

Faculty Evaluation in Higher Education: A Theory-of-Action Case Study in Vietnam

Lan Anh Thi Nguyen

A thesis submitted in fulfilment of the requirements for the degree of Doctor of
Philosophy in Education, The University of Auckland, 2022

Abstract

The growth in neoliberal or market-driven higher education has challenged traditional approaches to evaluating faculty members¹. The involvement of multiple stakeholders (i.e., accreditation bodies, quality assurance officers, administrators, teaching faculty, and students) has led to different and sometimes conflicting needs in faculty evaluation. While extant literature generally suggests that faculty evaluation in contemporary higher education is strongly associated with accountability purposes, little is known about how key agents at the institutional level use evaluation for learning and improvement. Thus, this study attempts to identify the approaches to faculty evaluation that promote learning and improvement in higher education.

This study is a qualitative case study of Vietnamese higher education student evaluation of teaching (SET) and voting evaluation practices. It adopts the problem-based methodology (Robinson, 1993; Robinson & Lai, 2006) to investigate the theories of action of some key stakeholders of these faculty evaluation practices in Vietnam. The theories of action comprise the stakeholders' approaches to faculty evaluation, together with the constraints (that rule in or rule out specific approaches) and the consequences of the approaches. Participants included four quality assurance officers, 18 administrators², and 20 faculty members from seven public higher education institutions in Vietnam. The theories of action about faculty evaluation were constructed based on interviews and key evaluation documents.

Overall, the participants took three main approaches to faculty evaluation: (i) complying with the expected evaluation procedures and roles, (ii) taking a harmony-oriented and unilateral approach to problem solving, and (iii) disengaging from evaluation for learning and improvement. The participants generally fulfilled the evaluation policy requirements, but their approaches had a limited impact on learning and improvement at the institutional level. The constraint analysis suggests that the participants' approaches were not completely passive but were mainly oriented to managerial accountability demands (Røiseland et al., 2015). The participants' collectivist Confucian harmony-oriented norms hindered their approaches to problem solving and using evaluation for learning and improvement. However, there were a few participants who acted on their own values rather than being confined by political and cultural constraints. In spite of its scarcity, these participants' self-reliant approach suggests a potential influence of Buddhist principles of self-transformation, contextuality and reflexivity on improving practices (Chu & Vu, 2021; Thich Nhat Hanh, 1987).

¹ In this study, *faculty* refers to a group of teachers or departments in a university whereas *faculty members* emphasises individual teachers. Faculty and faculty members are also referred to as higher education teachers, instructors, lecturers or teaching staff in other studies.

² Vietnamese higher education administrators are faculty members who are appointed with additional managerial positions (see article 4, Circular 20/2020/TT-BGDĐT from <https://moet.gov.vn/tintuc/Pages/tin-tong-hop.aspx?ItemID=6852>). In this study, *administrators* refer to faculty members who play the managerial roles of faculty deans or department heads.

The study has two key implications. First, it suggests the improvement of faculty evaluation policymaking and implementation by revising several existing constraints and by adding some cultural and individual constraints to the current theories of action. The faculty evaluation problems could be solved by improving the quality of evaluation processes and data, reconceptualising SET and voting evaluation, and treating faculty members' underperformance as a collective problem. Second, the study suggests a framework to predict the likely success of future intervention on faculty evaluation. For instance, faculty evaluation for learning and improvement will be more feasible if policymakers and implementers prioritise transformational purposes and make joint efforts to foster dialogues and collaboration among individuals and groups. My study also highlights the need to understand faculty evaluation, and possibly other higher education practices in Vietnam, in religious contexts and from individual participants' spiritual values or philosophical backgrounds. The study implications are applicable to Vietnamese higher education and potentially to other educational settings with similar characteristics.

Acknowledgements

My PhD study is a wonderful journey that has profoundly transformed my academic and life experience. I would not have enjoyed the journey without the support and encouragement of my supervisors, my family, my beloved friends, and many others.

My heartfelt gratitude goes to my supervisors, Associate Professor Claire Sinnema and Associate Professor Mei Kuin Lai. Claire, your resourceful support and critical insight have helped me become an independent researcher. Mei, I love your constructive comments and structured guidance. Your constant encouragement has helped me overcome many emotional challenges during my PhD study. My deep gratitude also goes to Professor Viviane Robinson for her supervision during my provisional year. Viviane, your wisdom, clarity and your development of problem-based methodology have reinforced my belief in research that improves educational practices.

I am incredibly grateful to my beloved family, especially my parents, for always providing me with the best conditions to pursue my learning. To my husband and my soulmate, Hung Nguyen, I could never be who I am today without your selfless support. I am so proud when our sons call you “the best dad in the galaxy.” To my sons, Hieu Nguyen and Nam Nguyen, thank you for being so sweet. I am blessed to watch you grow your confidence and loving kindness every day.

My thesis journey would not have been more fruitful without my PhD buddies at the Faculty of Education and Social Work, University of Auckland (UoA). Thanks to Darren Hannah for giving critical feedback on my research proposal and early drafts of the participants’ theories of action. Many thanks to my UoA fortnight research-sharing group members, Hailey Chan, Harry Zhao, Ripi Kaur and many others, for broadening my knowledge through our engaging discussions on educational research. I very much appreciate my doctoral hub members, Junita Purwandari, Waseema Fikuree, Taqi Amini and Anny Bertoli, for sharing our enjoyable lunches and cultural talks. Thanks also to my UoA Vietnamese PhD candidates, Lan Nguyen, Hoa Pham, Quynh Anh Phan, Duy Trang Nguyen, Thu Trang Nguyen, and Tien Nguyen, whom I am really proud of.

Thanks to many adorable friends that I met face-to-face or virtually during my stay in New Zealand, my PhD journey has been more fulfilling. I sincerely thank friends from my children's virtue class, especially Pari, Mary, Alka, and Aashish, for teaching our lovely children to harmoniously live and unite regardless of their ages, races, and religions. My special thanks go to the whole AVSE Global – EduNet team, especially to our talented leader, Professor Hong Bui, for reinforcing my belief in the sustainable development of Vietnamese education.

My PhD journey would not have started without inspiration from Tam Duong, a long-time friend whom I truly value. Also, my thesis could not have been completed without insightful perspectives from the participating quality assurance officers, administrators, and faculty members to whom I am deeply indebted. My thesis is dedicated to these participants and to all teaching faculty members whom I consider the heart of any university. I firmly believe that an effective faculty evaluation system that promotes professional learning and improvement among faculty members would be the key to higher education quality enhancement. Below is my poem that summarises the story behind my thesis aspiration:

Why do I investigate faculty evaluation?

Why do I care about teaching faculty?

Because they are important to our Vietnamese.

It has been a crucial tradition of our beloved community to respect *thầy*³ and to appreciate learning opportunities since education is the best chance for our upward mobility⁴.

However, since Vietnam's transition to modernity, our teaching faculty seem to lose their professional autonomy. They are drawn to the culture of performativity, and they seem to pay less attention to their teaching quality.

Student evaluation of teaching has become a top priority. Over the benchmark, or you'll be rated as "underperformed employees." Another utmost faculty concern is research productivity, Getting enough research points unless you want to encounter adversity.

Oh, how important the faculty evaluation has shown to be: a catalyst to change the behaviours of our faculty. That's why I aspire to put it under scrutiny to investigate the faculty evaluation that improves the professional learning of Vietnamese teaching faculty.

(L. A. Nguyen, July 2021)

³ *Thầy* generally means teachers or teaching faculty members, and it also means father in Vietnamese.

⁴ Following a popular proverb in Vietnamese: "Muốn sang thì bắc cầu Kiều. Muốn con hay chữ phải yêu lấy thầy" (If you want a high-class house, build an elegant bridge. If you want well-educated children, respect the teacher).

Table of Contents

Abstract	ii
Acknowledgements	iv
List of Figures	ix
List of Tables	ix
List of Abbreviations	x
Chapter 1: Introduction	1
Global Context of Higher Education Faculty Evaluation.....	1
Accountability in Global Higher Education	1
Increased Managerial Focus on Faculty Evaluation	2
Negative Impact on Professional Focus of Faculty Evaluation.....	3
Vietnamese Context of Higher Education Faculty Evaluation	4
Vietnamese Cultural Context.....	4
Vietnamese Sociopolitical Context	6
Vietnamese Higher Education Context.....	7
Vietnamese Faculty Evaluation and Development Issues	9
My Positionality and Thesis Structure.....	9
My Positionality	10
Thesis Structure	11
Chapter 2: Literature Review	12
Faculty Evaluation Definitions	12
Faculty Evaluation Purposes	14
Summative Purposes	15
Formative Purposes	15
Transformational Purposes	16
Faculty Evaluation Criteria	18
Evaluation of Teaching.....	18
Evaluation of Research and Service.....	19
Faculty Evaluation Practices	19
SET: A Global Practice.....	20
Voting Evaluation: A Vietnamese Practice.....	24
Concluding Remarks	28
Chapter 3: Methodology	30
Research Approach	30
Research Questions.....	30
PBM	30

Qualitative Case Study Design	33
Participants, Cases and Context	34
Participants	35
Cases.....	38
Policy Context	40
Data Collection Methods	45
Interviews.....	45
Documents.....	48
Data Analysis Approach.....	50
Data Familiarisation.....	50
Individual ToA Construction.....	54
Group ToA Construction.....	59
Ethical Considerations.....	61
Addressing Dual Roles.....	61
Minimising Risks	62
Increasing Validity	63
Ensuring Transferability.....	66
Chapter 4: SET Theories of Action	67
QA Officers' SET ToA	67
QA Officers' Approaches to SET	67
Constraints on QA officers' Approaches	73
Consequences of QA Officers' Approaches	77
Administrators' SET ToA	79
Administrators' Approaches to SET	79
Constraints on Administrators' Approaches.....	85
Consequences of Administrators' Approaches	90
Faculty Members' SET ToA.....	91
Faculty Members' Approaches to SET	91
Constraints on Faculty Members' Approaches	96
Consequences of Faculty Members' Approaches.....	101
Concluding remarks	103
Chapter 5: Voting Evaluation Theories of Action	104
Administrators' Voting Evaluation ToA	104
Administrators' Approaches to Voting Evaluation	104
Constraints on Administrators' Approaches.....	110
Consequences of Administrators' Approaches	115
Faculty Members' Voting Evaluation ToA.....	117

Faculty Members' Approaches to Voting Evaluation.....	118
Constraints on Faculty Members' Approaches	125
Consequences of Faculty Members' Approaches.....	132
Concluding Remarks	134
Chapter 6: Discussion	136
The Predominance of Accountability-Oriented Approaches.....	136
Managerial Accountability: A Better Explanation of Compliance Than Institutional and Individual Passivity	138
Multiple Perspectives Towards Harmony-Oriented Approaches to Problem Solving	139
Exceptions to the Rule: Understanding Why Some Individuals Act on Their Own Values	141
Improvability of the Current Theories of Action	143
Can Faculty Evaluation be Improved?.....	143
How can Faculty Evaluation be Improved?.....	144
Chapter 7: Conclusion.....	148
Contributions of the Study	149
Methodological Contributions	149
Theoretical Contributions	150
Practical Contributions	152
Limitations of the Study	153
Suggestions for Future Research	153
References	155
Appendices	187
Appendix A. Interview Protocol.....	187
Appendix B. Sample of SET Report	190
Appendix C. Sample of Translation Excerpts	192
Appendix D. Individual ToA Example	194
Appendix E. Group ToA Example.....	198
Appendix F. Participant Information Sheet	200
Appendix G. Consent Form.....	203
Appendix H. Sample of SET Form.....	205

List of Figures

Figure 1 <i>Faculty Evaluation Purposes and Orientations</i>	14
Figure 2 <i>Single- and Double-Loop Learning</i>	32
Figure 3 <i>Member HEIs Structure and Governance</i>	36
Figure 4 <i>Categories of Faculty Evaluation Practices</i>	39
Figure 5 <i>Expected SET Procedures at Member HEIs' Level</i>	42
Figure 6 <i>Procedures and Roles of Key Stakeholders in Voting Evaluation</i>	44
Figure 7 <i>Data Analysis Approach</i>	50
Figure 8 <i>Identifying Categories of Consequence Component in NVivo 12</i>	58
Figure 9 <i>A Ladder of Inference</i>	64
Figure 10 <i>Summary of QA Officers' SET ToA</i>	68
Figure 11 <i>Summary of Administrators' SET ToA</i>	80
Figure 12 <i>Summary of Faculty Members' SET ToA</i>	92
Figure 13 <i>Summary of Administrators' Voting Evaluation ToA</i>	105
Figure 14 <i>Summary of Faculty Members' Voting Evaluation ToA</i>	118
Figure 15 <i>The Participants' Faculty Evaluation Orientations</i>	136

List of Tables

Table 1 <i>Defining Faculty Evaluation Related Key Terms</i>	12
Table 2 <i>Participating HEIs and Individuals</i>	37
Table 3 <i>Documents Used in the Study</i>	49
Table 4 <i>Key Analytic Steps</i>	55
Table 5 <i>Examples of Administrators' Pacifying Approaches</i>	82

List of Abbreviations

Abbreviations	Full Phrases
ASEAN	Association of Southeast Asian Nations
GoV	Government of Vietnam
HERA	Higher education reform agenda
HEI(s)	Higher education institution(s)
MoET	Ministry of Education and Training
PBM	Problem-based methodology
PIS	Participant information sheet
QA	Quality assurance
SET	Student evaluation of teaching
ToA	Theory of action
UAHPEC	University of Auckland Human Participants Ethics Committee
VND	Vietnamese Dong
USD	United States Dollar

Chapter 1: Introduction

This chapter introduces the global higher education context that has negatively influenced the professional focus of faculty evaluation. It further highlights the cultural and sociopolitical context of higher education in Vietnam where this study was undertaken. The chapter then presents my positionality, followed by the thesis structure.

Global Context of Higher Education Faculty Evaluation

Faculty evaluation is crucial in higher education as it directly influences the working quality of faculty members who play pivotal roles in promoting student learning and ensuring the quality of research and community engagement (Cherry et al., 2017; Hallinger, 2010; Posselt et al., 2020; Schreiner & Tobolowsky, 2018). An effective evaluation system must promote faculty learning and improvement to satisfy these roles. However, faculty evaluation in contemporary higher education worldwide has been negatively influenced by neoliberalism which is characterised by corporate logic and market-oriented ideologies (Field, 2015; Saunders et al., 2011; Sułkowski et al., 2020; Yokuş & Yelken, 2019). Neoliberalism is linked to various trends such as privatisation, new public management, entrepreneurialism, globalisation, internationalisation and massification of higher education (Roman et al., 2018; Tight, 2019). These trends have increased the demand for managerial accountability and performativity, which negatively impact the professional values of faculty evaluation in higher education (Buller, 2012; Kenny, 2018; Sułkowski et al., 2020).

Accountability in Global Higher Education

Accountability, which generally refers to reporting and justifying actions (Robinson & Timperley, 2000), means demonstrating the transparency and relevance of higher education in particular (Hazelkorn et al., 2018). There are two accountability principles: *professional accountability* which is grounded in internal values and professional competence through ongoing discussions and dialogues; and *managerial accountability* which is rooted in hierarchical structures and external control through audits, external incentives, and strong sanctions (Røiseland et al., 2015).

The culture of accountability in higher education emerged in the late 1960s in the United States due to social and political forces, including state budget cuts, demands for academics' pedagogical improvement, students' demands for more active roles in higher education governance, and public demands for more cost-effective higher education (Buller, 2012; Darwin, 2021; Jha, 2005). Since the 1980s, neoliberalism, a market-driven ideology, has become prominent in higher education in various countries, such as the United Kingdom,

the Netherlands, Australia and New Zealand (Jones et al., 2020; Kenny, 2018; Macfarlane, 2021; Mutch & Tatebe, 2017). Due to neoliberalism, accountability in higher education has been more managerially focused, shifting away from being “light-touch or bottom-up self-assessment and quality assurance tools” to “top-down regulatory processes of accreditation and performance-linked funding” at global, national and institutional levels (Hazelkorn et al., 2018, p. 8).

Managerial accountability, which is linked to the global ranking, national funding, and institutional performance-management systems, has changed the global higher education system. A tremendous change that “has washed over universities throughout the western world” (Grant, 2019, p. 12) is the performativity culture that has transformed the higher education landscape to be more externally oriented and outcome-driven (Hazelkorn et al., 2018; Hutaibat et al., 2021; Macfarlane, 2021; Posselt et al., 2020). For example, the government in New Zealand uses the performance-based research fund to grant funding to higher education institutions (HEIs) with more active research activities at the international and national levels (Mutch & Tatebe, 2017). The managerial accountability and performativity culture have also become more prevalent in various Asian countries, such as in China (Gu & Levin, 2021; Huang, 2020), Hongkong (W. V. Chan et al., 2020; Postiglione & Jung, 2017), Korea (Lee & Lee, 2013), and the Philippines (Sannadan & Lang-ay, 2021). At the national level, the emphasis on accountability and performativity has transformed HEIs in some nations, such as Singapore (Cheong, 2017), Taiwan (Shreeve, 2020), and China (Gu & Levin, 2021; Shi et al., 2018), into world-class universities. However, managerial accountability has transformed faculty evaluation in many global HEIs in ways that are not fruitful to professional learning and improvement among faculty members.

