes in the third person and avoids colloquial language. The tone is objective and evidential – statements are backed by citations (the excerpts show dense referencing, which is good for academic rigor). The language is for the most part clear and fluent. From the excerpts, grammar and vocabulary are at a high scholarly level. For example, sentences like *“The transparency and coherence of an EQMS are strongly dependent on its...”* or *“This study fills important gaps in the literature on higher education by including viewpoints that position students as active participants...”* show a professional academic tone. The candidate uses appropriate terminology (e.g. “moderating effects,” “partial least squares modeling,” “cross-case synthesis”) confidently.

One indicator of clarity is how the dissertation handles terminology and acronyms. Early on, the key constructs are given acronyms (UEQ, SL, ACB, AP) which are then used consistently. This not only saves space but also ensures the reader isn't confused by varying terms. The dissertation title itself is long but precisely reflects those constructs, and throughout the writing the author sticks to those terms (we don't see the author suddenly calling ACB something else; it's consistent). Additionally, the author spells out exactly what is meant by the “education quality assurance system” in context, ensuring readers from various backgrounds understand it's the institutional framework of processes and standards for quality in education (likely referencing definitions from QA literature). By the time the methodology and results are discussed, the reader has a clear idea of each construct and each piece of the model, which is a credit to the clarity of earlier chapters.

**Clarity in Presenting Data and Arguments:** The results chapter(s) appear to be clearly written as well. The candidate presents quantitative results with both statistical details and plain-

language interpretation. For example, instead of just stating “ $\beta = 0.725, p < 0.001$ ,” the text also says “Thus, hypothesis H2p is supported”, making it clear which hypothesis that was and the conclusion drawn. Similarly, qualitative findings are likely presented with illustrative quotes from interviewees (we saw some quotes like “It’s harder to change the schedule...”, which presumably are used to evidence points). The inclusion of verbatim quotes from participants is a good practice in qualitative reporting, and the candidate does include them (with clear attributions like “(Teacher 3)” or “(NCU Student 2)”). This enhances credibility and clarity, as readers can see the raw data behind the themes.

Furthermore, the writing remains focused on the research objectives. There do not seem to be irrelevant digressions. Even when providing broad context in Chapter 1 (like historical quality management concepts), the author ties it back to the current study’s context (discussing, for instance, how early quality concepts led to today’s educational quality approaches). The discipline in writing is evident: after laying out broad concepts, the author narrows down to specific hypotheses by end of Chapter 2, and doesn’t stray from addressing those in the results/discussion.

**Writing Quality:** From a language perspective, the dissertation seems well-written. There are minimal typos or grammar issues in the glimpses we have (the only minor thing seen is “supports for adaptive ...” where probably “supports adaptive” was meant, but that could be a transcription artifact). The sentence structures are generally complex yet comprehensible, befitting a PhD level. The candidate employs an academic vocabulary and proper citation style. The references to literature are seamlessly integrated (with author-year or numeric citation, etc., although in the excerpts we see the internal format, the final likely follows APA or similar). The tone remains formal; for instance, rather than saying “I think” or using first person, the author makes statements like “This study aims to explore...” or “It contends that...”, which keep the focus on the research, not the researcher – a stylistic choice often expected in dissertations (unless otherwise directed by a style guide).

**Overall Readability:** Despite the length and the potentially dense theoretical content, the dissertation is reader-friendly due to the structured approach. The use of bullet points or numbered points for research questions and hypotheses, the division into sub-sections with descriptive titles (e.g., “2.5 Academic Citizenship Behaviour and Academic Performance”, “4.5 Limitations and Future Research”) all help the reader navigate. Additionally, the summary of key findings (supported vs unsupported hypotheses) is concisely given in one place which is a nice way to remind the reader of the quantitative outcomes before diving into discussion.

**Consistency in Languages and Terminology:** It appears the dissertation might also have a Polish summary or section (given some Polish text in find results), but the main content is in English. The author manages the bilingual aspect (if any) by clearly separating it (possibly an abstract in Polish at the end). This doesn’t intrude on the English argumentation.

The dissertation’s structure and writing meet high academic standards. The argumentation is logically consistent and each part of the text serves the overall narrative. The clarity of writing ensures that even a complex mixed-methods, multi-variable, cross-cultural study can be understood by the reader without undue effort. Any minor stylistic improvements that could be suggested (e.g., slightly reducing length in some theoretical sections or adding a

diagram of the qualitative themes) do not detract from the overall impression that this is a clearly written and well-argued piece of scholarship. For a PhD candidate writing in a non-native language (assuming English is not the first language here), the proficiency is excellent. Importantly, the lightly positive stance of the writing – while objective – shines through in the way the author remains constructive and analytical throughout, which in turn makes the dissertation convincing and authoritative.