Increased Managerial Focus on Faculty Evaluation

The neoliberal emphasis on managerial accountability has led to increased managerial focus, manifested by a global culture of performativity in higher education (Cherry et al., 2017; Gonzales & Núñez, 2014; Field, 2015; Saunders et al., 2011; Sułkowski et al., 2020; Yokuş & Yelken, 2019). The performativity culture has resulted in three key characteristics of faculty evaluation in many global HEIs: competitiveness, homogeneity and commodification (Gonzales & Núñez, 2014; Tight, 2018). First, the ranking game reinforces competition between nations, HEIs and individuals based on publications, awards, or other research-oriented activities (Stack, 2021). There have been national trends for assessing and managing individual faculty performance, such as using contracts linked to objectives or using a research excellence framework in the United Kingdom, France, Australia and Germany (Kenny, 2017; Musselin, 2013). Individual competition is more intensified in many contexts where single authorship receives more credibility than multiple authorship (Ramirez, 2013).

Second, HEIs tend to measure faculty work in various settings homogeneously based on a single and relatively narrow definition of quality as publication and citation metrics (Gonzales & Núñez, 2014; Marginson, 2010; Stack, 2021). The standardised or “one-size-fits-all” performance measurement (Tight, 2018, p. 7) usually favours research outputs over teaching quality (Ramirez, 2013). For example, on the QS World University Rankings, one of the most renowned ranking systems, the two most dominant criteria: “academic reputation” and “citation per faculty,” account for 40% and 20%, respectively (Top Universities, 2021). In contrast, teaching quality only accounts for 20% of the total. Teaching is measured by faculty member/student ratios, which the system considers as “the most effective proxy metric for teaching quality” (Top Universities, 2021, n.p.). The prevalence of these ranking systems has led to many HEIs’ emphasising research rather than teaching. Research output is highly rewarded because it is more easily presented and quantified than teaching effectiveness (K. Chan, 2001). Consequently, the scholarship of teaching and learning has been overshadowed by the maximisation of research publications (Alvesson & Spicer, 2016).

Third, faculty work has become commodified, which means faculty performance is considered a commodity in a competitive market (Gonzales & Núñez, 2014). Faculty research and scholarship, usually measured by quantitative outputs, demonstrate HEIs’ effectiveness and efficiency and promote institutional branding (Blanco & Metcalfe, 2020). Therefore, various HEIs worldwide have used measurable performance to consider tenure and promotion (Blanco & Metcalfe, 2020; Cherry et al., 2017; Jones et al., 2020; Kenny, 2018). Many HEIs also link faculty evaluation with a pay-for-performance mechanism in which faculty members receive incentives and rewards based on their performance (Lu, 2014; Taitt, 2011; Williams & Preziosi, 2004). For example, the United Kingdom appraisal system has moved from collegial and developmental to more managerial means such as objective-based evaluation or 360-degree appraisal (Shelley, 1999). Many HEIs in Poland and Malta use a metric calculation of indexed research productivity for pay rises, promotion and funding (Sułkowski et al., 2020). In China, merit-based evaluation, which was officially mandated in the national education system in 2008, rewards faculty members based on their performance (Lu, 2014). Overall, outcome-based performance management or “cash-for-publication” (Macfarlane, 2021, p. 464) reinforces the tendency among HEI’s leaders and academics to optimise publication outputs for economic rather than professional benefits.

Negative Impact on Professional Focus of Faculty Evaluation

Existing literature has demonstrated several negative effects of predominant neoliberal ideas with an increased managerial focus on faculty evaluation in higher education. One negative effect of neoliberal overemphasis on managerial accountability is decreased professional values of teaching, learning, community engagement and public interest

(Alvesson & Spicer, 2016; K. Chan, 2001; Mutch & Tatebe, 2017). While many HEIs pay more attention to the faculty members' economic efficiency, they invest minimally in improving their teaching and learning practice (Giovannetti, 2015). The HEIs' evaluation policies, therefore, tend to be "more as a gesture than a serious attempt to provide a pathway to improvement," and lack the willingness to invest "economic and institutional capital" in faculty members – "one of [their] most important resources" (Giovannetti, 2015, pp. 195–196). Consequently, there has been a decrease in faculty members' intrinsic motivation for learning and innovation in teaching (Cadez et al., 2017; Kallio & Kallio, 2014; Ko, 2001). The competitiveness in HEIs, which results from typical government budget cuts and institutional scarcity of financial resources, causes faculty members to be sceptical, demotivated and intimidated by evaluation (Arreola, 2007; Field, 2015; Sułkowski et al., 2020; Yokuş & Yelken, 2019).

Furthermore, the dominant neoliberalism, characterised by a Western-centric global ranking and performance management, has placed faculty members from the Global South, the lower-income countries or less-recognised disciplines, on an "uneven playing field" (Stack, 2021, p. 26). The Anglo-American domination in the publication and ranking regimes, manifested by an "inside/outside binary," or the "North/South and West/ East hierarchies", has excluded "non-English language work" and "endogenous knowledge" (Marginson & Xu, 2021, pp. 4–5). The global ranking logic and imposed norms from the West have shifted and even deteriorated national, local, and institutional priorities in many less developed countries (Shreeve, 2020; Stack, 2021). In many Asian countries, the "world-class movement" has diminished local knowledge and cultural status, leaving faculty members at the periphery of academia (Lo, 2011, p. 212).

Vietnamese Context of Higher Education Faculty Evaluation

My study explored higher education faculty evaluation in Vietnam, a non-Western lower middle-income nation in the Asia-Pacific region (World Bank, 2021). Vietnam is a populous country with above 98 million people living over a land area of 310,070 km² (World Population Review, 2022). Vietnam represents a peripheral nation (Lo, 2011) as it has relatively slow progress in global knowledge-economy capacity measured by institutional rankings and citation metrics (Marginson, 2010, 2011). The distinctive cultural and sociopolitical contexts have influenced Vietnamese higher education policies and practices, including the evaluation and development of faculty members.

Vietnamese Cultural Context

Like many countries in the East Asia region, the Vietnamese higher education tradition is influenced by collectivist Confucianism and Buddhism (T. M. Le & Yu, 2019; H. T. Ngo, 2020; T. N. Nguyen, 2019; H. V. Vu, 2019). These ideologies have influenced Vietnamese

people's ways of viewing learning, teaching, and engaging in educational practices like evaluation.

Collectivist Confucianism. *Collectivism* is a set of beliefs that prioritise groups' interests and power over individuals (Hofstede et al., 2010). Collectivism is one of the Vietnamese people's most significant shared identities (N. T. Nguyen, 2019). Collectivist values are corroborated by Confucian ideas of promoting hierarchy by accepting socially-expected responsibilities and respecting unequal status relationships among community members (Hofstede et al., 2010; H. T. Ngo, 2020; N. T. Nguyen, 2019). Confucian ideology has reinforced Vietnamese people's values of education and harmonious relationships (H. T. Ngo, 2020; T. L. Ngo, 2020). The Confucian heritage in Vietnam may serve as both an enabler and a hindrance to student learning. Confucian teaching, which promotes relationships and hierarchies, may encourage students' holistic or pluralist ways of viewing things (T. L. Ngo, 2020). In contrast, Confucian support of teachers' centeredness and authority may cause faculty members to be unreflective about their teaching approach, limiting the effectiveness of their teaching evaluation (K. T. Nguyen et al., 2006).

Buddhism. Despite commonly having nonreligious identification, Vietnamese people traditionally follow many teachings, practices or rituals, including Buddhism (Hoskins & Ninh, 2017). While Confucian ideologies promote political purposes of educating Vietnamese people to respect hierarchy and fulfil moral responsibility (H. T. Ngo, 2020), Buddhism enriches the Vietnamese people's spiritual lives with individual values of self-liberation, compassion, and nonattachment (Vuong et al., 2018). Buddhism can be viewed either as a religion, a spiritual set of values or philosophical teachings (Hoskins & Ninh, 2017). The world-renowned Vietnamese Zen master Thich Nhat Hanh⁵ (1987) emphasised that Buddha (i.e., Siddhartha Gautama) was not a God but a compassionate teacher. Buddhism encourages self-enlightenment and discovering the truth or knowledge through one's open-mindedness based on one's empirical experience. Underpinned by the principles of flexibility and freedom, recent Buddhism movements, such as Engaged Buddhism, have significantly promoted the ethical and egoless approaches to leadership and management in Vietnamese organisations (Chu & Vu, 2021). Buddhist managerial practices, which are characterised by context-specific, nonextreme and sustainable approaches, help leaders respond to the problems of morality and social-trust issues (Vu & Tran, 2021). While Confucianism has been related to "backwardness and superstition," Buddhism, especially the engaged Buddhist principles, has become more impactful in contemporary Vietnamese life (Chu & Vu, 2021, pp. 7–9). The Buddhist emphasis on reflexivity and contextuality, together with its sustainable approaches

⁵ Zen Master Thich Nhat Hanh (1926-2022) is lovingly referred to as *Thầy* (teacher in Vietnamese). He is also known as "the father of mindfulness" for his tremendous impact on practices of Buddhism in Western countries, from <https://time.com/5511729/monk-mindfulness-art-of-dying/>.

to management (Chu & Vu, 2021; Thich Nhat Hanh, 1987), arguably has a positive influence on Vietnamese approaches to individual faculty members' learning and improvement in various practices, including faculty evaluation in higher education.

Vietnamese Sociopolitical Context

Vietnam is one of a few existing communist countries, along with China, Cuba, and Laos (World Population Review, 2022). The single-party Vietnamese communist government, which has held power since 1975, rules the country with high interference in all facets of life, including education (London, 2010).

Since the nation's economic reform, "Doi Moi" (Renovation) in 1986, Vietnam's higher education governance has shifted remarkably from a centrally planned economy to a market-based economy (K. A. Le et al., 2019). Doi Moi has transformed Vietnam from one of the world's poorest to a middle-income nation, dropping the poverty rate from 70% to below 6% (World Bank, 2021). Along with this substantial economic transition, Vietnam has transformed its higher education system from Soviet-style socialist models to market-oriented ones (K. A. Le et al., 2019; London, 2010; L. H. Phan & Doan, 2020). Vietnam's dualism in governance, which means the government's adoption of two seemingly conflicting ideologies (market orientation and communism), has made the Vietnamese higher education system a distinctive case within the global higher education context (L. H. Phan & Doan, 2020).

On the one hand, the higher education system in Vietnam sustains the communist Soviet principles commonly seen in China, Russia, or other post-Soviet nations (L. H. Phan & Doan, 2020). At the institutional level, most important decision making, such as development plans or managerial posts, are required to consult the institutional communist party committee (T. L. Pham, 2012). Historically, Vietnam has officially unified since 1975, after 3 decades of war (T. H. Nguyen, 2016). In 1975, the Vietnamese communist government applied the Soviet-modelled higher education system across the whole country. The period 1975–1985 was marked by public mono-disciplinary HEIs with a mission to produce the "new socialist all-round developed citizens" to construct a socialist nation (T. L. Ngo, 2020, p. 127). Before 1986, the primary function of Vietnamese HEIs was teaching or training human resources, not research (H. T. Nguyen, 2020). Thus, faculty members' research evaluation has become a relatively new practice among Vietnamese HEIs.

On the other hand, the Vietnamese higher education system has been drastically transformed following "a strong Western-oriented spirit, energy and outlook" (L. H. Phan & Doan, 2020, p. 4). Various international and regional organisations, such as the World Bank and the Association of Southeast Asian Nations (ASEAN), have promoted neoliberal ideas of globalisation, internationalisation, and modernisation of various higher education policies and practices (H. T. Ngo, 2020; A. N. Phan, 2021). These neoliberal trends are marked by the

government's Resolution 4 in 1993 on restructuring and expanding higher education and Resolution 14 (also known as the Higher Education Reform Agenda [HERA]) (Government of Vietnam [GoV], 2005) on modernising higher education (Hayden & Chinh, 2020; T. L. Pham, 2012). The massification is marked by the government's approval of various types of HEIs. In contrast, the internationalisation of higher education is characterised by international (primarily Western) influence on Vietnamese higher education (L. T. Tran & Marginson, 2018a). The student evaluation of teaching (SET) is one of many practices resulting from the internationalisation of Vietnamese HEIs.

Vietnamese Higher Education Context

Vietnamese higher education is generally less developed than other counterparts in the ASEAN region, such as Singapore, Thailand, and Malaysia (C. H. Nguyen & Shah, 2019). Since the establishment of the first university in 1075, the Vietnamese higher education system has undergone various changes in its model, following French colonisation (1858–1942 and 1945–1954) and American imperialism (1954–1975), and the leadership of the Communist Party of Vietnam after 1975 (Agent, 2007; T. L. Ngo, 2020). Since Doi Moi, the Vietnamese government has demonstrated high aspirations for higher education to be on a par with world-class universities, aiming that by 2020 some of their HEIs would be ranked among the top 200.⁶ Specifically, HERA (GoV, 2005), one key Vietnamese policy document, indicates a national objective to radically reform the higher education system to satisfy the “national requirements for industrialisation, modernisation and international integration” (p. 1). Vietnamese higher education has witnessed size and scope achievements over 30 years since Doi Moi; however, there are persistent concerns over Vietnamese higher education quality.

Vietnamese Higher Education Growth in Size and Scope. The Western internationalisation and marketisation of Vietnamese higher education have expanded the number of HEIs and the scope of QA activities within the Vietnamese higher education system. First, Vietnamese higher education has experienced tremendous growth in size and diversity for the past 35 years. By 2019, Vietnam had 237 HEIs, of which 172 were public and 65 were private institutions, with over 73,000 full-time faculty members accommodating over 1.5 million students (Ministry of Education and Training [MoET], 2020). There has been an improvement in the number of academic staff having PhD degrees from 10% to over 22% during 2014–2017, and research productivity doubled during 2012–2017 (K. A. Le et al., 2019). Second, the Vietnamese MoET has developed a nationwide quality-assurance system that oversees a range of QA activities (C. H. Nguyen & Shah, 2019). The QA system has many signs of growth

⁶ Following Decision 121/2007/QĐ-TTg dated 27/7/2007 by the GoV, from <https://chinhphu.vn/default.aspx?pageid=27160&docid=33899>

in scope, from establishing the first accreditation unit in 2002 to developing legal accreditation documents and a framework with 700 accreditors by 2017 (N. T. Tran & Vu, 2019). There have been 217 out of 235 universities completing self-evaluation reports, with 124 HEIs being nationally accredited and six HEIs being internationally accredited by 2019 (MoET, 2019; N. T. Tran & Vu, 2019). Various practices have been introduced at the institutional levels, ranging from accreditation of training programmes to surveying stakeholders' opinions, such as SET (N. T. Tran & Vu, 2019). These QA instruments are expected to display HEIs' accountability towards the governing bodies and the public and be used to improve teaching and curricula (T. L. Pham, 2012).

Concerns Over Vietnamese Higher Education Quality. Despite growth in size and scope, Vietnamese QA progress has been relatively slow, and its higher education quality is still questionable (C. H. Nguyen & Shah, 2019). Despite high aspirations, Vietnamese HEIs have not climbed up the world ranking as expected by the state leaders. It was not until June 2021 that two of the Vietnamese HEIs were listed among the top 1000 according to the QS World University Rankings (The World and Vietnamese Report, 2021). After more than 15 years since the HERA (GoV, 2015) aimed to reform the Vietnamese higher education system comprehensively, its QA system is still at the pilot stage, lagging behind neighbouring countries such as Singapore, Indonesia, and Malaysia (H. T. Pham & Nguyen, 2019).

The development of the QA system is also inadequate to solve the existing issues of low-quality higher education in Vietnam. Also, the massification and internationalisation of higher education without adequate attention, governance, and resources have led to considerable concerns about the quality of Vietnamese HEIs (L. T. Tran & Marginson, 2018b). Many aspects of the national curriculum framework that Vietnamese HEIs must follow are obsolete in the current context (T. L. Pham, 2012). The training programmes in most HEIs in Vietnam are claimed to provide students with "pure disciplinary theories and political indoctrination" (L. T. Tran & Marginson, 2014, p. 105), hindering Vietnamese students' "independence, creativity, and problem-solving capacity" (Luu, 2006, p. 4). Consequently, Vietnamese graduates fail to cope in a rapidly changing world due to a lack of proper attitude and practical skills necessary for employment (L. T. Tran & Marginson, 2014). Also, there have been some critical issues of decreased educational values and social inequality due to the commercialisation of higher education, where Vietnamese HEIs pay more attention to profit-making than sustaining and improving teaching and research quality (Q. T. Do 2019; Hayden & Le-Nguyen, 2020).

Vietnamese HEIs' potential for higher education quality improvement is hindered by the dualism in HEI governance (L. T. Tran & Marginson., 2014). The Vietnamese government's combination of two contrasting ideologies, namely the market-led versus state-centralised vision, has created tension, confusion, and slow progress within HEIs (H. Tran,

2009). Although following Western university development models, the Vietnamese government retains a top-down approach to higher education governance, in which “government policy and regulation are not always insightful or internally consistent” (Tran & Marginson, 2018b, p. 254). This dualism has marked the Vietnamese higher education reform as an “experimental” process in which the success and failure of the previous policy have become lessons informing the next stage of policymaking (H. Tran, 2009, p. 104).

Vietnamese Faculty Evaluation and Development Issues

Although the Vietnamese government aims to invest in and develop a mechanism for attracting, employing and properly remunerating high-quality faculty members (GoV, 2005; Law on Higher Education 2003), their stated intention has not been realised. Consequently, there have been several issues with the quantity and quality of Vietnamese faculty members. The number of faculty must increase four times to accommodate massive student enrolments and a least a quarter of faculty members needs to acquire a doctoral degree (Do et al., 2017). The Resolution of the 8th Central Committee (the 11th term) indicates that the teaching staff in Vietnamese HEIs have not met the requirements for innovative teaching and research. Numerous Vietnamese staff are weak in research with low publication productivity, and some even violate professional ethics (T. X. Duong, 2016). While Vietnamese foreign-trained faculty members are more active in research, many of them have left public HEIs or moved to work overseas, which has caused the problem of “brain drain” in Vietnamese higher education (L. T. Tran et al., 2020, p. 297).

A lack of a practical faculty evaluation and development scheme is a core reason for Vietnamese faculty members' low quantity and quality (K. D. Nguyen, 2000; H. T. Pham & Nguyen, 2020). The current faculty evaluation in Vietnamese HEIs is characterised by unpersuasive evaluation criteria and ritual processes, which has resulted in a limited impact on motivating faculty members' continuous teaching and learning improvement (K. T. Nguyen et al., 2006; T. H. Nguyen, 2016; H. T. Pham, 2014). As faculty members are the key drivers for improving the HEIs' quality of teaching, research, and community engagement, it is imperative to investigate approaches to evaluation that promote Vietnamese faculty members' learning and improvement. More effective approaches to faculty evaluation would arguably contribute to a better student learning experience and serve the broader public interest.

My Positionality and Thesis Structure

My research topic was informed by the lack of professional focus and the limited impact of faculty evaluation in the global and Vietnamese higher education contexts. The research choice was also inspired by my positionality related to my personal values and previous experience with faculty evaluation.

My Positionality

This thesis grew out of my positionality as a continuous learner and a previous faculty member in a Vietnamese HEI where I experienced various faculty evaluation methods, such as SET, self-evaluation, or annual voting evaluation.

I value the learning opportunity in every professional practice that I have engaged in, including faculty evaluation. My learning focus is rooted in our Vietnamese tradition of value for education and respecting the teaching faculty who guide, inspire, and promote student learning. Following the Buddhist philosophy, I tend to view things “with the eyes of interbeing” (Thich Nhat Hanh, 1991, p. 98), which means to look deeply to see the ways people and their practices are interrelated and mutually dependent (Asher, 2003). Thus, I view teachers as continuous learners, and I also promote a sustainable leadership approach (Chu & Vu, 2021; Thich Nhat Hanh, 1987) to link faculty evaluation and faculty development. The Buddhist knowledge implies that mental need is equally important as physical need in determining if a person can work well or not (Thich Nhat Hanh, 2014). Hence, I paid careful attention to faculty members' professional needs and values, which are important to investigate, take care of, and nurture. However, I found many annual evaluation practices not very useful for my learning and improvement during my 8-year working in the HEI. Although I generally valued student feedback about their learning, I found the formal SET confusing and unhelpful. For example, I once got very negative SET rating scores without knowing what I was wrong and how I could improve my teaching. I also regularly attended whole faculty evaluation meetings where we appraised, proposed, and voted each other for communist-modelled titles such as “advanced fighter” or “Ho Chi Minh prize.” Although I appreciated the institutional effort to reward us with emulation titles and small material incentives, I did not find many learning opportunities from these collective evaluation meetings. Hence, I aspired to investigate the faculty evaluation practices that would be fruitful for faculty members' learning and improvement in Vietnamese public HEIs.

I strongly believe that my study's focus on faculty evaluation for learning and improvement is imperative to counter the negative effects of increased managerial focus on faculty evaluation in higher education (Beerkens, 2018; Cherry et al., 2017; Posselt et al., 2020; Tight, 2018). By focusing on a case study of faculty evaluation in some Vietnamese public HEIs, my study potentially sheds light on how institutional faculty evaluation in a lower middle-income country has been influenced by the national and global agenda. My study findings are expected to provide insights into improving higher education faculty evaluation in Vietnam where faculty members desperately long for an evaluation system that supports their professional learning and development. The lessons learnt from the cases of the SET and voting evaluation in the Vietnamese public HEIs are likely to have applicability to other similar contexts.

Thesis Structure

This thesis has seven chapters. Chapter 1 introduces the global and Vietnamese context of higher education that influenced my research focus on the professional values of faculty evaluation. The chapter then presents my positionality, followed by the thesis structure.

Chapter 2 defines the terms, criteria and purposes of faculty evaluation which are critical for understanding the evaluation implementation. The chapter then reviews the empirical literature on the SET and voting evaluation, especially the effectiveness and problems of these practices in the global and Vietnamese higher education contexts.

Chapter 3 introduces the research approach, namely the research questions, its methodology and qualitative case study design. Before presenting the data collection and analysis procedures, the chapter introduces the participants, selected cases, and policy contexts. The chapter then reflects some ethical considerations that address my dual roles, risks, validity and transferability of the research.

Chapters 4 and 5 present key research findings based on the analysis of the SET and voting evaluation policy documents and the participants' interview accounts. Chapter 4 presents the SET theories of action of the participating QA officers, administrators, and faculty members. Chapter 5 presents the voting evaluation theories of action of the participating administrators and faculty members.

Chapter 6 discusses key actions and constraints of the SET and voting evaluation compared to broader contexts. Then it discusses the improvability of the current theories of action concerning faculty evaluation.

Chapter 7 concludes the thesis by summarising its main contributions, limitations, and suggestions for future research.

Chapter 2: Literature Review

This chapter begins by clarifying the definition of faculty evaluation and its criteria, purposes and practices. It then elaborates on the key definitions summarised in Table 1. The chapter ends by highlighting the effectiveness and problems of two faculty evaluation practices (i.e., the SET and voting evaluation) that were selected as the cases of this study.

Table 1

Defining Faculty Evaluation Related Key Terms

Key Terms	Definition
Faculty evaluation	<ul style="list-style-type: none"> • a socially constructed process of collecting, interpreting, and using data about faculty performance (Lonsdale, 1999) • includes but is not limited to performance appraisal, performance management or quality assurance
Faculty evaluation criteria	the principles, standards, or indicators used to judge faculty performance.
Faculty evaluation purposes	<p>Key reasons for carrying out faculty evaluation; each of the purposes can either be:</p> <ul style="list-style-type: none"> • espoused or • evident in practice
Faculty evaluation practice	<p>An educational practice, which comes from the literature on theories of action, is defined as</p> <ul style="list-style-type: none"> • the “delivery of educational services,” together with • the “beliefs about what is important and about how what is important can be realised in particular circumstances,” and the “formulation of policies intended to influence those services” (Robinson, 1993, p. 5). <p>A faculty evaluation practice involves both</p> <ul style="list-style-type: none"> • the faculty evaluation implementing agents’ actions, and • the faculty evaluation policies and set of beliefs that influence these actions.

Note. The definition of faculty evaluation is adapted from “Performance Appraisal, Performance Management and Quality in Higher Education: Contradictions, Issues and Guiding Principles for the Future” by A. Lonsdale, 1998, *Australian Journal of Education*, 42(3), p. 304. The definition of educational practice is adapted from *Problem-Based Methodology: Research for the Improvement of Practice* (p. 5), by V. M. Robinson, Pergamon Press. Copyright 1993 by Pergamon Press Ltd.

Faculty Evaluation Definitions

Faculty evaluation is a social construct that is open to various interpretations. *Evaluation* linguistically means “worth” or “value,” but it also refers to a process, method, or

the purpose of determining the worth or the value of something (Mark et al., 2006, p. 7). *Faculty evaluation* (*đánh giá giảng viên* in Vietnamese) is defined as a process of collecting, interpreting, and using data about faculty performance to inform individual or institutional decision making (Lonsdale, 1998). Although *đánh giá* (evaluation) generally implies a value judgement or a conclusive opinion (Hoang, 2003), this study does not limit its focus on judgemental purposes but on the broader use of faculty evaluation. My chosen topic embeds an implicit concern over the use of evaluation for teaching improvement as *giảng viên* (teaching faculty members) focuses explicitly on those who teach (Hoang, 2003). Notably, faculty evaluation is not a one-off instance but a “social practice” or an “evolving process” (Saunders et al., 2011, p. 3), depending on the involvement of multiple agents. Thus, I view faculty evaluation practice as being constantly moving and socially constructed by different agents or stakeholders.

Faculty evaluation is an umbrella term that refers to various processes of performance appraisal, performance review, performance management, or QA of faculty members’ work (Buller, 2012; C. E. Cardno & Piggott-Irvine, 1997; Lonsdale, 1998; Mills & Hyle, 1999). QA commonly means ensuring that the quality of programmes, courses, or faculty members’ working quality is up to the predetermined standards (Tight, 2018). QA has recently been linked more to audit or accreditation systems – government instruments to measure and accelerate institutional performance and accountability (Jungblut et al., 2018; C. H. Nguyen & Shah, 2019). Many authors would see that QA is separate from quality enhancement because the former practice concerns diagnosing educational problems rather than improving practices. However, several authors (Elassy, 2015; Jungblut et al., 2018) suggest combining these quality concepts into a single continuum towards continuous quality improvement. This study attempts to link the concepts of QA and quality enhancement by investigating the use of QA, as in the case of SET, for faculty learning and improvement.

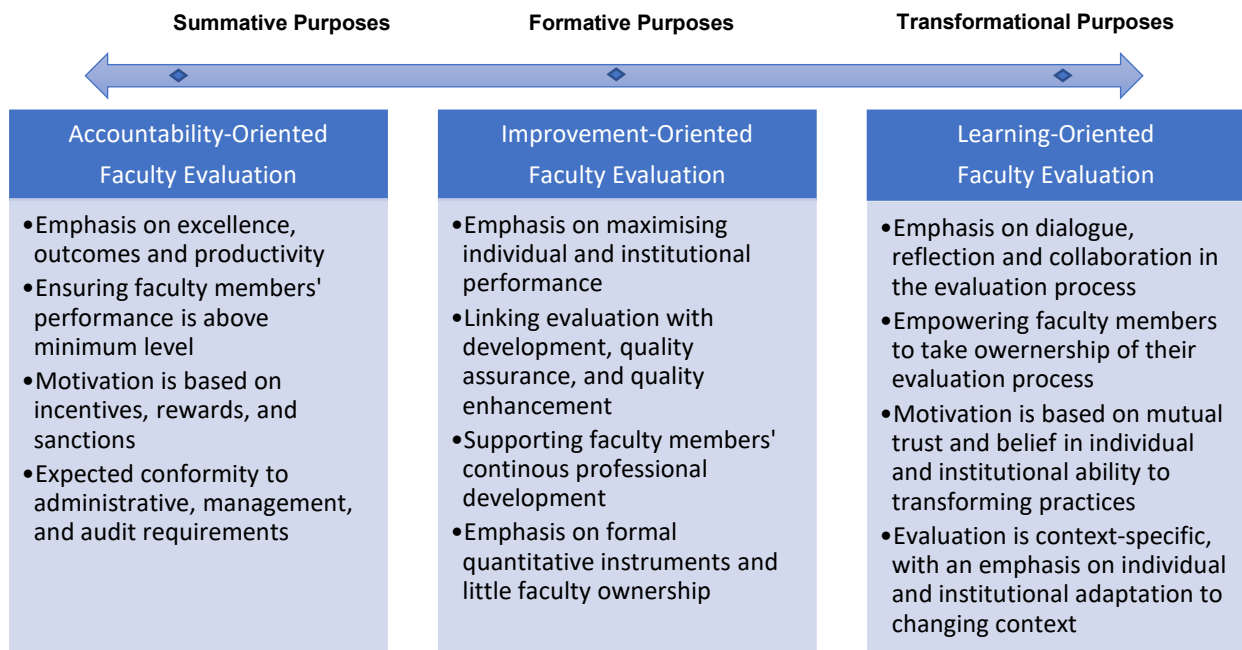
Performance appraisal of faculty members generally involves measuring faculty performance for managerial purposes such as granting rewards, ensuring accountability, or promoting organisational efficiency (Field, 2015; Khtere, 2020). Performance appraisal is a dynamic process of evaluating individuals or teams to determine the extent to which each contributes to an organisational goal at a specific time(s) during a year (Mampane, 2020). Performance appraisal is a central practice of performance management that has become more prevalent in contemporary higher education (Mampane, 2020). The notion of *performance* has recently focused more on “what academics produce through their teaching, and issues of academic outputs and outcomes” rather than on student learning and colleagues’ cooperation and shared responsibility (Hazelkorn et al., 2018, pp. 6-7).

Faculty Evaluation Purposes

Faculty evaluation success depends on the identification and prioritisation of major evaluation purposes that specify various aspects, such as collecting data and disseminating evaluation results (Seldin, 2006). Existing literature reveals commonly binary purposes of faculty evaluation are “personnel management vs self-improvement,” “hard vs soft,” “judgmental vs developmental,” “managerialist vs developmental,” and “institutional vs individual” (Mills & Hyle, 1999, p. 352). Instead of the binary view, I would suggest three evaluation purposes: summative, formative, and transformational purposes which are placed on a single continuum and are linked to three orientations of faculty evaluation: accountability, improvement (Lonsdale, 1998) and learning (Dahler-Larsen, 2009) (see Figure 1). While summative and formative evaluations are usually accountability- and improvement-oriented, transformational evaluations are learning-oriented. However, depending on their actual evaluation orientations, the evaluation purposes are either espoused or evident in practice. An evaluation practice can be espoused to serve either formative, summative (Scriven, 1991), or transformational purposes (Patton, 1996), but evidence in practice would determine if the evaluation is for accountability, improvement, or learning.

Figure 1

Faculty Evaluation Purposes and Orientations



Note. Accountability-oriented and improvement-oriented evaluation approaches are adapted from “Performance Appraisal, Performance Management and Quality in Higher Education: Contradictions, Issues and Guiding Principles for the Future” by A. Lonsdale, 1998, *Australian Journal of Education*, 42(3), 314-315. Learning-oriented evaluation is adapted from “Learning-Oriented Educational Evaluation in Contemporary Society” by P. Dahler-Larsen, in K. E. Ryan & J. B. Cousins (Eds.), *The SAGE international handbook of educational evaluation*, 2009, SAGE Publications.

Summative Purposes

Summative evaluation mainly involves judgmental determination or classification of faculty performance for personnel decision-making or for complying with managerial and accreditation requirements (Buller, 2012; Cashin, 2003; Shao et al., 2007). As Scriven (1991) stated, evaluation benefits are half determining the best candidates and half identifying the weaker ones. Summative purposes are closely associated with *accountability-oriented faculty evaluation*, which emphasises the use of external control, rewards and sanction to promote excellence, outcomes and productivity (Cherry et al., 2017; Field, 2015; Gonzales & Núñez, 2014; Lonsdale, 1998; Saunders et al., 2011; Sułkowski et al., 2020; Yokuş & Yelken, 2019) (see Figure 1). This orientation favours analytical methods and quantitative instruments to determine if individuals or institutions have satisfied the minimum performance requirements (Guba & Lincoln, 1989; Lonsdale, 1998; Scriven, 1991).

However, the effectiveness of the accountability-oriented evaluation is elusive because its key assumptions are inconsistent with the contemporary theory of effective management and leadership that aims at empowering employees and valuing collective performance (Lonsdale, 1998). The accountability-oriented evaluation usually involves minimal ownership or decision-making power of faculty members (Lonsdale, 1998; Scriven, 1991), causing them to suffer from potentially unfair and unequal treatments due to their limited power during the evaluation process (Guba & Lincoln, 1989). The generalised evaluation outcomes, which are not context-specific, may have superficial value to practice improvement at the local or grassroots levels (Guba & Lincoln, 1989). The oversimplified evaluation outcomes may prevent HEI leaders and administrators from being morally responsible for the teaching and learning quality, as they may reason that their roles are complete “when the report is delivered” (Guba & Lincoln, 1989, p. 38).

Formative Purposes

Formative evaluation mainly involves providing faculty members with constructive feedback, assistance, or guidance for improvement (Buller, 2012; Seldin, 1984; Scriven, 1991). This process is also intended to offer data for HEIs to plan professional learning and development for faculty members (Koçke et al., 2017; Wei, 2015; Yao & Grady, 2005). However, many authors, including Scriven (1991), consider the formative process is another step towards summative evaluation, and such process should be separated from the job of the evaluators. Formative purposes relate to *improvement-oriented faculty evaluation*, which concerns linking evaluation with developing faculty members to maximise institutional and individual performance (Lonsdale, 1998). Such an evaluation process may involve shifting back and forth between formative and summative evaluation by first identifying faculty members' strengths and weaknesses and then providing them with developmental or

mentoring programmes (Buller, 2012; Lonsdale, 1998; Scriven, 1991). However, an improvement-oriented evaluation may still be inadequate to transform practice (Hallinger, 2010; Lonsdale, 1998) for several reasons.