## 6. Presentation and Discussion of Findings

The dissertation's presentation of results and its discussion are thorough and reflective, effectively combining the quantitative and qualitative findings to address the research questions. This section of the review examines how the findings are presented, the extent of triangulation between methods, the handling of limitations, and the implications drawn for theory and practice.

**Presentation of Quantitative Findings:** The quantitative results are clearly presented, with appropriate use of tables and figures. The author first validates the measurement model for the survey constructs, ensuring that the indicators of UEQ, SL, ACB, AP (and cultural values) meet reliability and validity criteria – all of which bolsters confidence in subsequent structural results. For instance, the dissertation provides tables of internal consistency (Composite Reliability and Cronbach's alpha) and convergent validity (AVE), showing that all constructs are measured reliably (CR values mostly well above 0.8 as seen in an excerpt). Such detail indicates careful attention to methodological soundness before hypothesis testing.

The structural model results are then presented for each country (Poland and Norway). The author wisely separates these (H1–H7 tested perhaps in each subsample, and H8–H11 tested through moderation analysis or multi-group comparisons). The findings are summarized in the text in a very digestible format: the author enumerates which hypotheses were supported and where. For example, *“Supported Hypotheses: H1, H2, H3, and H4 were significant in both countries. H6 was significant only in Poland, while H5 was significant only in Norway. Cultural moderation was confirmed for H9 in Poland, H10 in Norway, and H11 in Poland. Unsupported Hypotheses: H7 and H8 showed no significance in either country.”*. This succinct summary is excellent for readers – it condenses a lot of complex analysis into a clear set of outcomes. It also directly answers many of the hypotheses, indicating the pattern of cross-cultural similarity and difference. The text around this summary likely provides the numeric results (e.g., “UEQ → SL:  $\beta = 0.72$ ,  $p < 0.001$ , supporting H2”). The clarity in linking back to hypothesis numbers ensures the reader can map findings to the earlier stated expectations.

The quantitative results presentation also addresses the magnitude and significance of effects, not just their existence. For instance, the coefficient for UEQ → SL is quite high ( $\beta \sim 0.725$  in Poland), indicating a substantial effect, which the author presumably discusses as an indication of the strong role of perceived quality in driving student loyalty – a point of practical importance for universities. Additionally, the moderation results are explained in text; for example, H9 being supported in Poland implies that Norway's low masculinity (high femininity) context might weaken the link between citizenship and performance, whereas Poland's higher masculinity strengthens it – the author likely interprets this in discussion.

**Presentation of Qualitative Findings:** The qualitative results are presented in a structured manner aligned with the research questions. The dissertation separates the findings by case (NCU and NTNU) for each question, and then provides a cross-case synthesis. For instance, under RQ1 (measurements in QAS), the author describes NCU's tools (e.g., course evaluation surveys via the USOS system, specific quality indicators like a teaching quality index), then describes NTNU's tools (likely similar surveys but also perhaps other metrics), and then compares. The presence of details like "*NCU implements multiple measurement tools: primarily course evaluations through the USOS system, complemented by student satisfaction surveys, graduate career tracking, etc.*" shows the richness of detail captured. For RQ2 (procedures for improvement), RQ3 (QAS contribution to quality improvement), etc., the author would similarly narrate what each university does. The comparative analysis is explicitly done: "*then, a cross-case synthesis is conducted to highlight the similarities and differences between the QAS of the two universities*". This is where key differences emerge, such as the earlier-cited hierarchical vs participatory approach in QA processes. Presenting the qualitative data with direct quotes enhances credibility – for example, a quote from an NCU teacher: "*When employees ignore the system, students will too*", or a student noting the difficulty of changing certain aspects via QA. These quotes humanize the findings and provide evidence for the author's interpretations (e.g., communication issues or limits of QA changes).

The author uses qualitative findings not only to describe but also to help explain the quantitative patterns. This is the essence of triangulation in a mixed-methods study. For instance, why was academic citizenship behaviour linked to performance only in Poland? The qualitative data might reveal that Polish students believe engaging in extra activities (like tutoring peers or active class participation) can improve their grades or are recognized by teachers, whereas Norwegian students might not see such a direct payoff due to more individualistic or egalitarian classroom norms. If any interviewees touched on this (perhaps a Norwegian student saying academic performance is seen as one's own responsibility regardless of extra activities), the author would bring that insight to bear. Similarly, the lack of power distance moderation could be explained by interviews showing that both universities have fairly student-friendly environments despite national culture differences, meaning even in Poland students felt they could voice opinions – thus PD didn't change the effect of quality on loyalty. Another example: the strong effect of quality on loyalty in both countries might be illustrated by students' comments praising good teaching or curriculum (which fosters their sense of belonging and willingness to continue). The dissertation indeed notes that both universities "*implemented comprehensive QAS frameworks following European Bologna Process requirements*", implying a baseline similarity, but also had differences in implementation approach. The discussion likely synthesizes these: both cases had robust QA, which may explain why in both cases students respond positively (hence H1–H4 working in both), yet the differences in approach align with differences in outcomes for performance-related hypotheses.

**Triangulation and Integration:** The integration of findings is explicitly done in Chapter 4 (titled as an integrated findings chapter). The author does not treat the qualitative and quantitative as separate silos; instead, they are woven together in the discussion. For each major finding or theme, the author cross-references the evidence from both strands. For example, the author might state: *Quantitatively, UEQ significantly increased SL in both*

*countries, which is consistent with numerous prior studies and was expected. Qualitative interviews reinforced this result: students at both NCU and NTNU frequently mentioned that when they perceive courses to be well-organized and teaching to be of high quality, they feel more satisfied and proud of their institution (NCU Student 1: "Good teachers make me want to recommend my university.") – such a hypothetical statement shows how triangulation can validate a finding. Likewise, for a divergent finding: Quantitatively, ACB predicted AP in Poland but not in Norway. Interestingly, Polish faculty interviews suggested that students who engage in extra activities often are noticed and encouraged, sometimes translating to better capstone project opportunities or informal learning gains, whereas in Norway teachers emphasized equality and did not differentiate students based on extra contributions (NTNU Teacher 2: "Whether a student is active or not, we grade purely on performance"). While I am conjecturing based on logical interpretation, the dissertation likely contains similar reasoning gleaned from actual quotes and observations. This integrated approach greatly strengthens the study's conclusions; it demonstrates that the candidate can "tell the story" of the research in a multifaceted way, moving beyond just numbers or isolated quotes to a coherent narrative.*

**Discussion of Findings:** In the discussion (or integrated results discussion), the author addresses each hypothesis and research question in light of the results, relating them back to theory and expectations. The candidate shows critical thinking in interpreting why some hypotheses did not hold. For example, H7 (ACB mediating UEQ→AP) was unsupported; the discussion acknowledges this and offers explanations – perhaps because one of the links (UEQ→ACB or ACB→AP) was weak in one context, the mediated effect didn't materialize. The author might reference that *to the best of current knowledge, this was the first test of that mediation, so the finding itself is informative*, and suggest that maybe direct effects of quality on performance overshadow indirect ones, etc. Similarly, H8 (PD moderation) being null is discussed: the dissertation indicates that despite large country-level PD differences, the individual-level measures and the student context did not show a moderating effect. The author thoughtfully notes that by measuring culture at the individual level, we see variation and some hypothesized effects might *"deviate from theoretical expectations"* because students within a country are not monolithic. This nuance is well-articulated and is actually a contribution to theory (reminding researchers that cultural effects can be complex).

**Limitations:** The dissertation has a dedicated section for limitations and future research (Chapter 4.5), which the author addresses candidly. Key limitations discussed include:

- *Sample composition:* Only business students from two European universities were surveyed. The author notes that while the cross-cultural sample adds breadth, it is still limited to similar institutional types and doesn't cover other disciplines or non-European contexts. They acknowledge that having more Norwegian respondents would improve balance and that future research should include more universities or countries (e.g. non-Bologna systems). This is an appropriate limitation that they not only acknowledge but turn into a suggestion for comparative research beyond Europe.
- *Self-report and common method bias:* The reliance on self-reported data from students is mentioned, with the risk of common method variance since predictors and outcomes came from the same survey instrument. The author likely mentions

any steps taken to mitigate this (perhaps assuring anonymity, mixing question order, etc.) and cites Podsakoff's recommendations. This transparency is good; many dissertations might not mention it at all.