First, while evaluation for improvement is linked to quality enhancement, the notion of quality is flexibly interpreted depending on the context. While much of the debate on higher education revolves around quality, there is no consensus on the understanding and the use of quality in higher education (Darojat et al., 2015; H. T. Pham & Starkey, 2016). Hazelkorn et al. (2018) found various interpretations of quality as teaching and learning, research engagement, institutional leadership, meeting social needs, “excellence” (p. 6) in reaching different objectives, good governance, or national competitiveness. From an educational perspective, *quality* is linked to the relationship between different agents (Ehlers, 2009). In this sense, a *quality culture* involves a synergy of various elements to increase staff commitment, ownership, empowerment, and knowledge that lead to continuous improvement of practice (Bendermacher et al., 2017). However, due to the lack of a holistic understanding of quality, quality is usually equated with bureaucratically documenting performance rather than promoting individual members’ continuous learning and improvement (Ehlers, 2009).

Second, improvement effort lacks effectiveness due to the emphasis on using formal quantitative instruments and little faculty ownership over the evaluation process (Guba & Lincoln, 1989). Although the improvement-oriented approach to faculty evaluation emphasises the ongoing enhancement of staff performance by identifying staff members’ developmental needs, this approach to evaluation has turned out to be unsuccessful with inadequate evidence of sustained performance improvement and lack of real commitment from staff who saw appraisal as a primarily administrative or symbolic activity (Hallinger, 2010; Lonsdale, 1998).

Transformational Purposes

Transformational evaluation involves viewing evaluation as a continuous learning and adapting process that recognises and integrates the needs of primary intended evaluation users⁷ such as administrators or faculty members (Patton, 1996). Transformational purposes include:

- generating general knowledge about and principles of program effectiveness
- developing programs and organisations
- focusing management efforts
- creating learning organisations

⁷ *Primary intended users* are defined as “specific people, in a specific position, in a specific organization who will use the evaluation findings and who have the capacity to effect change”, from https://www.betterevaluation.org/en/rainbow_framework/frame/identify_primary_intended_users

- empowering participants
- fully integrating evaluation into the intervention, and
- stimulating critical reflection on the path to more enlightened practice. (Patton, 1996, p. 142)

Transformational purposes are closely linked to *learning-oriented faculty evaluation*, which involves increasing faculty members' opportunities for learning and discovery through increased ownership, adaptation, interaction, and trust-building (Dahler-Larsen, 2009). The learning-oriented evaluation is rooted in the evaluation theories and practices that emphasise the local context, practitioners' engagement, and the value of dialogues that exchange multiple views between evaluators and evaluatees (Guba & Lincoln, 1989; Ryan & Cousins, 2009). *Learning* is defined as a multifaceted notion that involves a "complex, contextually sensitive, and dialogical process" among those who are "attentive and mindful of what is said on the spot during the evaluative process" (Niessen et al., 2009, p. 375). As Dahler-Larsen (2009) put it,

learning refers to several types of social and individual adjustments to changing circumstances... learning is difficult to standardise as the feedback loops of reflexivity take place in many forms, on many levels in the social system, and under constantly unstable circumstances. (p.311)

Transformational evaluation is a relatively new idea in response to a critique of limited utility or little impact of traditional evaluation for practice improvement (Guba & Lincoln, 1989; Lonsdale, 1998; Morris et al., 2007). Learning-oriented faculty evaluation with transformational purposes may work better in countries like New Zealand or Australia, where "issues of social justice and human rights are of central concern" (Mertens, 2010, p. 9). In the communist-embedded context of Vietnam, this new concept might be challenged by the centralised leadership's political and social context and lack of institutional autonomy (L. H. Phan & Doan, 2020). However, the transformational purposes are worth investigating since lessons learnt from the Vietnamese case can hopefully help understand and resolve prevalent problems of the current accountability- and improvement-oriented evaluation approaches in similar contexts. The learning-oriented evaluation aims to engage implementing agents in scientific inquiry; thus, it increases the voices of those who are often unheard in evaluation science (Patton, 2018). Although an emphasis on evaluation for learning is not a new concept within the education sector (Sinnema, 2005), much of the previous work that focuses on teaching and learning improvement deals with the K-12 sectors, not higher education (Shaked, 2021). Thus, findings from my study are expected to broaden the learning-oriented approach to evaluation in higher education.

Faculty Evaluation Criteria

Establishing clear, objective, and consistent criteria or standards for faculty evaluation is essential for any effective evaluation model. An explicit set of criteria or standards presented in the faculty handbook or employment contract helps reduce disagreements during the evaluation process. Faculty evaluation criteria generally concern three primary responsibilities: teaching, research and service (Arreola, 2007; Khtere, 2020). However, establishing faculty evaluation criteria is challenging in many HEIs because it is difficult to define and quantify such a broad scope of assessing faculty work (K. Chan, 2001; Teelken, 2018).

Evaluation of Teaching

There have been multiple perspectives on what counts as effective teaching. One common view of evaluating teaching is to attain absolute objectivity with a formula of the most effective teaching (Berk, 2018; Cranton, 2001). Under this view, teaching evaluation involves using student rating forms based on the faculty members' behaviours and styles for various disciplines. In a broader sense, teaching is any interaction that enables student learning (Arreola, 2007; Norsworthy & Sanders, 2021). Under this view, teaching evaluation needs to consider faculty members' efforts to design meaningful interactions that facilitate student learning. For example, HEIs can use a teaching evaluation model comprised of five broad dimensions to evaluate teaching effectiveness: (a) content expertise, (b) instructional design skills, (c) instructional delivery skills, (d) instructional assessment skills, and (e) course management skills (Arreola, 2007, p. 19). Evaluation committees can observe faculty members' instructional delivery skills through oral presentation, written communication, or creating a favourable learning environment (Arreola, 2007).

There have been many changes in the methodologies used for teaching evaluation across HEIs with more systematically collected evidence (Lohman, 2021; Ory, 2000). Teaching evaluation has shifted from being the primary responsibility of the dean or department chairs to a formal system using multiple methods such as student ratings, peer feedback, self-evaluation, or teaching portfolios (O'Leary & Savage, 2020). Despite a more systematic teaching evaluation system, many faculty members still found the process more summative than formative (Lyde et al., 2016). Thus, a transformational process of evaluating faculty members' teaching is pivotal.

One way to promote the transformational teaching evaluation process is to increase faculty members' ownership of the process through ongoing reflection, learning and improving their scholarship of teaching and learning (Boyer, 1991; Dewar & Perkins, 2021; Jeffs et al., 2021; Smith, 2001; Theall & Centra, 2001). This transformational process involves taking an interpretive and critical teaching evaluation approach in which faculty members design

teaching portfolios or participate in the individual and group reflection process (Cranton, 2001). For example, SET can be considered as a GIFT-giving process in which students and educators engage with the SET feedback to promote student learning (Norsworthy & Sanders, 2021). When interacting with evidence on different facets of teaching from multiple data sources, faculty members can critically reflect on their assumptions, perspectives, norms, beliefs, and values (Cranton, 2001). Such activities would stimulate faculty members to become more skilful in teaching and become critically reflective contributors of teaching and learning scholarship (Dewar & Perkins, 2021; Smith, 2001).

Evaluation of Research and Service

Despite the supposedly equal importance of teaching and research, many HEIs have prioritised research performance in their faculty evaluation system (Bogt & Scapens, 2012). The reliance on research evaluation on peer review and metrics has resulted in “an ethical crisis in scholarly communication and the reward system in science”, giving rise to “the proliferation of predatory journals and conferences” (Madhan et al., 2018, p. 1). While research productivity has commonly been a primary criterion in faculty evaluation, it does not guarantee teaching quality, and it fails to consider numerous demands, tasks, and creative works of faculty members (Cadez et al., 2017; Stephens, 1999). Likewise, the service evaluation criterion usually involves counting how many times faculty members’ participate in tasks such as reviewing academic journals (Arreola, 2007; Sampson et al., 2010; Seldin, 2006). An emphasis on quantity rather than quality has created tension for many HEIs and faculty members between maximising their research outcomes and contribution to student learning and social betterment (Cenamor, 2021; Lei et al., 2021; Shields & Watermeyer, 2020).

More inclusive evaluation criteria that recognise a broader range of tasks that faculty members fulfil to contribute to research and scholarship would motivate their external engagement in research advancement and social contribution (Atta-Owusu & Fitjar, 2021). Broadening the scopes of faculty work in the evaluation system by substituting the research criteria with scholarly and creative activities such as the scholarship of proficiency, discovery, dissemination, and translation would help recognise faculty members’ continuous engagement in professional learning, sharing and translating their research findings to benefit the professional and broader community (Arreola, 2007).

Faculty Evaluation Practices

There are multiple faculty evaluation practices, each of which has its strengths and weaknesses depending on the contexts and manners of those involved (Berk, 2018; Hightower, 2010). My study focused on the two most common practices: SET – an

international practice relatively new in Vietnam and voting evaluation – a national communist-embedded practice strongly tied to Vietnamese public HEIs.

SET: A Global Practice

SET, which involves surveying students' opinions about courses, lecturers, or other teaching faculty members, was formally introduced in the USA in the 1920s (Darwin, 2016). SET was widespread in the USA, Australia, and New Zealand in the 1960s, and now it has become a common practice in global HEIs (Barrow & Grant, 2016; Darwin, 2016).

SET in Global Higher Education. There has been a shift in the motives for using SET in global higher education over the past several decades. The original motives for SET were advancing faculty members' pedagogical knowledge and skills and responding to students' protests for increased participation in university governance (Darwin, 2016), which is associated with satisfying professional accountability (Røiseland et al., 2015). Students' feedback is helpful to inform changes in various aspects of teaching, such as course content, teacher-student interaction, and the clarity of faculty presentation and student advising (Seldin, 2006). More importantly, SET contributes to the democratic process of letting students' voices be heard, maintaining continuous dialogues about learning and improvement, and reinforcing institutional norms for effective teaching (Palermo, 2013; Serdyukova et al., 2010).

However, SET has recently been used primarily to satisfy various audit, review, accreditation, or managerial requirements (Darwin, 2016; Thiel, 2019). For example, in New Zealand, SET has been used as an instrument for institutional compliance with the audit-governing bodies (Barrow & Grant, 2016; Norsworthy & Sanders, 2021). SET in New Zealand, for instance, has shifted from its original purpose of supporting teaching and learning in the 1960s to the more bureaucratized practice, which is a part of the audit system (Barrow & Grant, 2006). Norsworthy and Sanders (2021) questioned if the students' voices are authentically heard when online SET is used with a low response rate (within 30% to 40%). Like those in New Zealand, studies in various higher education contexts have raised concerns over the consequences of the managerially focused SET, which is associated with low validity and reliability, grade inflation and negative influence on faculty members' identities.

Concerns Over SET Validity. Various studies (Clayson, 2017; Curby et al., 2020; Oon et al., 2017; Uttl et al., 2017) have questioned SET validity which refers to the accuracy of SET rating scales for measuring what needs to be measured. One SET problem involves invalid SET measurement scales or rating tools that pose potential bias against faculty members. Ray et al. (2018), in a study into 1,074 questions from 55 SET forms across a range of HEIs in the USA, revealed that faculty member performance questions include components beyond their control, such as student participants or course-related issues. In a similar vein, a

meta-analysis of over 2,000 samples of SET by Curby et al. (2020) showed that measurement errors account for nearly a quarter of the ratings. The SET rating centres on other factors than the quality of the instructors. Invalid SET instruments and the way SET is conducted have resulted in unreliable SET scores, which means SET scores do not necessarily reflect faculty members' teaching effectiveness. The ratings are influenced by nonteacher factors such as class size, student tiredness, course timing, instruction levels, or academic disciplines (Dawson et al., 2020; Hajdin & Pažur, 2012; Roxå et al., 2021). Student ratings are also impacted by faculty members' physical attractiveness, gender, and rapport with students (Lammers et al., 2017). Even when SET tools focus on instructor performance, they tend to measure instructor popularity rather than effectiveness due to questions about students' satisfaction levels (Hornstein, 2017).

Grade Inflation. Another potential risk of SET is grade inflation, or faculty members' lenient grading of student assignments in exchange for or anticipation of high SET scores (Eiszler, 2002; Hunt, 2008). In America, positive relationships have been empirically established between teaching faculty members' lenient grading and higher student rating scores (Eiszler, 2002; Hunt, 2008). Similarly, in a Taiwanese HEI, both students' final grades and course failure rates were predictors of student evaluation scores of teaching faculty (Chen et al., 2017). Unfortunately, faculty members' increased grading leniency and tendency to please students with lower course requirements are necessarily related to improved student learning (Berezvai et al., 2021; Chen et al., 2017).

Negative Influence on Faculty Members. SET has several demoralising effects on faculty well-being and self-image, especially among those with continuously negative student ratings (Hammer et al., 2018; Mutch & Tatebe, 2017). For example, Hammer et al. (2018) found that about one-third of 2,323 faculty members from over 20 HEIs in Israel felt the SET ratings underestimated their teaching contribution. A subgroup of 9% felt strongly misjudged or even victimised by their students. These faculty members had more criticism and less trust in student evaluations. With their sense of being underestimated by students, those teaching faculty members would be prone to hostility and resistance to feedback. Faculty members' teaching would be unlikely to improve without a receptiveness to feedback (Hammer et al., 2018). In the same vein, Mutch and Tatebe (2017) reflected on their disheartened experience of receiving negative SET and experiencing students' resistance to the courses they had passionately taught for many years. Such demotivating experience with SET left the faculty members with an uncertain sense of professional identity after many years of serving in higher education (Mutch & Tatebe, 2017).

Learning and Improvement Attempts. Existing literature has revealed a lack of systemic improvement resulting from SET deployment (Lohman, 2021; Shah et al., 2017). For example, in Scotland and Australia, various HEIs took initiatives to increase students' sharing

of their learning experience and participation in institutional committees; however, these efforts resulted in little influence (Shah et al., 2017). Some HEIs took alternative approaches to include more qualitative data for SET from several sources in the USA; however, their approaches have not resulted in a systematic and effective process for improving faculty teaching effectiveness (Lohman, 2021). One explanation for the ineffectiveness of using SET for instructional improvement is that top-down policies are distant from the values and beliefs of those involved (Henderson et al., 2011; Roxå et al., 2021). As stated by Muliaina (2018), “any social or educational change or development that is not anchored in the knowledge and value system of the target population is destined to fail” (p. 519). Thus, the potential of using SET for teaching and learning improvement is only possible when the whole SET processes integrate the worldview, values, and beliefs of faculty members and students (Norsworthy & Sanders, 2021; Roxå et al., 2021).

Some attempts have been made to tackle the unintended effects of SET and optimise its use for learning and improvement (Babad et al., 2021; Mutch & Tatebe, 2017; Norsworthy & Sanders, 2021; Roxå et al., 2021; Stein et al., 2021; Thiel, 2019). One suggestion was to redesign SET to focus on affective aspects rather than just on cognitive aspects of the faculty member’s performance (Babad et al., 2021). Another suggestion was to reframe SET as continuous or reflexive dialogues between HEIs, faculty members, students (Roxå et al., 2021; Stein et al., 2021). Instead of focusing on quantitative measures or imposing managerial control over faculty members, HEIs could turn SET into a meaningful collective experience for faculty members in a community of practice with a shared goal of learning and care (Mutch & Tatebe, 2017).

SET in Vietnamese Higher Education. SET is a practice that originated in the West, such as the USA and Europe, before spreading its influence across Asia countries, including Vietnam (Pham, 2019). SET in Vietnamese higher education has been tightly linked to the development of the QA system over the past 15 years. Since HERA (GoV, 2005), the Vietnamese government has developed a national QA and accreditation system across Vietnam. This reform agenda paved the way for the pilot stage of SET between 2008 and 2009. Until 2010, the MoET (2010) mandated all HEIs to start collecting SET of all teaching faculty members. The new correspondence by MoET (2013) reflected its aspiration to encourage SET for the dual purposes of improvement and accountability.

SET as an Important QA Task. In Vietnamese higher education, SET is considered one of the critical pieces of evidence that demonstrate Vietnamese HEIs’ efforts to assure the teaching quality of their staff. Besides, both international and national QA frameworks emphasise HEIs’ efforts to evaluate and improve the quality of faculty members. For example, Criterion 5 in programme accreditation of the ASEAN University Network (2020) indicates eight variables, including the evaluation of academic staff based on:

- articulated competence that is linked to faculty roles and professional ethics,
- a merit system or performance management including reward and recognition, and
- the improvement of faculty teaching, research, and services (p. 28).

Following the QA standards of the ASEAN University Network (2020), the MoET published policy requirements about ensuring the quality of faculty members in Vietnamese HEIs (Pham & Nguyen, 2020). To be fully accredited by international or national accreditation agencies, Vietnamese HEIs are required to demonstrate their attempt to assure many aspects of their faculty members' performance, including teaching.

Reasons for Limited SET Impact. Despite Vietnamese HEIs' widespread use of SET since the MoET (2010) mandate, the use of SET for teaching and learning improvement is generally limited for several reasons (Le & Do, 2017; N. D. Tran & Nguyen, 2015; Wei, 2015). First, SET is constrained by the slow progress of QA development, resulting in a limited impact on quality improvement across Vietnamese HEIs (C. H. Nguyen & Shah, 2019). Second, several cultural obstacles hinder Vietnamese HEIs' effective use of SET. SET is generally considered a Western practice incompatible with the Vietnamese Confucian tradition of respecting teachers (K. D. Nguyen & Mcinnis, 2002; T. T. Tran, 2018). Under these Confucian views, faculty members are considered authority figures, whereas students might be viewed as less mature adults whose perceptions are less unreliable. Administrators, therefore, are unwilling to use SET for formal faculty evaluation due to their fear of threatening "the privileged position held by teachers in Vietnamese society" (K. D. Nguyen & Mcinnis, 2002, p. 156). Third, Vietnamese HEIs' effective use of SET is constrained by their low funding for improving human resources and their leaders' incapability to formulate development plans for faculty members' professional learning and teaching (K. D. Nguyen, 2003).

Conditions for Increased SET Impact. Despite the currently limited effect of SET on teaching and learning improvement, Vietnamese HEIs have some conditions for increasing the impact of SET (P. V. Nguyen, 2020; N. D. Tran & Nguyen, 2015; T. T. Tran, 2018). Under the Confucian legacy, Vietnamese faculty members tend to have intrinsic motivation to undertake the teaching job, characterised by their learning attitude and appreciation of students' respect and affection (Pham, 2018). The future potential of SET is possible if faculty evaluation policies and practices incorporate faculty members' intrinsic motivation. Despite Confucian teacher-centredness, many Vietnamese faculty members demonstrate a positive attitude about SET when SET implementation involves careful consideration of institutional norms and individual beliefs (P. V. Nguyen, 2020; N. D. Tran & Nguyen, 2015; T. T. Tran, 2018). Specifically, T. T. Tran (2018) suggested the following strategies to redesign SET would accommodate Vietnamese faculty members' values, expectations and needs and reduce their hostility towards SET:

- revising of SET forms based on open discussion of administrators or faculty members, including more qualitative questions, and removing faculty members' perceived meaningless questions;
- eliminating bias in student responses by using online instead of manual feedback collection;
- making SET quantitative scores more informative by adding the average scores of faculty members across the divisions, departments, and institutions;
- keeping SET results confidential and safe for faculty members.

Overall, the use of SET for learning and improvement in Vietnamese higher education, as in the case of SET, is an uneasy but not impossible task.

Voting Evaluation: A Vietnamese Practice

Voting evaluation is a practice that is unique within the communist-embedded Vietnamese public organisations, including public HEIs. The voting practice is a part of the national emulation movements and a key annual performance appraisal within Vietnamese HEIs. Despite the popularity of the practice, literature on voting evaluation, like those about the Vietnamese emulation movement, has been “entirely lacking” (Homutová, 2018, p. 19). Previous studies (Hung, 2013; K. D. Nguyen, 2000; V. K. Nguyen, 2015; H. T. Pham, 2014; N. Phan, 2014; X. B. Tran, 2010) view Vietnamese HEIs' faculty evaluation as a taken-for-granted or culturally constrained performance appraisal practice. However, existing research has not delved into the sociopolitical context and the underlying ideologies of the practice. My study, which integrates the view of evaluation as a national emulation movement, will deepen sociopolitical understanding of the Vietnamese HEIs' faculty evaluation.

Voting Evaluation as an Emulation Movement. The voting practice is also known as *hoạt động bình bầu thi đua khen thưởng* (voting, emulation and commendation/ reward practice). The practice has been widespread from national to the grassroots levels, mandating leaders and people at public organisations to launch and participate in various emulation campaigns⁸ (Homutová, 2018). Following the National Law on Emulation and Commendation 2003, all public HEIs organise the voting evaluation is mandatory for all employees, including full-time faculty members. This practice involves individual faculty members completing self-achievement reports, sharing their reports with other faculty members before or during a whole faculty evaluation meeting then voting (usually on discreet ballot papers) among themselves for the emulation titles of various ranks. As voting was the primary means to finalise faculty

⁸ The emulation campaign, also referred to as the emulation movement, is used in socialist countries to denote a series of organised small-scope activities within a broad movement, to achieve a particular (usually political) aim (Homutová, 2018)

members' evaluation results, this evaluation is often called the “voting evaluation”⁹ by the participants.

Emulation campaigns in Vietnam started in 1948 when President Ho Chi Minh, the Vietnamese government's first representative, announced *The Calling for Patriotic Emulation* to encourage Vietnamese citizens and soldiers to show their patriotism by fighting against the enemies of poverty, illiteracy, and foreign invasion (Thanh, 2016). The early emulation campaigns in Vietnam imitated Soviet or Chinese counterparts to improve people's productivity and reinforce the socialist ideology, but they were contextualised to be well linked with “patriotism,” the love for protecting the country and the tradition of “heroism” or personal sacrifice to preserve the fatherland (Homutová, 2018, p. 273). Ho Chi Minh's popular quote, “emulation means patriotism, those who are patriotic must emulate” (Ho Chi Minh, 1952 as cited in Bui et al., 2011, p. 402), has continuously been consolidated and promoted nationwide by the government led by the Vietnamese Communist Party until today.

The voting evaluation is one key practice of Vietnamese public HEIs' to fulfil the essential political mission of public organisations (T. V. Pham, 2016). The emulation campaign's initial purpose was to motivate citizens of all ages to contribute to the protection and construction of the nation, whereas the voting evaluation at Vietnamese public HEIs aim to identify, incentivise and multiply individuals with outstanding achievements (Grand University, 2015; Ministry of Home Affairs, 2020). Outstanding faculty members and leaders who have been selected at institutional levels will be invited to a national congress. In the most recent congress on emulation in education in 2016–2020, the Minister of Education and Training awarded 400 academic staff and student delegates with honourable titles such as the “Labour Hero”, the “People's Teachers”, or “Outstanding Students” for their outstanding achievements such as winning medals in competitions (Ministry of Home Affairs, 2020).

Ineffective Voting Evaluation. As an annual performance appraisal, the voting evaluation in the Vietnamese context was perceived to be important but ineffective by many Vietnamese academics (K. D. Nguyen, 2000; H. T. Pham & Nguyen, 2020; X. B. Tran, 2010). At the institutional level, many Vietnamese HEIs assessed their faculty members based on the quantity rather than on the quality of their work (T. H. Nguyen, 2016). Many faculty members found the evaluation criteria more politically than professionally oriented, such as how well faculty members maintain their relationships with other colleagues or how good they understand Ho Chi Minh's ideology (Pham & Nguyen, 2020). Evaluation titles such as Advanced Labourers are outdated and fail to distinguish the outstanding and committed members from the less achieving ones (Bach & Nguyen, 2009). Even when one institution

⁹ For this reason, I use the term “voting evaluation” throughout the thesis to refer to the “evaluation for emulation and commendation” or “evaluation for emulation and rewards” (công tác thi đua và khen thưởng) at the participating public HEIs in Vietnam.

attempted to adapt the criteria, faculty members still did not see the appraisal as a source of motivation because of its low material rewards and its lack of professional values (Pham & Nguyen, 2020). The material rewards are constantly low; for example, those for the satisfactory and excellent performance range from 50.000 Vietnamese Dong (VND) (about 2 USD) to 300.000 VND (about 12 USD) (K. D. Nguyen, 2000). Together with an ineffective salary framework and the fact that numerous teaching faculty must moonlight for extra income, it is challenging to promote faculty engagement in the evaluation and advancement of quality in Vietnamese HEIs (H. T. Pham, 2012; H. T. Pham & Nguyen, 2020).

At the national level, the voting evaluation is associated with the *disease of achievement (bệnh thành tích)*, a phenomenon commonly reported in Vietnamese society and within the education sector (Homutová, 2018). Several patterns of bad habits during annual voting processes (*bệnh mùa bỏ phiếu*), among public employees, potentially including faculty members, are

- “spontaneously good” disease (*bệnh tốt đột xuất*)
- “comparing a bunch of chopsticks to (not) choose a flagpole” disease (*bệnh so bó đũa [nhưng không] chọn cột cờ*)
- “defaming others” disease (*bệnh đim hàng, hạ bệ*)
- “saving the best (titles) for the boss” disease (*bệnh nhường sếp*)
- “honouring intimate fellows” disease (*bệnh bè phái*)
- “say what you don’t do” disease (*bệnh nói một đằng, làm một nẻo*). (Song, 2021, n.p.)

As the list of common diseases suggests, the collective voting process potentially causes faculty members to behave opportunistically to gain the personal advantage of being voted with higher titles. As Song (2021) explained, these behaviours “have appeared in many localities, agencies and units,” and they have a “strong effect [on public employees] before each year-end contest using the vote of confidence or vote of trust for human resource planning” (n. p.). Within the Vietnamese HEIs, the overemphasis on achievement during the voting evaluation has driven Vietnamese public HEIs to focus more on “polishing” the reports rather than on improving practices (Huynh, 2016, p. 197).

Reasons for Ineffective Voting Evaluation. There are several cultural and sociopolitical constraints leading to the ineffectiveness of the voting evaluation in Vietnamese HEIs.

Cultural Constraints. Several cultural constraints hinder the Vietnamese HEIs’ voting evaluation. Vietnam is considered a country “without a culture of evaluation” (H. T. Pham, 2019, p. 182) since the judgemental sides of evaluation are incompatible with the Confucian culture of face-saving and harmony maintenance. The QA of faculty performance is

considered a mechanism recently imported from the West, and many would doubt its feasibility and practicality in Vietnamese HEIs (H. T. Pham, 2019; H. T. Pham & Nguyen, 2020). Thus, many Vietnamese academics are resistant to the judgemental aspects of evaluation (Pham, 2014). Also, the voting processes involve competition for outstanding titles, resulting in conflicts of interest and potentially unfair evaluation results. The problem is that Vietnamese public staff of various positions and seniority levels, including administrators and faculty members, often compete by voting among themselves for the limited titles of excellence¹⁰. Consequently, like many public employees, teaching staff in public HEIs and schools are subjectively evaluated by considering their relative positions in the organisations (Homutová, 2018; “Xét chiến sĩ thi đua,” 2017). Given the collectivist Confucian culture of Vietnamese people paying respect to relationship and hierarchy, behaviours such as “saving the best (titles) for the boss” and “honouring intimate fellows” (Song, 2021, n.p.) is understandable.

Sociopolitical Constraints. Several sociopolitical constraints inhibit the Vietnamese HEIs’ voting evaluation. First, Vietnamese HEI leaders paid inadequate attention to evaluating and developing faculty members due to their lack of institutional autonomy in resource management (Dao & Hayden, 2019; X. B. Tran, 2011). The faculty performance appraisal of Vietnamese HEIs, especially the public institutions, is highly dependent on the government and MoET’s regulations and guidelines (Pham & Nguyen, 2020). Vietnamese HEIs are expected to comply with two key policy documents that direct the performance appraisal in Vietnamese HEIs, including the Law on Emulation and Rewards 2003 and the Law on Public Employees 2010. Although faculty members engage in self-assessment and peer assessment, the final decision is then considered again at the faculty and institutional levels. HEIs’ institutional autonomy is heavily restricted by the line-management control mechanism in which HEIs’ major decisions about human and financial issues are subject to a governing body (Lam, 2010). According to Lam (2010), Vietnamese universities and academies are under the governance of 18 ministries and more than 60 state governing bodies, including provincial governments. Except for two national universities, which are under the direct governance of the Cabinet, the MoET retains its line-management responsibility for most of the leading universities. The Cabinet oversees issuing and implementing regulations related to the Higher Education Law, and the MoET is responsible for enforcing these regulations across the higher education sectors. However, due to the fragmentation and disparities resulting from multiple line-management arrangements, the government and MoET’s capacity to foster policy implementation at the local level have become weakened (Lam, 2010).

Also, Vietnamese HEIs and individuals demonstrated limited levels of accountability in various practices, including faculty evaluation. Although *accountability* has been more

¹⁰ From <https://giaoduc.net.vn/giao-duc-24h/khong-phai-cu-dat-chien-si-thi-dua-chac-chan-la-giao-vien-gioi-post216718.gd?>

explicitly defined as the demand for HEIs to be held accountable for their activities (trách nhiệm giải trình)¹¹, the Vietnamese government, represented by the MoET, has not had an effective mechanism for enforcing HEIs' commitment to regulations (Salmi & Pham, 2019, p. 110). Teaching staff at public HEIs also maintain a secure place and do not hold accountable for their poor performance. (T. L. Pham, 2012). Accountability in Vietnamese HEIs is likely to be "limited to standards or criteria for accreditation" (Felix, 2020, p. 116). Also, Vietnamese HEIs leaders' decision-making power is constrained by the government and MoET's centralised governance. Within Vietnamese HEIs, the central authority for decision making includes three bodies: a Vietnamese Communist Party representative, the rector, and the governing board (Salmi & Pham, 2019). Although the governing boards are supposed to function independently in improving HEIs, they are, in practice, under the Vietnamese communist leadership since the chairs are usually Vietnamese Communist Party members who reinforce the HEIs' compliance with the MoET or other line-management instructions (L. H. Phan & Dang, 2020). In the revised Higher Education Law in 2018, the Vietnamese government expressed a solid determination to reinforce institutional autonomy and accountability by emphasising the role of university councils – not as a bureaucratic unit but as the core organisation to fortify institutional autonomy. However, it is unclear if the autonomy and accountability levels of Vietnamese HEIs will improve significantly in the long run because "the culture of governance and management" within these HEIs still preserves its "legacy of centralised control" (Hayden & Le-Nguyen, 2020, p. 33).

Concluding Remarks

The literature review has highlighted three key purposes of faculty evaluation: summative, formative (Scriven, 1991) and transformational (Patton, 1996), which are linked respectively to accountability, improvement (Lonsdale, 1998), and learning orientations (Dahler-Larsen, 2009). The two key practices of this study, the SET and voting evaluation, have been reviewed considering their effectiveness concerning accountability, learning and improvement orientations. Existing literature reveals that many global HEIs emphasise the use of SET for managerial accountability (Røiseland et al., 2015) in response to neoliberal forces. Despite some attempts to use SET for learning and improvement, these efforts have not led to systemic change. In the Vietnamese higher education context, the use of SET for learning and improvement is limited by various cultural and institutional constraints. Unlike SET, voting evaluation is a unique practice in communist-led Vietnamese higher education. Within Vietnamese higher education, previous studies reveal that some cultural and sociopolitical

¹¹ Mentioned in the Amendment of the Higher Education Law 2018

constraints have hindered the use of faculty evaluation for accountability, learning, and improvement.

Overall, there are several gaps in the current literature. First, previous studies focus on highlighting faculty evaluation problems without examining the combined possibilities and conditions that sustain effective use of the evaluation for learning and improvement (Cherry et al., 2017). Even in some cases where the conditions for improvement were explored, they tend not to see the faculty evaluation problems in “the whole problem set” (Robinson, 1993, p. 8). Thus, current higher education faculty evaluation research lacks emphasis on the degree of improvability within the current evaluation system. Furthermore, previous studies investigated faculty evaluation using preexisting theories rather than the theories of those involved. It is unclear from the existing literature as to which actors/ agents were involved in the problem resolution and how they sustained the problems of faculty evaluation ineffectiveness. These gaps in existing literature demand research on problems of faculty evaluation practices and an actionable plan to improve the system, specifically from the perspectives of active and responsible actors (Robinson, 1993). An educational problem is “a gap between an existing and a desired state of affairs” (Robinson, 1993, p. 25), so this study examines the problems of the participants’ potential lack of using the faculty evaluation practices for learning and improvement.

Chapter 3: Methodology

The problem this study intended to address was faculty evaluation potentially not being used for learning and improvement purposes. I chose problem-based methodology (PBM) (Robinson, 1993) to gain insights into the participants' theories of action (Argyris & Schön, 1974) concerning faculty evaluation. The constructed theories of action helped reveal the current evaluation practices, the consequences, and, most importantly, the set of constraints that guide the current evaluation practices (Robinson, 1993). The constraint analysis, which involved looking into the sociopolitical context and ideological influences on the participants' actions, is expected to address the lack of "philosophically robust ground" (T. N. Nguyen, 2019, p. 17) for understanding and improving higher education faculty evaluation in Vietnam. This chapter will elaborate more on how my study was designed and conducted under PBM.

Research Approach

My study used PBM and a qualitative case study to examine faculty evaluation practices in Vietnamese public HEIs. Being informed by PBM, my study explored the theories of action (Argyris & Schön, 1974) concerning faculty evaluation practices from the perspectives of key stakeholders (i.e., QA officers, administrators, and faculty members).

Research Questions

Under PBM, with the adoption of theories of action (Argyris & Schön, 1974), my research aimed to explore the participants' approaches to faculty evaluation, together with the constraints and consequences of these approaches. Overall, the study intended to address the following questions, each corresponding to a key theory-of-action component (approaches, constraints that explain the approaches and consequences of the approaches):

- (i) What approaches are currently taken to faculty evaluation?
- (ii) What explains the approaches?
- (iii) What is the impact of the approaches?