- *Academic Performance measure*: The author explicitly notes that AP was captured as a single (probably self-reported GPA or perceived performance) variable, which is a simplification. They reflect that student success is multi-dimensional and influenced by personal factors (motivation, study habits) beyond institutional quality. This is a very insightful limitation to raise. In response, they suggest future research incorporate additional outcomes or mediators like self-efficacy, engagement, learning strategies to get a fuller picture. By doing so, the author is practically outlining next steps that could build on this work – a hallmark of a reflective discussion.
- *Use of Hofstede/national culture*: The author notes a limitation regarding cultural measures: while national indices provide a starting point, individuals vary, and thus cultural effects might not perfectly align with country averages. They actually handled this by using individual-level scales, but they still caution about interpreting results solely at country level. This is a nuanced discussion of limitations, showing the candidate's awareness of potential over-generalization. They also cite literature (Au & Cheung, 2004; Straub et al., 2002) on pitfalls of using national culture mean scores, which is excellent – it shows engagement with methodological literature on culture.
- *Qualitative sample limits*: The author acknowledges that the number of student interviewees was small (only 5 in total) and thus “constrains the diversity of perspectives”. They recommend including more students across faculties and levels in future qualitative research. This shows humility and understanding that their case insights, while valuable, are not exhaustive. Also, they might mention that faculty and admin perspectives were more represented, which could bias the qualitative narrative towards the institution's view rather than student view – again something future work can rectify by larger, more diverse interview samples.
- *Institutional scope*: Both cases are European public universities in developed countries. The author notes this and suggests expanding to different contexts (private universities, other regions, etc.). They even elaborate that including non-European contexts with different governance or economic conditions could reveal how QAS works under different regimes. This forward-looking limitation acknowledges that findings might not directly transfer to, say, an Asian or American context without further study.
- *Temporal factors*: Although not explicitly mentioned in excerpt, it might be implicit: the data was collected at one point in time (2024), and the author might mention that a longitudinal design would be useful to see how changes in QAS or student perceptions over time affect outcomes. The mention of COVID-19 in future directions – the digital transformation of education – hints that the author is aware how context is rapidly changing and that their study, while timely, captures a snapshot and that ongoing changes (like the rise of online learning and digital QA tools) will need investigation.

The open discussion of limitations, coupled with constructive suggestions for future research, indicates a mature understanding of the research process. It invites other scholars to build on this work and addresses possible critiques head-on, thereby strengthening the

dissertation's credibility. None of the limitations identified fundamentally undermine the validity of the findings; rather, they provide context and boundaries for interpretation.

**Implications:** The dissertation delineates both theoretical and practical implications of the findings:

- *Theoretical Implications:* The author discusses how the findings contribute to existing theories. For instance, the support of many hypotheses reinforces theories like Social Exchange and Expectancy-Disconfirmation in a higher education setting, extending their applicability. The partial support of cultural moderation adds nuance to cross-cultural OB theories, suggesting that not all theoretically expected cultural effects manifest uniformly in an academic context (e.g., the null PD effect could imply that student populations might be converging in some values, or that institutional subcultures moderate national culture impact). The dissertation likely notes that this research "*contributes to a more nuanced understanding of the complex interplay between UEQ, SL, ACB, and AP, while accounting for culture*". This is indeed a theoretical implication: it provides an integrated model that future research can test in other contexts, and it raises new questions (like the conditions under which OCB benefits performance in education).
- *Practical Implications:* For university management, the implications are quite direct. The study finds that improving perceived educational quality (teaching, curriculum, support services) can significantly enhance student loyalty across cultures, which in turn fosters positive behaviors. This implies universities should invest in quality teaching and student experience, as it builds a virtuous cycle of engagement. Moreover, the fact that student citizenship behavior is beneficial (especially in certain contexts) suggests that encouraging student involvement (through volunteering, mentorship programs, student organizations) could have academic payoffs. Importantly, the cross-cultural differences imply that policy and practice should be culturally informed. The author explicitly argues against one-size-fits-all QA policies, advocating for *adaptive, culturally appropriate quality assurance programs* that resonate with local attitudes while meeting global standards. This might mean, for example, that in a high power-distance context, formal mechanisms with clear top-down communication might be needed to assure students of improvements, whereas in low power-distance contexts, involving students in decision-making (like NTNU's reference groups) is more effective. The dissertation likely provides such insights. It might also advise that in cultures like Norway's (feminine, individualistic), universities might need to make extra effort to illustrate to students how their engagement can still be beneficial, or else consider other ways to boost performance, since voluntary behaviors might not naturally translate into grades there. Conversely, in cultures like Poland's, harnessing students' competitive spirit via recognition of their contributions could reinforce both ACB and performance.
- Another practical aspect: The study can guide international accreditations or partnerships – showing that while European standards provide a framework, the implementation can differ. Perhaps the author mentions that both case universities were pursuing AACSB business accreditation, and how internal QA efforts align with such external expectations. The implication for managers is to blend external best practices with internal culture.

Finally, the discussion wraps up by tying these implications back to the original research problem: it demonstrates that understanding the role of QAS is crucial and that this dissertation has shed light on it, but also that continuous improvement is needed (which is fitting, given the subject matter is quality improvement itself!).