Answers to the research questions, which were based on the analysis of document and interview data, would reveal the intended purposes and actual uses of the selected faculty evaluation practices for accountability, learning and improvement. Such answers enabled me to attain the overarching research goal of identifying approaches to higher education faculty evaluation practices that promote learning and improvement.

PBM

The research adopted *PBM* which is defined as a methodology used to examine "the theories of action relevant to the problem situation" and, if necessary, evaluate and alter them

(Robinson, 1993, p. 7). The primary purpose of PBM research is to reduce the gap between research and practice (Robinson, 1992, 1993) by doing research “in ways that contribute to the improvement of practice” (Robinson & Lai, 2006, p. 80). PBM was formulated by Robinson (1993) as a methodology providing educational and social researchers with a way of conducting research that directly contributes to practice improvement. PBM was then expanded by Robinson and Lai (2006) to guide practitioner-researchers to investigate and improve their practices.

PBM Versus Conventional Research. PBM has some distinctive qualities compared to conventional research. Unlike conventional research which usually starts with a “theoretically derived hypothesis” (Robinson, 1993, p. 14), PBM research is based on theories that are relevant to those employed by the practitioners who are actors, agents or stakeholders of a practice. General education research is usually theoretically driven and statistically controlled, so its findings may not be relevant to the practice culture (Robinson, 1993). The answers to researchers’ generated problems might be influenced by constraints different from practitioners’, so PBM research addresses the limited impact of conventional research by exploring the theories of action of those involved in educational practice. Like action research, PBM research shares similar concerns about practitioners’ actions and improving practice (Herr & Anderson, 2014; Robinson, 1993). However, PBM is distinctive in its emphasis on understanding the participants’ theories about the practice before making any effort to change or improve it (Robinson & Lai, 2006). This distinctive quality makes PBM stand out as “an innovative, comprehensive, and useful methodology,” which is “arguably the best extended methodological treatment of action research currently available” (Haig, 2013, p. 356).

Theory of Action (ToA). PBM is based on *ToA* which is defined as people’s conscious and unconscious reasoning process, including their goals, values, beliefs or “perceptions of institutional constraints such as legal, regulatory, and resource issues” (Robinson & Lai, 2006, p. 22), that guide or explain their actions (Argyris & Schön, 1974; Dick & Dalmau, 1991, Robinson, 1993). Simply speaking, a *ToA* describes the links between what people do (*strategies* or *actions*), the values and beliefs that explain or influence the actions (what Robinson calls *constraints* in PBM), and the *consequences* of taking those actions (Robinson, 2011). PBM is drawn on constraint-inclusion theory, which defines a problem as a demand to achieve a goal and the constraints on achieving that goal (Nickles, 1981; Robinson, 1993). Thus, to solve a problem of practice involves understanding and altering the theories people attribute to themselves and others based on their values, beliefs and motives:

To understand an educational problem, therefore, is to understand the theories of action of relevant agents and the factors that sustain those theories. To resolve an

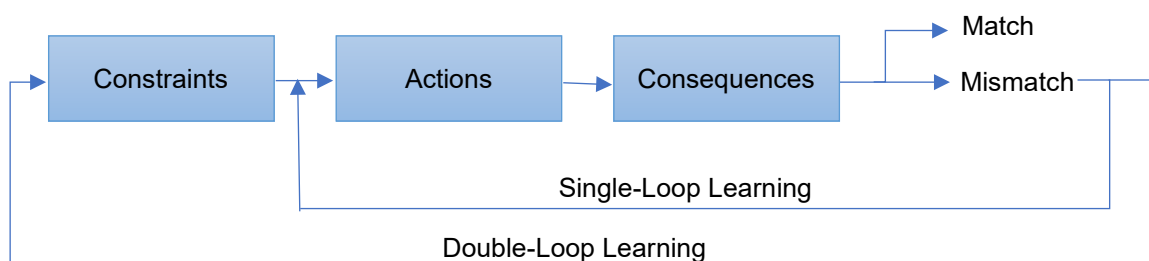
educational problem is to change those theories of action to produce consequences that are no longer judged to be problematic. (Robinson, 1993, p. vii)

Notably, a ToA can be espoused (“talk theories”) or in use (“walk theories”) (Robinson, 2018, p. 17). An *espoused* ToA refers to what people say they have done or intend to do; in contrast, a *theory in use* describes what people actually do (Argyris & Schön, 1974; Robinson & Lai, 2006). Sometimes, a person’s espoused theory can be incongruent with their theory in use, and many people are unaware of gaps between their espoused theories and theories in use (Anderson, 1997; Argyris & Schön, 1974).

Single-Loop and Double-Loop Learning. The improvability of a ToA depends on the quality of the theory that enables or hinders improvement, especially regarding the types of learning resulting from the feedback loop (Robinson, 1993). Two types of learning associated with the ToA concept are single-loop and double-loop learning (Argyris & Schön, 1996; Robinson, 2018). *Single-loop learning* refers to “instrumental learning that changes strategies of action or assumptions underlying strategies in ways that leave the values of a theory of action unchanged,” whereas *double-loop learning* involves changes in “the values and norms that govern their theories in use” (Argyris & Schön, 1996, pp. 20–21) (see Figure 2). On examining the participants’ theories of action regarding problem solving during faculty evaluation processes, my study would reveal if the participants’ approaches were associated with single- or double-loop learning. The participants’ approaches to problem solving during faculty evaluation would be associated with single-loop learning when they resorted to different actions or strategies to reduce the mismatch between intended and actual consequences. In contrast, the participants’ approaches would be linked to double-loop learning when they changed the beliefs and values that drove their subsequent actions.

Figure 2

Single- and Double-Loop Learning



Note. Reprinted from *Reduce Change to Increase Improvement* (p. 21), by V. M. Robinson, Corwin. Copyright 2018 by Corwin.

Rationale for Using PBM. PBM was the best methodology to help me understand the problems and the degree of improvability of faculty evaluation practices in Vietnamese higher

education for several reasons. First, PBM enabled me to expand my understanding of why faculty evaluation problems occurred, who was involved in the problem-solving process, the action patterns of those involved and how the problems could be solved (Robinson, 1992). Previous research into Vietnamese higher education faculty evaluation was very scant, and among those few available (e.g., Hung, 2013; K. D. Nguyen, 2000; N. Phan, 2014), most described the evaluation problems without emphasising who was involved in the problem-solving process and how the problems could be addressed. Thus, my adoption of PBM helped resolve a gap in the existing literature about faculty evaluation in Vietnamese higher education.

Second, my research goal was to provide implications for improving faculty evaluation practices, and PBM is a proven methodology for improving educational practices (e.g., Donald, 2013; Eastham, 2017; Finnerty, 2020; Meyer & Slater-Brown, 2020; Nock, 2017; Slater-Brown, 2016). Though most PBM studies are in New Zealand, using PBM to understand and improve educational practice in an Asian context or among Asian participants is promising (Hannah et al., 2019; Robinson & Lai, 1999; Svay, 2017). PBM allowed me to examine holistic and long-term problem solutions to higher education faculty evaluation in an Asian nation like Vietnam. Like the study by Robinson and Lai (1999), my study also challenges key features of practice culture from the perspectives of the Asian participants, who had a collectivist norm of sharing resources and were possibly unfamiliar with the Western notions of “authorship” and “plagiarism” (p. 194). Without challenging the practice complexity, the problem solution or practice improvement will only be partial and short-lived (Argyris et al., 1985).

Qualitative Case Study Design

The research employed a qualitative case study research design (Creswell, 2013; Merriam & Tisdell, 2016; Yin, 2018) to explore the participants’ theories of action about two cases of the SET and voting evaluation practices. I interviewed three groups of 42 faculty evaluation stakeholders (i.e., four QA officers, 18 administrators, and 20 faculty members) from seven public HEIs in Vietnam. Besides gaining insights into the participants’ theories of action from their accounts of practice, I explored the SET and voting policy documents that provided a more in-depth explanation regarding national and institutional policy constraints that influenced the participants’ approaches to the evaluation practices.

Although PBM can be compatible with quantitative and qualitative methods (Robinson & Lai, 2006), I found that a qualitative research design best matched my research aims. A *qualitative case study* is defined as an approach in which a researcher investigates “a real-life, contemporary bounded system (a case) or multiple bounded systems (cases)” through “detailed, in-depth data collection involving multiple sources of information” (Creswell, 2013, p. 97). My PBM research resonates with qualitative research because its “focus on discovery, insight, and understanding from the perspectives of these being studied offers the greatest

promise of making a difference in people's lives" (Merriam & Tisdell, 2016, p. 1). A qualitative case study allowed me to collect natural or real-world data and capture the complexity and the contextuality (Stake, 1995) of the selected faculty evaluation practices in the participating Vietnamese public HEIs. A qualitative case study design allowed me to generate "richly descriptive" PBM research findings (Merriam & Tisdell, p. 37), especially to gain insights into the participants' values, beliefs, or the reasoning processes that guide their actions (Robinson & Lai, 2006) in the real-life context.

Selecting cases for a study is challenging because a case can be defined and bounded in multiple ways, such as an individual person, an event, an entity, a phenomenon or a service identified by "spatial, temporal, and other explicit boundaries" (Yin, 2018, p. 31). My study used a *multiple case study approach* (Yin, 2018) to examine the faculty evaluation problems of practice from the two cases of SET and voting evaluation. The bounded cases were the SET and voting evaluation taken by three different groups of participants from seven public HEIs within the Grand University (pseudonym) in Vietnam in the pre-COVID period (before 2020). My study was an *embedded multiple case study* (Yin, 2018) as each case of SET and voting evaluation focused on several units of analysis at the same time. The *units of analysis* for the SET were theories of action of the QA officers, administrators, and faculty members. However, the units of analysis for the voting evaluation were theories of action of the administrators and faculty members only (the QA officers were not directly involved in the voting evaluation). The choice of multiple cases allowed me to generate in-depth findings with rich data from diverse perspectives (Merriam, 2009; Yin, 2018). The choice of seven HEIs under the same university system was to strengthen potential findings through *literal replication* logic, which means the study of different cases under the same conditions to examine if the study generated similar results (Yin, 2018). The SET and voting evaluation represented a global and a typically Vietnamese communist-oriented practice, respectively. Thus, the selection of these practices representing different contexts might contribute to the *theoretical replication* logic or the extent to which these two practices generate contrasting results (Yin, 2018).

Participants, Cases and Context

One important step of case study research involved identifying participants, determining cases and understanding the context of the selected cases (Creswell, 2013). My study involved analysing policy documents and individual interviews with 42 participants working at the seven member HEIs of the Grand University.

Participants

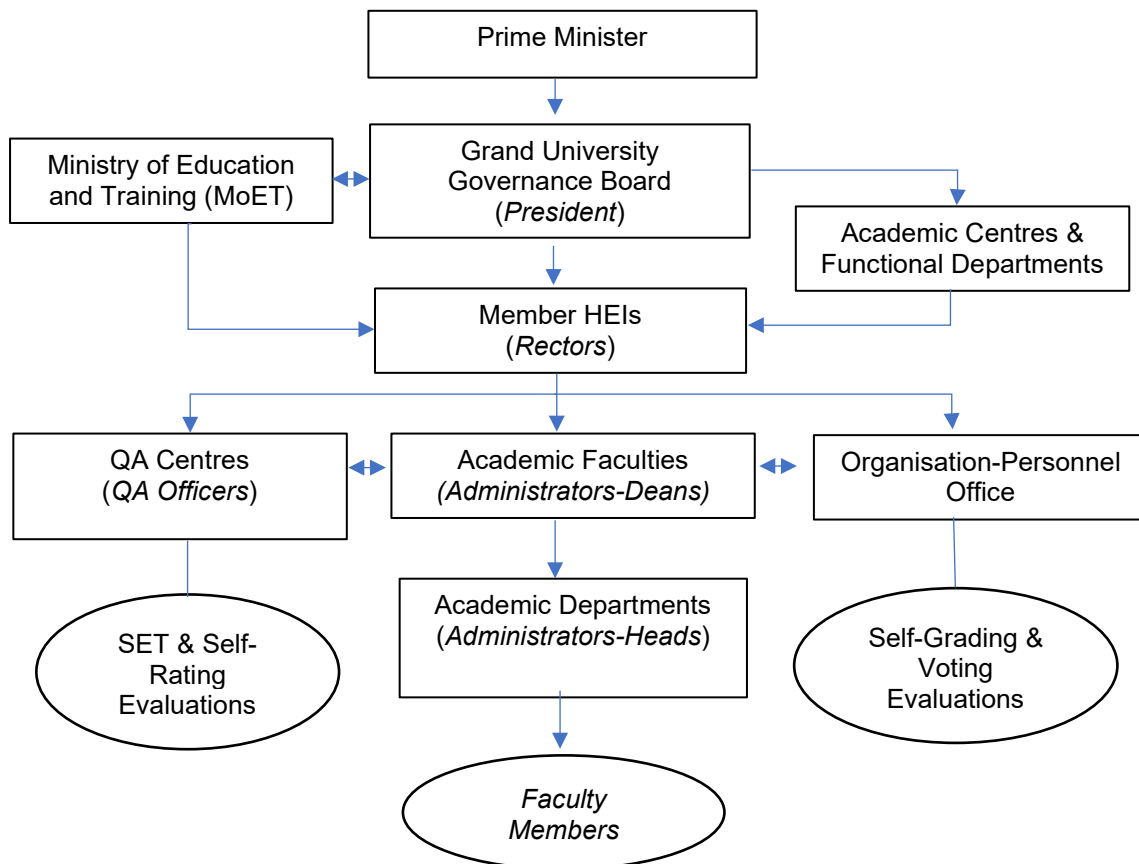
Selecting participants involved identifying the Vietnamese public HEIs and individual participants that optimised the understanding of the practice problems.

Participating HEIs. The problem identified in this study was the potential lack of stakeholders using faculty evaluation for learning and improvement purposes in Vietnamese public HEIs, even within the top-tier ones. To validate that assumption, I selected seven HEIs under the governance of both MoET and the Grand University, one of the exemplary public HEIs in Vietnam. As Figure 3 illustrates, the QA officers, administrators and faculty member participants were under several layers of governance.

At the Grand University level, a governance board, with the president being the chair of the board, includes representatives from Vietnam Communist Party members and rectors of member HEIs. The Grand University has multiple offices and functional departments to regulate policy documents to instruct the member HEIs to follow its goals. For example, the Grand University's Institute of QA staff directs members' HEI implementation of QA activities, including SET. The Grand University's organisation-personnel department staff oversee all the member HEIs' personnel-related activities, including voting evaluation.

At the member HEI level, each rector directs a member HEI, which has their juridical entity status, is entitled to use their seals and possess their institutional financial account. These HEIs consist of functional offices such as the QA centre, an organisation-personnel office that coordinates with academic faculties and departments to implement various faculty evaluation practices. For example, member institutions' centres for QA are responsible for organising and monitoring QA activities such as SET or faculty self-rating evaluations. The organisation-personnel officers assist the rectors in managing the institution's various organisation and human resource tasks, including coordinating annual voting evaluations.

At the faculty level, each faculty has one to three administrators who are deans (or vice-deans) who oversee all the academic activities and faculty performance management. A large faculty is divided into smaller departments in which a head of department (also referred to as an administrator in this study) works directly with faculty members to support the faculty deans with their academic activities and faculty performance management.

Figure 3*Member HEIs Structure and Governance*

Note. HEIs: higher education institutions, MoET: Ministry of Education and Training, QA: quality assurance, SET: student evaluation of teaching.

The participating member HEIs were selected using *purposeful sampling*, which means selecting cases that potentially provide deep understanding and rich insights into the practice under study (Merriam & Tisdell, 2016). I adopted *maximum variation*, meaning I purposefully selected seven member HEIs that differed in size, discipline, and period of establishment. This sampling allowed me to reflect on diverse perspectives from multiple cases (Creswell, 2013). There were variations in the types of institutions, with different faculty sizes or approximate numbers of faculty members, ranging from just over 40 to nearly 600. In terms of academic disciplines, recruited institutions offered programmes in various disciplines, from social science (e.g., language, history, law, business, or pedagogy) and natural sciences (e.g., informatics, mathematics)¹². Following the University of Auckland Human Participants Ethics Committee (UAHPEC) standards, I contacted the rectors from the participating HEIs to ask for their consent for me to undertake this research as outlined in the participant information

¹² To protect the anonymity of the participating HEIs, I do not provide a detailed description of these HEIs.

sheet (PIS). The rectors granted access to this research and assured that their staff's participation or nonparticipation would not affect their employment in any way.

Individual Participants. The sample consisted of 42 interview participants from 16 faculties under seven member HEIs of the Grand University (see Table 2). Among 42 participants, 20 were faculty members, 18 were administrators, and four were QA officers. These participants were relatively diverse in academic positions, genders, ages, and educational backgrounds.