In conclusion, the presentation and discussion of findings in this dissertation are a strong culmination of the research. The results are communicated clearly, the mixed-methods evidence is triangulated to provide a rich, validated understanding, limitations are openly discussed, and meaningful implications are drawn. The candidate shows an ability to critically analyze the findings and to situate them in both scholarly and practical contexts – exactly what is expected at the PhD level.

## 7. Critical and Polemical Aspects – Points for Challenge, Extension, or Improvement

While the dissertation by Mengyu Cao is a substantial and high-quality piece of research, it is fruitful to consider areas where the work could be critically examined or extended. These critical and polemical aspects do not detract from the overall value of the dissertation but rather highlight opportunities for further debate and development of the research ideas.

**Ambitious Scope and Complexity:** One critical observation is that the dissertation undertakes a very ambitious scope, attempting to link four major constructs and cultural factors using two methods in two countries. Managing this complexity is challenging. The author succeeds overall, but one could question whether too many hypotheses were advanced for one study. With 11 hypotheses (and numerous sub-findings), the narrative risks becoming diffuse. For instance, some mediation and moderation effects turned out non-significant; a skeptical reviewer might ask if all those hypotheses were necessary, or if the study could have focused on fewer core relationships with more depth. The counterargument (which the author would likely make) is that exploring these many paths was necessary to fully map the theoretical model and that even null results contribute knowledge. Nonetheless, future research might isolate segments of this model for deeper exploration (for example, conducting a dedicated study on the mediating role of loyalty in several contexts, or an in-depth qualitative study on how academic citizenship is rewarded in different cultures). The broad scope is a double-edged sword: it yields comprehensive insights but also exposes the research to many points of potential critique (each hypothesis introduces a possible point of failure). In this case, the author handled the breadth well, but the complexity could have been overwhelming – a testament to the author's ability, but something future scholars might streamline for replication studies.

**Measurement of Academic Performance:** The use of a single or simplified measure of academic performance (likely GPA or self-assessed performance) can be challenged. Academic performance is multifaceted – encompassing not just grades, but skills acquisition, personal development, etc. By focusing mainly on grades (if that was the case), the study might under-represent what “performance” means. Additionally, self-reported GPA can be biased or inaccurate. A polemical question is: *Does a high GPA truly reflect learning or educational success?* Some educators would argue that focusing on GPA is too narrow and might encourage a transactional view of education (contrary to the spirit of academic citizenship). The author is aware of this limitation and suggests including things like learning gain or engagement in future research. Still, a critic might say that without objective

performance data (like actual transcript records or a standardized test), the link between quality and performance is hard to conclusively establish due to potential self-report inflation by students who are loyal (i.e., loyal students might overrate both quality and their own performance). This raises the classic issue of common source bias – which the author did attempt to address conceptually, but statistically it remains a concern in any single-survey study. A more polemical standpoint could also argue that academic performance is influenced by so many external factors (prior preparation, socioeconomic status, etc.) that attributing differences to QAS influence (especially in cross-sectional data) might overstate the causal inference. The dissertation interprets associations in line with theory, but skeptics might call for experimental or longitudinal studies to truly claim a causal “quality improves performance” relationship.

**Cultural Interpretation Nuances:** While the study treats national culture via Hofstede dimensions, one could challenge whether national culture is the best or only lens for interpreting the differences observed. Poland and Norway differ not only in culture but in other aspects: educational traditions, economic contexts, and student demographics. For example, Norwegian higher education is free and has a welfare-state support system; Polish students might face different pressures (some might work part-time jobs, etc.). Some differences in loyalty or behavior might stem from these socioeconomic factors rather than deep cultural values. The dissertation acknowledges that institutional contexts outside Europe might yield different results, but even within these two cases, culture is one explanatory layer among many. A polemical viewpoint might argue that *organizational culture* (the specific ethos of NCU vs NTNU) could be just as influential as national culture. Perhaps NTNU’s leadership fosters open dialogue, whereas NCU’s leadership style is more top-down – this might be a result of individual leaders or institutional history rather than broad national culture. Hofstede’s dimensions are quite general and sometimes criticized for being oversimplifications. The author did well to measure individual orientations and not rely solely on country means. Yet, if one were to challenge the findings, one might ask: *How do we know the observed differences are due to cultural values per se and not other country-specific differences?* For instance, the finding that quality directly influences performance in Norway but not in Poland – could it be that Norwegian courses are smaller or have more continuous assessment (so quality teaching immediately helps grades), whereas Polish courses might rely on big final exams (where perhaps extra studying or tutoring – related to ACB – matters more)? These pedagogical or structural differences could produce the same outcome observed, without invoking culture. The dissertation would benefit from a discussion acknowledging such alternative explanations. Indeed, the author’s case descriptions hint at differences in QA approach (hierarchical vs participatory), which could be cultural or could be policy choices. In a polemical debate, one might press the candidate on why those differences exist – are they truly because Polish culture is higher PD (so students are not given a voice, hence a hierarchical QA), or because Poland’s university governance regulations mandate a certain procedure? The interplay of culture and policy is complex.

**Students as Customers vs. Citizens – Philosophical Debate:** The dissertation embraces a dual view of students as customers and as citizens, aiming to reconcile them. However, some academics are deeply critical of the notion of students as “customers,” arguing it undermines academic values. A critical discussion point is whether emphasizing student

loyalty (a concept from customer satisfaction literature) is appropriate in education. Could it encourage the idea that universities should satisfy students at the expense of academic rigor (a common polemic in higher ed literature)? The dissertation's results show that quality (presumably including academic rigor) actually enhances loyalty, which is reassuring, but detractors of the student-as-customer model might still worry about unintended consequences. Conversely, one could challenge the concept of "academic citizenship behavior": is every extra-role behavior by students necessarily positive? Could some be performative or strategic (e.g., a student being overly participative just to curry favor for grades, which might distort fair evaluation)? The dissertation, in framing ACB positively, doesn't address potential negatives (like could too much citizenship behavior distract from study time, or cause burnout?). These are perhaps beyond its scope, but they are areas for further inquiry or debate. The author's introduction of ACB is novel and mostly framed as beneficial, which aligns with OCB literature, yet a critic might question if all forms of student extra-role behavior are indeed beneficial to performance or to the institution (for example, outspoken "citizens" might also criticize the institution – is that still ACB or not? If a student protests for better quality, is that an ACB showing civic virtue, or disloyal behavior? The definitions matter). The dissertation likely had to delineate what counts as positive citizenship behavior (the OCB scale captures altruism, etc., which are by definition positive). Still, this could be a polemical angle in a defense: the nuance of what student citizenship means and whether faculty universally welcome it.

**Generality and Applicability of the Quality Assurance Findings:** The dissertation finds that both NCU and NTNU have comprehensive QAS in line with Bologna Process standards. One might challenge: if both have "mature QAS frameworks," what exactly is the *added value* of those systems as opposed to general good teaching practices or institutional culture? In other words, how do we know it is the formal QAS that shapes the relationships, rather than simply the underlying educational quality itself? The study title posits the "role of the QAS in shaping relationships among UEQ, ACB, etc.", but the QAS is not directly measured as a variable, it's more of a contextual factor explored qualitatively. One could argue the dissertation provides strong evidence on relationships between perceived quality, satisfaction, behavior, performance, and how they differ by context – but slightly less direct evidence on *how the design of the QAS per se* influences those perceptions and outcomes. The qualitative portion does address this (with RQ3 about QAS contribution to quality improvement, etc.); however, a skeptic might say: "We see that in both places with QAS, quality leads to loyalty and so on, but would that not be true even without a formal QAS, if you just had good teaching and responded to student feedback informally?" Essentially, isolating the effect of having a systematic QAS vs. not having one is not within the scope (since both cases had one). Thus, one could challenge the title's implication that the QAS's role is proven – the study shows how QAS functions and is perceived, but doesn't experimentally demonstrate an institution without QAS for comparison. The author would likely respond that the QAS is embedded in what constitutes "perceived quality" at these universities (since QAS drives improvements that students then perceive as quality), and indeed interviews might have indicated that many quality improvements (like curriculum changes, facility upgrades) happened due to QAS procedures, thereby indirectly affecting student perceptions. This is a subtle point but one that could be raised: the causal chain involving QAS design is inferred rather than directly tested. An extension of this work could be to compare an institution with a very rigorous QAS and one with a very minimal QAS to

see differences in student outcomes; or a longitudinal study of before-and-after implementing certain QAS initiatives.

**Data and Methodological Constraints:** Some critical remarks on methods were already mentioned (e.g., cross-sectional self-report data limiting causality inference). One might also critique the interview sample composition – for instance, only a few students were interviewed, which might bias the qualitative findings towards the perspectives of QA officers or faculty, who could have institutionalized viewpoints. Perhaps students may feel differently; indeed, the interviews that were done did reveal some student skepticism (communication problems, etc. ). A future improvement would be systematically including more student voices. Also, the disproportionate number of interviews at NTNU vs NCU (9 vs 6) might mean NTNU's perspective is richer; though cross-case synthesis was still done, one could question if the Polish case got as deep an exploration as the Norwegian one. The author's familiarity with NCU likely offset that, but it's a valid question.