Table 2

Participating HEIs and Individuals

Grand University member HEIs ¹³	Faculties	Research Participants			Total number
		Administrators	Faculty members	QA officers	
HEI 1	1	I4A1	0		11
	2	I1A2, I1A4	I1T1, I1T2, I1T4, I1T5, I1T6	I1Q1	
	3	I1A3	I1T3		
HEI 2	4	I2A1	I2T1, I2T2		9
	5	I2A2	0	I2Q1	
	6	I2A3	0		
	7	I2A4	I2T3, I2T4		
HEI 3	8	I3A1	I3T1, I3T2	I3Q1	7
	9	I3A2	I3T3, I3T4		
HEI 4	10	I4A1	0	I4Q1	12
	11	I4A2, I4A4	I4T4		
	12	I4A3	I4T1		
	13	I4A5, I4A6	I4T2, I4T3, I4T5		
HEI 5	14	I5A1	0	0	1
HEI 6	15	I6A1	0	0	1
HEI 7	16	0	I7T1	0	1
Total		18	20	4	42

Note. HEI: higher education institution. In a participant's name code, the letter *I* stands for an institution, *A* for administrator, and *T* for teaching faculty member. The assigned name code for each participant was based on the order of the participants' interviews. For example, Participant I2T3 was the third teaching faculty member interviewee from HEI 2.

To protect the anonymity of the HEIs, I do not provide the detailed background, disciplines, or personal information of the participants. Generally, participants included both females (n=27) and males (n=15). Their ages ranged from 25 to 64. Except for three participants who held bachelor's degrees, the rest had attained master's or doctoral degrees. Among those who had master's or doctoral degrees, over 40% (n=17) had attained their

¹³ Each member HEI will be called by its number from now on, e.g., HEI 1.

degrees overseas. The participants' disciplines varied, from social sciences such as linguistics, pedagogy, philosophy, history, tourism, and business to natural science such as mathematics and informatics. This diversity in participants' backgrounds allowed me to generate multiple perspectives on faculty evaluation practices.

The individual participants were selected using *snowball sampling*, which means I identified "the cases of interest from the people who know what cases are information-rich" (Creswell, 2013, p. 158). This sampling also involved identifying and interviewing a few key participants who met my listed sampling criteria before asking them to refer me to other participants (Merriam & Tisdell, 2016). While recruiting QA officers, administrators, and faculty members, I intentionally asked my interview participants to suggest new participants that matched the inclusion criteria outlined in the PIS. For example, the teaching faculty members were limited to those who

- held a full-time¹⁴ position,
- had been in the institution for more than 2 years, and
- had a recent experience (in the previous year) with institutional faculty evaluation practices.

The snowball sampling strategy (Creswell, 2013; Merriam & Tisdell, 2016) was helpful for me to gain access to participants when they knew me from a trusted network. Those who introduced new participants needed to ensure the new participants were entirely voluntary and their employment would not be affected in any way.

Cases

I selected the cases of SET and voting evaluation based on the interview participants' suggestions because I aimed to understand practices in ways that were most relevant to their perspectives. I used "exploratory" rather than "checking" questions during my interviews by asking the participants to identify the practices that were most relevant to them (Robinson & Lai, 2006, p. 146). I categorised various faculty evaluation practices based on their formality and relevance (see Figure 4). The participants identified both formal and informal evaluation practices as important to them. Examples of informal evaluation included administrators gathering information about a faculty member's performance through informal conversations with colleagues or faculty members collecting their student feedback about teaching by themselves. However, as informal faculty evaluations varied across the participating HEIs, I excluded them and selected only formal practices for investigation. The participants' interview

¹⁴ Full-time faculty members are those who have signed long-term contracts (including a contract of 1 year or more, and an indefinite-term contract). Faculty members include both those who directly teach and those who concurrently fulfil teaching and managerial duties (e.g., heads of department or faculty deans) (Grand University, 2014).

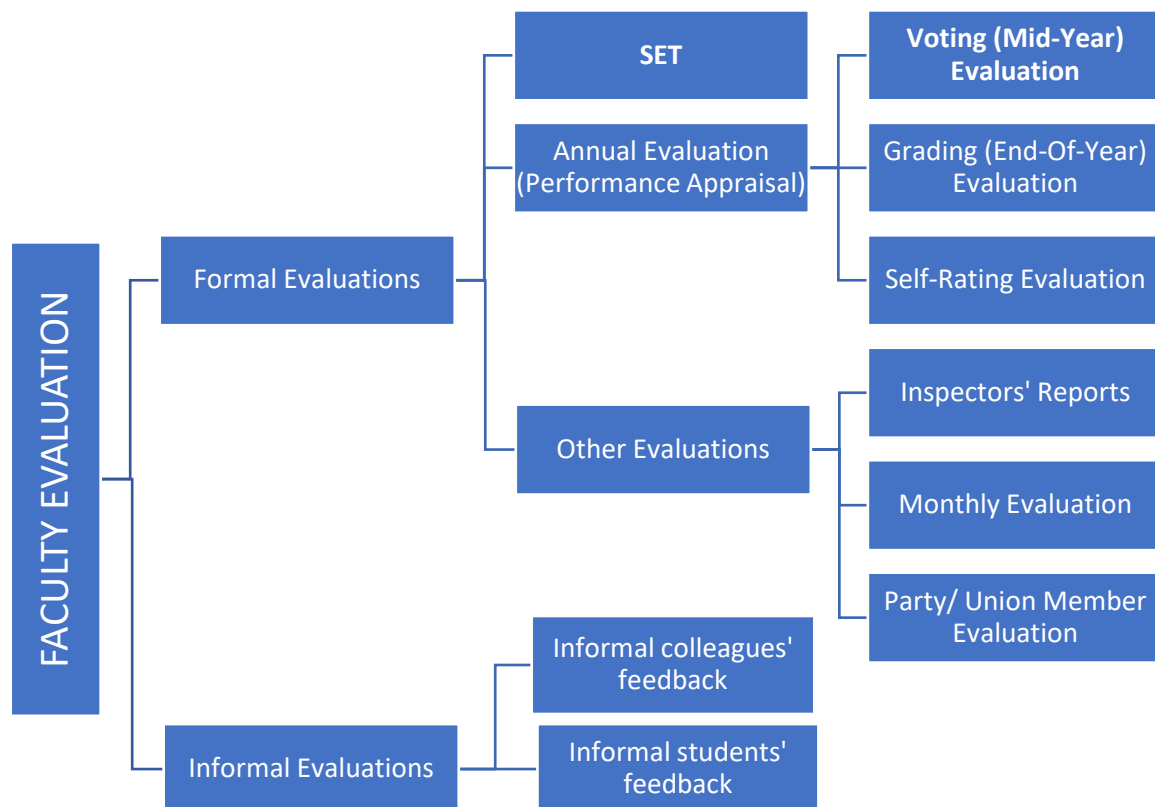
responses revealed five formal evaluation practices being regularly conducted at the participating HEIs:

- (i) SET
- (ii) Voting evaluation: Mid-year evaluation for emulation and commendation
- (iii) Classification evaluation: End-of-year classification of cadres, civil servants, and public employees¹⁵
- (iv) Self-rating evaluation¹⁶
- (v) Inspectors' reports¹⁷

I selected SET and voting evaluation as they were two practices most frequently mentioned by the participants.

Figure 4

Categories of Faculty Evaluation Practices



Note. SET: student evaluation of teaching.

¹⁵ Translated from the Vietnamese evaluation practice of “đánh giá, xếp loại chất lượng cán bộ, công chức, viên chức”. This practice involves individual faculty members filling in (usually end-of-year) self-grading reports and administrators approving the reports. The participants often referred to this practice as the “grading evaluation”.

¹⁶ Translated from the Vietnamese evaluation practice of “hoạt động giảng viên tự đánh giá”. This practice involves faculty members filling in self-rating evaluation forms sent down by QA officers. In this practice, administrators are not required to approve the individual faculty self-rating evaluation, and QA officers usually generate a summarised report basing on faculty provided forms.

¹⁷ Translated from the Vietnamese evaluation practice of “báo cáo thanh tra”. This practice involves inspectors board monitoring, recording, and reporting faculty members' rule-compliance or (non)disciplinary behaviours.

Policy Context

As policies are among key constraints that influenced the participants' actions (Robinson, 1993), I sought to understand the policy contexts of the selected cases of SET and voting evaluation. I analysed the key SET and voting evaluation policy documents at the national and institutional levels to determine the policy intentions or espoused theories of these practices.

SET Policy Context. The MoET's (2010, 2013) national policy documents on SET mandate all Vietnamese HEIs to conduct student evaluations of all faculty members within the HEIs, starting from 2010–2011 (section V). At the national level, SET is considered a separate practice that is intended to be used for summative, formative, and transformational purposes. It is stated in the MoET (2010) policy that HEIs are expected to collect SET to

- contribute to the implementation of the regulation on democracy in HEIs; build a team of faculty members with ethical qualities, professional conscience, qualifications, high expertise, advanced and modern teaching methods, and styles;
 - create more information channels to help faculty members adjust teaching activities; enhance the sense of responsibility of teachers in the implementation of training objectives of HEIs;
 - strengthen the sense of responsibility of learners with rights and obligations to study and practice on their own; create conditions for learners to reflect their thoughts, aspirations, and express opinions about the teaching activities of the lecturers.
- (section I)

In terms of accountability purposes, the SET national policy (MoET, 2010) highlights the “sense of responsibility” among faculty members and students. While this policy emphasises SET improvement purposes, such as supporting faculty members to adjust their teaching, it also introduces the SET transformational purposes of democratising HEIs. In the updated correspondence a few years later, the MoET (2013) emphasises the SET accountability purpose of ensuring accreditation and behavioural reinforcement:

Contributing to the quality accreditation of higher education institutions; helping managers in higher education institutions have more grounds to comment and evaluate lecturers; contributing to preventing negatives in teaching activities at higher education institutions, detecting and replicating good examples among the teaching staff (section I)

Following the national mandate, the Grand University (2010, 2014) issued policy guidelines that specify SET purposes, expected procedures and key agents' roles.

Institutional SET Purposes. At the institutional level (Grand University, 2010), SET is embedded in the course evaluation practice at the institutional level and focuses mainly on summative and formative purposes, but not the transformational ones. First, the institutional SET policy (Grand University, 2010) indicates SET summative purposes of making judgmental classifications of faculty members' teaching based on predetermined SET performance requirements. SET is expected first to be used for improvement and later for personnel decisions related to emulation titles. Second, the SET institutional policy (Grand University, 2010) also highlights that the formative purposes of SET are to help faculty members adjust teaching activities, build a team of faculty members with ethical qualities, and improve the quality of training. The SET policies expect the HEI leaders and administrators to use SET to plan faculty training and development. However, while the national SET policies indicate a transformational purpose of increasing students' participation in the Vietnamese HEIs' governance (MoET, 2010, section 1), none of the statements within the institutional SET policies reflects such democratic spirit nor the ownership of the administrators or faculty members in institutional SET design, collection, or utilisation. This absence reflects a gap in involving grassroots participants in the institutional SET processes compared to the national-level ones.

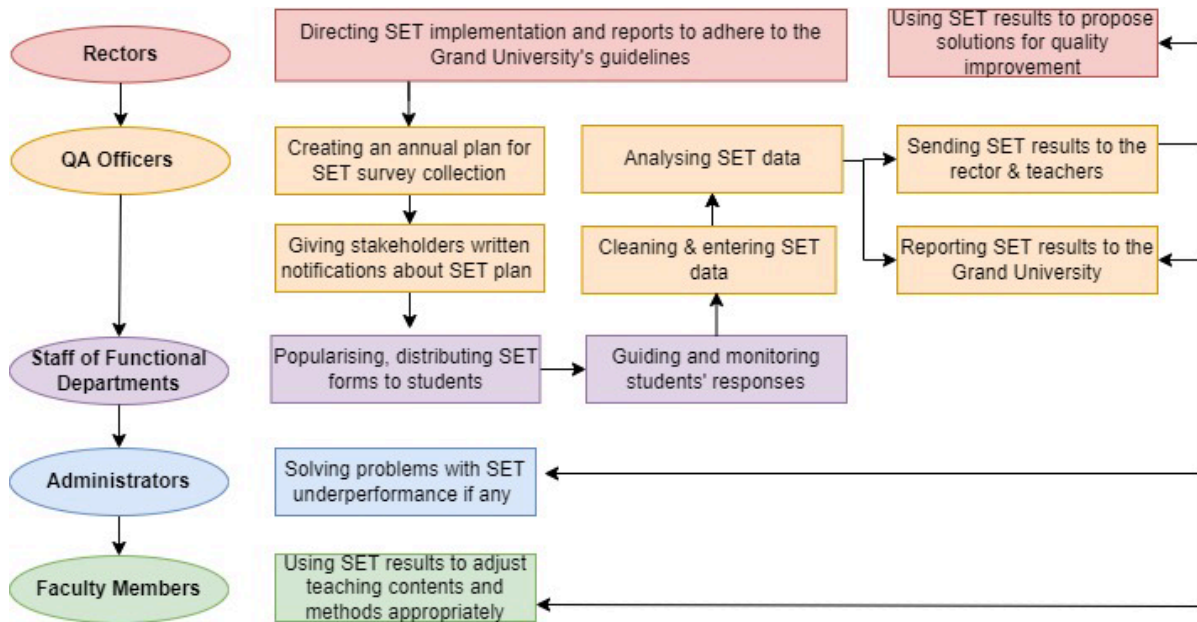
Institutional SET Procedures and Agents' Roles. The Grand University's (2014) SET policy document outlines the roles of various stakeholders at two levels. At the Grand University level, QA officers at its QA Institute must design the core content of SET survey forms and monitor SET implementation across member institutions. Following the SET plans, directions, and goals from the upper institutional level, the rectors of the member HEIs must direct the SET implementation across their institutions under the support of QA officers.

At the member HEIs' level, several expectations are set out in the policy about how each stakeholder uses SET results (see Figure 5). First, the member HEIs' QA officers must follow their rectors' directions to execute critical SET tasks such as planning analysing SET data. Other functional departments support QA officers with less important tasks like distributing or collecting SET forms. After analysing SET results, QA officers must report SET results to different stakeholders such as institution rectors, administrators, and faculty members. Second, administrators must solve problems if faculty members' SET results are under a particular benchmark or cut-off score (e.g., 3 out of 5) (Grand University, 2014, p. 26). Third, faculty members must use SET results to make relevant adjustments to their teaching. QA officers are also responsible for keeping SET results confidential and handling faculty members' feedback or complaints about SET practices (Grand University, 2014, pp. 4–5). HEIs' rectors, under the support of QA officers, must propose solutions for quality improvement based on SET results. Overall, the institutional SET policy requires various

stakeholders to contribute to planning, collecting, disseminating SET results, solving SET underperformance, and using SET to solve quality issues.

Figure 5

Expected SET Procedures at Member HEIs' Level



Note. Based on the Grand University's policy document (2014, ss. 2.2, 2.3). SET: student evaluation of teaching, QA: quality assurance, HEIs: higher education institutions.

Voting Evaluation Policy Context. The national policy documents that guide the voting evaluation are the *Law on Emulation and Commendation 2003* and the *Law on Emulation and Commendation Amendment 2013*. These emulation and commendation policy documents explicitly identify the overall meaning and purposes of the voting evaluation.

As stated in the *Law on Emulation and Commendation 2003*, *emulation* refers to “organised activities voluntarily participated by individuals and collectives to attain the best achievements in national construction and defence” (article 3). In the same section, *commendation* refers to “forms of recognition, praise and honouring of individuals and collectives that have recorded achievements in emulation movements”. The overall purpose of the emulation and commendation practice is to “create a motive to mobilise, attract and encourage all individuals and organisations to promote their patriotic tradition, dynamism and creativity in striving to accomplish well the assigned tasks for the objective of a prosperous people, a strong country, an equitable, democratic and civilised society” (article 5). In the *Law on Emulation and Commendation Amendment 2013*, the overall purpose of the emulation and commendation practice is still the same, with a slight change from “a strong country, an equitable, democratic and civilised society” towards “a strong, democratic, equitable and

civilised country” (article 5). This slight and hardly noticeable change reflects that the overall purpose of the practice remained unchanged after a decade.

Institutional Purposes of Voting Evaluation. The voting policies at national and institutional levels both emphasise a summative purpose (Buller, 2012; Scriven, 1991) of judging faculty members’ performance for personnel decision making and maximising institutional productivity. The institutional voting policy (Grand University, 2015) shares with the *Emulation and Commendation Law 2003* the purpose of discovering, nurturing, and multiplying outstanding achievers. Institutional voting involves individual faculty members and subunits of member institutions reporting on their and others’ strengths and weaknesses. Faculty members must attend whole faculty evaluation meetings where they share, comment on, and evaluate their own and others’ performance via votes of confidence. One key reward is “emulation titles,” which means the forms of recognition, praise and honouring of merits bestowed on the achievers of the emulation movements (Law on Emulation and Commendation 2003, article 3).