Another angle: The PLS-SEM approach is fine, but a purist might ask if a covariance-based SEM (like multi-group analysis in AMOS or Mplus) would confirm the measurement invariance across cultures and test paths slightly differently. PLS is great for exploratory analysis but provides less conventional model fit statistics. If one were skeptical, they might question if the measurement model was truly equivalent in both countries (did the constructs mean the same to Polish and Norwegian students?). The author does not mention conducting measurement invariance testing explicitly – it's somewhat advanced but increasingly expected in cross-cultural surveys. Perhaps the small sample in Norway limited that. This isn't a major flaw, but a technical point that could be raised: without full measurement invariance, comparing path coefficients across groups can be tricky. However, given the author's meticulous approach, they may have considered it or at least ensured each construct was reliable in each group.

**Opportunities for Extension:** The dissertation itself suggests many extensions: exploring additional mediators/moderators (student motivation, etc.), expanding to other contexts, examining technology-enabled QA (post-COVID digital learning quality). These are sensible directions. One intriguing extension is the effect of digital transformation on QAS – as the author notes, the pandemic accelerated digital evaluation methods and analytics. A follow-up study could investigate how digital feedback tools (e.g., learning analytics dashboards, online course evaluations) are changing student engagement and how that interacts with culture (maybe students in different cultures use digital feedback differently). This shows the candidate is forward-thinking, though that topic is broad enough to be a separate thesis in itself.

In a polemical discussion, one could push further: is formal quality assurance becoming obsolete in the face of agile, student-driven feedback loops (like immediate feedback on social media or course forums)? The dissertation stands on the premise that QAS is important; a contrarian view might claim that too bureaucratic a QAS can be counterproductive, stifling innovation and genuine feedback. Some faculty criticize QA as "tick-box exercises." The dissertation did capture some criticisms (like communication issues, and constraints in changing certain things). It would be interesting to challenge the candidate on how to ensure QAS processes are meaningful and not just formalities. The

candidate's work actually supports that when QAS is responsive, it yields positive student attitudes. But maintaining that requires continuous effort; the dissertation might prompt questions about the sustainability of positive effects if, say, students become apathetic to surveys (one teacher quote noted poor response rates and linked it to faculty attitudes). So, in improvement, one might examine strategies to combat "survey fatigue" or to incentivize ACB (like recognition of student representatives).

In conclusion, the points above highlight that while the dissertation is strong, it opens up debate on measurement choices, interpretation of cultural effects, the generalizability of results, and the practical implementation of QAS. These are not so much flaws as they are rich avenues for scholarly dialogue. The author has demonstrated awareness of many of these issues (through limitations and discussion), which mitigates potential criticisms. Each challenge noted can be turned into a constructive question or future research proposal, which indicates the dissertation's success in advancing a conversation in the field. A hallmark of good research is that it raises new questions even as it answers its initial ones, and that is certainly true here.

After this comprehensive evaluation of the dissertation, I will pose a few substantive questions to the PhD candidate, and then provide my concluding recommendation.

#### Questions for the Candidate:

1. Research Design – Mixed-Methods Integration: You chose a mixed-methods case study design combining PLS-SEM analysis with qualitative interviews. Can you elaborate on how you balanced these two components during the research? For example, did the qualitative findings ever lead you to re-interpret or contextualize a surprising quantitative result (such as an unsupported hypothesis), and how did this integration influence the conclusions you drew? In retrospect, would you adjust the design in any way (e.g. sequence of data collection or emphasis) to strengthen the integration of the methods?
2. Cultural Interpretation of Findings: Your results showed notable cross-cultural differences, such as academic citizenship behaviour impacting performance in Poland but not in Norway, and the use of hierarchical vs. participatory mechanisms in the two QA systems. How do you interpret these differences in cultural terms? Specifically, to what extent do you attribute them to underlying national cultural values (Power Distance, Masculinity, Collectivism, etc.) versus other contextual factors (like institutional policies, economic conditions, or student support systems)? In other words, what makes you confident that culture is the key explanatory factor for those differences, and could you envision any scenario where an outlier institution might not follow the national culture pattern?
3. Generalizability and Transferability of Findings: Given your study focused on two European universities with mature quality assurance systems, how generalizable do you believe your findings are to other settings – for instance, universities in non-European countries, or those without well-established QA frameworks, or even to different types of students (e.g. part-time or online students)? What cautions would you advise if an institution outside your study's context wanted to apply your insights (for example, about fostering student loyalty or academic citizenship) to improve

academic performance? Are there specific conditions under which you suspect the relationships you found (like UEQ → SL → ACB) might behave differently?

### Conclusion and Recommendation

In summary, Mengyu Cao's dissertation is a highly commendable piece of scholarly work that tackles a complex interdisciplinary problem with rigor, depth, and cross-cultural insight. The research is clearly motivated by important gaps in the literature and carries significant implications for both theory and practice in higher education management. The candidate demonstrates strong capabilities in conceptualizing a theoretical model, executing a mixed-methods research design, and analyzing and interpreting data in a nuanced manner. The writing is organized and articulate, reflecting a solid command of academic discourse.

Notably, the dissertation contributes original knowledge by illuminating how educational quality assurance systems can influence student attitudes and behaviors across different cultural contexts. It confirms several expected relationships (underscoring the value of quality improvement efforts), while also uncovering culturally specific dynamics that add a new layer of understanding to the literature on service quality and organizational citizenship in education. The work shows a balanced perspective – it highlights positive outcomes of robust QAS implementations, yet also critically discusses limitations and areas for improvement, indicating the author's reflective and scholarly mindset.

Having carefully reviewed the dissertation in all the specified aspects, I find the work to be of high quality, meeting the standards required for a doctoral degree. The research problem is relevant and addressed convincingly, the methodology is sound and well-executed, the theoretical framing is comprehensive, the analysis is thorough, and the presentation is clear and professional. Any minor shortcomings or open questions are outweighed by the dissertation's contributions and the author's awareness of how to further advance the research.

In light of the thorough research, coherent analysis, and meaningful contributions demonstrated in this dissertation, I confidently recommend that Mengyu Cao's dissertation be accepted for the final public defence of the PhD degree. The work fulfills the high standards of academic rigor and originality required for the doctoral degree, and Mengyu Cao has shown the capability to conduct significant independent research in the field of Management and Quality Sciences. The dissertation merits approval, and I endorse the conferral of the PhD.

### Konkluzja i Rekomendacja

Podsumowując, dysertacja Mengyu Cao stanowi wysoce godne uznania dzieło naukowe, które podejmuje złożony problem interdyscyplinarny z należytą rzetelnością, głębią oraz międzykulturowym wglądem. Badania zostały wyraźnie zainspirowane istotnymi lukami w literaturze i niosą ze sobą znaczące implikacje zarówno dla teorii, jak i praktyki w obszarze zarządzania szkolnictwem wyższym. Doktorantka wykazuje silne kompetencje w zakresie konceptualizacji modelu teoretycznego, realizacji badań o mieszanej metodologii oraz analizy i interpretacji danych w sposób subtelny i pogłębiony. Tekst jest dobrze zorganizowany i klarowny, odzwierciedlając solidne opanowanie dyskursu akademickiego.

Co istotne, rozprawa wnosi oryginalny wkład do wiedzy poprzez ukazanie, w jaki sposób systemy zapewniania jakości kształcenia mogą wpływać na postawy i zachowania studentów w różnych kontekstach kulturowych. Potwierdza kilka oczekiwanych zależności (podkreślając wartość działań na rzecz doskonalenia jakości), a jednocześnie odkrywa kulturowo specyficzne mechanizmy, które wnoszą nową perspektywę do literatury dotyczącej jakości usług i organizacyjnego obywatelstwa w edukacji. Praca prezentuje zrównoważone podejście – wskazuje pozytywne rezultaty wdrażania solidnych systemów QAS, a zarazem krytycznie omawia ograniczenia i obszary wymagające poprawy, co świadczy o refleksyjnej i naukowej postawie autorki.

Po wnikliwym przeanalizowaniu rozprawy we wszystkich wymaganych aspektach stwierdzam, że praca reprezentuje wysoki poziom, spełniając standardy wymagane dla stopnia doktora. Problem badawczy jest istotny i został przekonująco opracowany, metodologia jest poprawna i starannie zastosowana, ramy teoretyczne są szerokie i spójne, analiza – dogłębna, a prezentacja – jasna i profesjonalna. Wszelkie drobne niedoskonałości lub otwarte pytania są niewielkie wobec znaczenia wkładu naukowego oraz świadomości autorki co do dalszych kierunków rozwoju badań.

W świetle oceny rzetelności badań, spójnej analizy oraz wartościowych kontrybucji przedstawionych w dysertacji, z pełnym przekonaniem rekomenduję dopuszczenie rozprawy Mengyu Cao do publicznej obrony stopnia doktora. Praca spełnia wysokie standardy rygoru akademickiego i oryginalności wymagane dla stopnia doktorskiego, a Mengyu Cao udowodniła zdolność do prowadzenia istotnych, samodzielnych badań w obszarze nauk o zarządzaniu i jakości. Rozprawa zasługuje na aprobate, a ja popieram nadanie stopnia doktora.

29.08.2025, Warszawa