Institutional Voting Evaluation Procedures and Roles. The Grand University’s Decision on Emulation and Commendation (2015) indicates the voting evaluation procedures and roles. Following this policy document, all full-time faculty members (except those with 1-year or under 10-month contracts) must participate in this evaluation practice. Faculty members vote on emulation titles for themselves and others during the voting evaluation. Following the national law’s (2003) list, the emulation titles for considering the Grand University’s staff, including faculty members, are ranked from the lowest to the highest as follows:

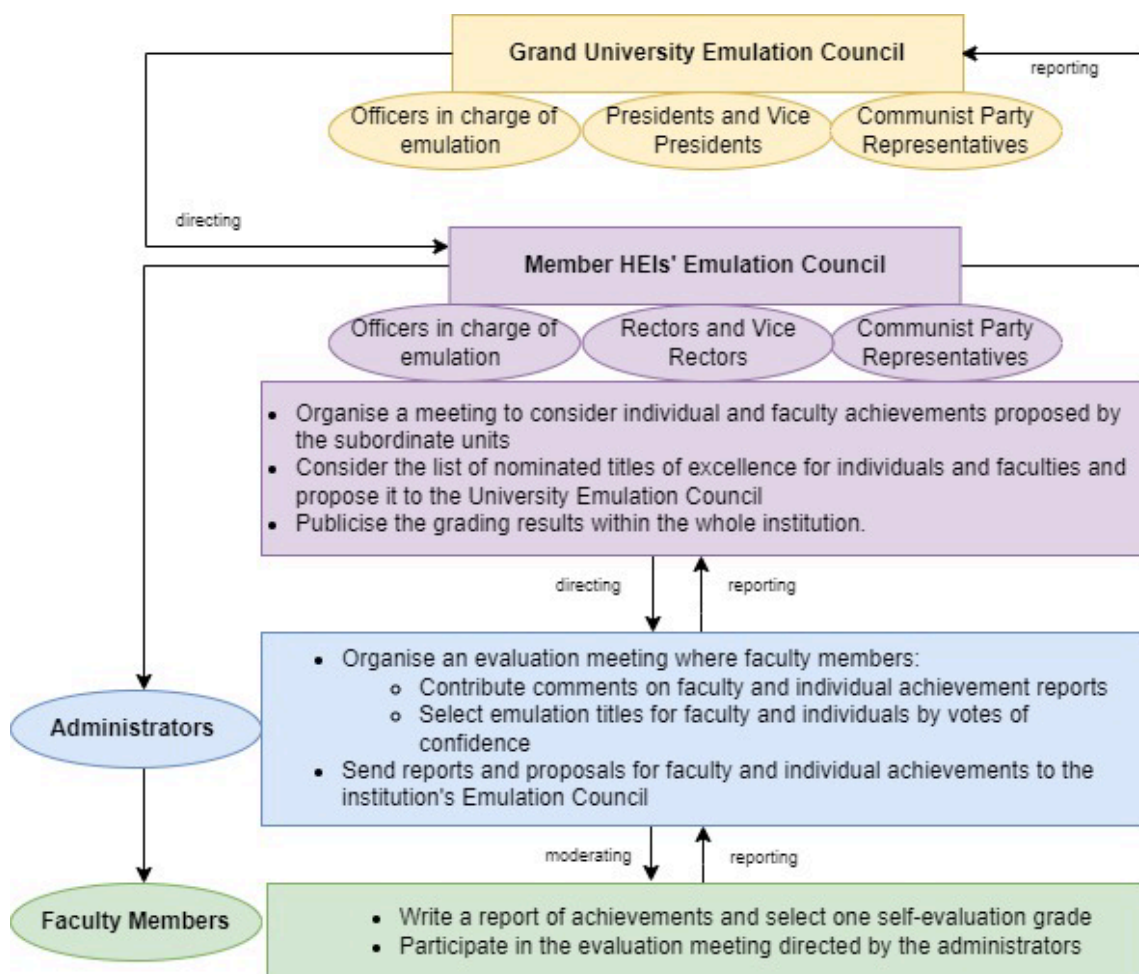
- (i) Advanced labourer, advanced fighter
- (ii) Emulation fighter of the grassroots level
- (iii) Emulation fighter of the university level
- (iv) Emulation fighter of the ministerial level
- (v) Emulation fighter of the national level
- (vi) Young face of excellence at the grassroots level
- (vii) Young face of excellence at the university level. (Grand University, 2015, article 10)

According to the Grand University (2015), the voting procedures for evaluating and considering honourable titles and rewards involve key stakeholders complying with the various steps (see Figure 6). At the Grand University level, an emulation council must launch annual emulation campaigns following the national policies (Law on Emulation and Commendation 2003; Amended Law on Emulation and Commendation Amendment 2013). Following these campaigns, the emulation council at the member HEIs must direct their voting implementation across their faculties and departments. Officers in charge of emulation must forward the voting

documents to administrators, who must then forward the documents to faculty members. Faculty members must report achievements and rate their performances based on given titles. Administrators must organise meetings in which all faculty members share their self-reports, comment, and select honourable titles for themselves, for other members and the whole faculty by a vote of confidence (usually in a secret ballot). The minimum requirement for individuals or collectives (e.g., departments, faculties) to be considered for the proposed titles in the next round is to get support from at least two-thirds of the voters. Then the member institutions' emulation council organises a meeting to consider honourable titles for individuals and collectives proposed by the faculties and departments. The member HEIs' emulation council must select and report the institutional list of honourable titles to an emulation council at the Grand University, which approves titles proposed by the emulation council at the member HEI level.

Figure 6

Procedures and Roles of Key Stakeholders in Voting Evaluation



Note. Adapted from Grand University's Decision on Emulation and Commendation (2015, articles 27, 28, 31). HEIs: higher education institutions.

Data Collection Methods

A qualitative case study usually involves collecting multiple data sources such as observations, interviews, documents and audio-visual materials (Creswell, 2013). Two data sources used in this study were semistructured interviews and documents. The data collection lasted 6 months, from September 2018 to February 2019.

Interviews

I conducted individual face-to-face interviews with 42 participants, each taking about 60 to 90 minutes at locations convenient for the participants. The interviews' purpose was to gain profound insight into each participant's theories of action concerning faculty evaluation, especially the explicit links between their actions, constraints, and consequences (Robinson & Lai, 2006). Individual interviews helped me attain rich and personalised information (Algozzine & Hancock, 2016) to answer the three key research questions based on the participants' understanding and interpretation of their faculty evaluation practices. Individual interviews enabled me to inquire into the participants' unobservable constraints (Merriam & Tisdell, 2016), such as their perceptions, beliefs, and values, that influenced their approaches to faculty evaluation. The interview questions were semistructured based on the PBM framework and learning conversation principles (Robinson & Lai, 2006). Interviews were conducted in Vietnamese. The interview questions were piloted using two mock interviews.

Semistructured PBM Interviews. All my interviews were *semistructured*, which means that I used "flexibly worded" and "a mix of more or less structured interview questions" (Merriam & Tisdell, 2016, p. 110). I employed PBM themes (i.e., actions, constraints, and consequences) to guide my interview and discussion with the participants. However, I did not strictly follow a rigid list, but I allowed my questions to flow flexibly along with my interviews (Travers, 2019). The use of semistructured interviews enabled me to gather information about the participants' theories of action as intended while simultaneously responding to the interviewees' "emerging worldview" (Merriam & Tisdell, 2016, p. 111).

Each interview started with a general introduction, followed by two main phases that aimed to inquire into the interviewees' theories of action about faculty evaluation (see Appendix A for the Interview Protocol). In the introduction, I explained the purposes and procedure of the interview, reviewed the PIS and clarified ethics with the participants. I made sure that the participants were aware that their reflections on and experiences with the faculty evaluation practices were valuable to my research. I also ensured that the participants were entirely voluntary and they were comfortable before the main interviews took place. In the main interview, I started by familiarising myself with the participants, asking them a few questions about their roles, responsibilities, and their experiences with various faculty evaluation practices (as evaluators or evaluatees). These questions provided me with some

background information that I could use at a later stage to probe more into the interview participants' experience with the evaluation practices.

Next, I elicited the individual participants' theories of action by asking questions that were related to the three topics of actions, constraints, and consequences. To explore the participants' approaches to evaluation, I probed their actions, including their responses to evaluation requirements, their engagement in the evaluation processes and their uses of evaluation. I asked the interviewees to describe their most recent faculty evaluation where they participated as an evaluator or evaluatee. I probed for constraints influencing their approaches to faculty evaluation. For example, I asked the participants questions like "why did you ...?" and "what influenced your decision to ...?" These constraints were the goals, beliefs, attitudes, values, norms, or conditions that promoted or hindered faculty evaluation actions. I also probed for the evaluation policies that drove the participants' faculty evaluation practices. I noted down the policies that the participants referred to, and I also asked if they could kindly provide related documents if possible. Lastly, I elicited the participants' perception of the impact of their approaches to faculty evaluation at both individual and institutional levels. These impacts were those on institutional personnel decisions, faculty members' improvement in professional learning and teaching, and their overall satisfaction.

Learning Conversation Principles. I drew on *learning conversation* principles which mean that I did not treat myself as "someone to be won over," but I contributed to the process of "describing, explaining, and evaluating" the participants' theories of action (Robinson & Lai, 2006, p. 42). During the interviews, I constantly checked the accuracy and adequacy of my own assumptions with the participants (Robinson & Lai, 2006). After eliciting the participants' theories of actions, I disclosed my understanding of the participants' responses and inquired into the participants' viewpoints (Robinson & Lai, 2006). For example, I frequently summarised and checked the accuracy of my own understanding of the participants' statements by asking questions such as "you have identified several evaluation practices that are common ... Have I understood you correctly?" I also checked possible coherence issues in the interviewees' accounts of practice, for example, by asking for explanations of differences between the stated policies and purposes and actual implementation: "It is stated in the guidelines about ... that ... But you mentioned that ..." During the interviews, I also applied other interview techniques such as critiquing the participants' accounts of practice by inviting them to explain the basis of their opinion or to consider alternative ways of explaining their practices (see Appendix A).

The adoption of the learning conversation (Robinson & Lai, 2006) in this study was beneficial for several reasons. First, unlike in traditional interviews, I did not withhold my assumptions about the participants' key points. Instead, I explicitly summarised, checked, and invited them to reflect on what they said and how I understood their account. The learning conversation enabled me to concurrently analyse data while interviewing through the process

of going “back and forth between thinking about the existing data and generating strategies for collecting new, often, better data,” of recognising and making necessary changes to my “blind spots” or potential subjectivity (Miles et al., 2014, p. 70). The learning conversation interview enabled me to understand the participants’ accounts of practice while having evidence to confirm or disconfirm the established causal links among these PBM components during the later data analysis stage. Second, by checking and probing my assumption, I provided the participants with an opportunity to reflect on their daily practice critically. One of the participants shared that these academics did not usually engage in a formal conversation about the evaluation practice as they usually took it for granted. To catalyse the participants’ engagement, I occasionally presented other viewpoints or scenarios different from theirs and invited their reflections, for example, by asking “what if...,” “other participants found... how do you react to this view?” Hence, the participants may have gained more knowledge, insight, and understanding of their theories of action regarding faculty evaluation practices by actively comparing and contrasting what they thought and otherwise.

Pilot Interviews. Before the real interviews took place, I conducted pilot interviews with two Vietnamese doctoral students who had been working as faculty members in Vietnamese public HEIs. Piloting an interview instrument allowed me to test the interview questions and add, modify, or refine questions in ways that yield more valuable data (Merriam & Tisdell, 2016). My study targeted participants from Vietnamese public HEIs, so I piloted my interview questions with two faculty members from two public HEIs in two different regions in Vietnam (i.e., Hanoi and Hue cities). As HEIs in Vietnam are under the same centralised governance of the Vietnamese government and the MoET, faculty evaluation practices in public HEIs in Vietnam share the same national policies as the Grand University. The two selected pilot interviewees were doctoral candidates, so they were able to provide me with “constructive feedback on the strengths and weaknesses” of my interview instrument (Robinson & Lai, 2006, p. 140).

After the two mock interviews, I made two modifications. First, I reworded my questions to make them more conversational and friendly. Second, I added reminder notes to ask more follow-up questions to elicit information about their thoughts or feelings about a specific experience during the faculty evaluation processes. I also learned from these mock interviews that there were many different faculty evaluation practices at their institutions during a year, besides profession-related faculty evaluations. Examples of nonprofessional assessments include Vietnamese Communist Party member evaluation, trade union evaluations, or voting for the title of “excellence at work and perfection at home” (giỏi việc nước đảm việc nhà)¹⁸ awarded only to female faculty members. As this study focused on examining the impact of

¹⁸ See, for example, from <https://giaoducthoidai.vn/ket-noi/phan-dau-90-nu-nha-giao-dat-danh-hieu-gioi-viec-truong-dam-viec-nha-lai.html>

evaluation on faculty members' professional learning and improvement, I decided to narrow down the question scope to focus on profession-related faculty-evaluation methods.

Documents

Documents refer to various public and nonpublic written materials (Merriam & Tisdell, 2016), such as the SET and voting evaluation policies, forms, and reports used in this study setting. When I started analysing the constraints, there were implicit references to evaluation policies that shaped the participants' actions. Thus, I selected the institutional evaluation policies that participants referred to as explanatory for their practice. I also selected the national policy documents which explain the identified ones at the institutional level. For example, the voting policy at the national level (Law on Emulation and Commendation 2003) was identified as the basis for the institutional voting evaluation policy expectations, requirements, and processes.

The public documents used in the study were sorted into three levels: national (including the National Assembly and MoET), the Grand University and member-HEI levels (see Table 3). I examined the national evaluation policies to understand the original policy intentions and expectations, whereas those at the institutional levels provided me with the stated purposes, criteria, and processes that guided the participants' SET and voting evaluation implementation. The documents at the member-HEI level were another source of evidence about the participants' practices. Besides the public documents, I also collected the SET and voting evaluation forms and reports provided by the participants as actual evidence of their evaluation practices (see Appendix B for a sample of an actual SET report). These selected documents helped me answer the second research question that aimed to explore the policy constraints of the participants' approaches to faculty. Documents were also used to triangulate or check the accuracy of claims made by the participants. For example, the interview data revealed some participants' limited engagement in using SET for faculty learning, and the in-depth analysis of the national SET policy document (2010) confirmed that the policy ambiguity was one key constraint. The inclusion of policy document analysis also enabled me to identify gaps between faculty evaluation policymaking and policy implementation and whether the participants' actions helped the intended purposes of the national and institutional evaluation policies.

Table 3*Documents Used in the Study*

No	Document name	Issued by	Reasons for sampling the data
National Policy Documents			
1	Law on emulation and commendation 2003 (number 15/2003/QH11)	National Assembly, 2003	To examine the original policy intentions and expectations for the implementation of the SET and voting evaluation
2	Law on emulation and commendation amendment 2013 (number 39/2013/QH13)	National Assembly, 2013	
3	Correspondence on collecting student evaluation of teaching (number 2754/BGDĐT-NGCBQLGD)	MoET, 2010	
4	Correspondence on organising the collection of student evaluation of teaching (number 7324/BGDĐT-NGCBQLGD)	MoET, 2013	
Grand University Policy Documents			
5	Guidelines on organising the SET collection	Grand University, 2010	To identify the stated purposes, criteria, and processes for the SET and voting evaluation at the institutional level
6	Guidelines on evaluating quality by collecting stakeholders' feedback	Grand University, 2014	
7	Decision on emulation and commendation	Grand University, 2015	
Member HEI Policy Documents			
8	Plan of SET implementation for semester II in the school year 2011–2012	HEI 3, 2012	To add and cross-check the participants' interview data regarding the implementation of the SET and voting evaluation at the member-HEI level.
9	Guidelines on implementing emulation and commendation practice	HEI 2, 2015	
10	Guidelines on summarising emulation and commendation practice in the school year 2017–2018	HEI 1, 2017	
11	Guidelines on implementing emulation and commendation practice in the school year 2017–2018	HEI 4, 2018	

Note. MoET: Ministry of Education and Training, SET: student evaluation of teaching, HEI: higher education institution. The National Assembly is the national legislature of the Socialist Republic of Vietnam (Quốc hội)

Data Analysis Approach

My data analysis was based on the PBM framework to construct the theories of action that explain the SET and voting evaluation at the participating HEIs. The data analysis started with familiarising with data, and then by constructing individual and group theories of action (Robinson & Lai 2006) (see Figure 7).

Figure 7

Data Analysis Approach



Note. ToA: theory of action

Data Familiarisation

I started my data analysis by familiarising myself with the data. This process helped me plan my later stages of constructing individual and group theories of action more carefully. The data familiarisation involved several steps as follows.

Transcribing and Sorting Information. I started the data analysis by conducting verbatim transcription of the participants' recorded interviews, which provided with me "the best database for analysis" (Merriam & Tisdell, 2016, p. 131). My self-transcription was time-consuming; however, it was helpful for me to get familiar with my data. After completing the transcription, I reread each interview transcript to "get a sense of the whole" and to "focus on the relationships" in the data (Robinson & Lai, 2006, p. 143) to make relevant decisions regarding my data analysis. For example, I found that the participants, especially faculty members, shared the two most important formal practices of SET and voting evaluation. Thus, I decided to focus on analysing the theories of action related to these two practices. I also decided to group individual participants by their roles (i.e., QA officers, administrators, and faculty members) rather than by institutions. I found the participants with the same roles had more commonalities in their tentative approaches to the SET and voting evaluation due to their shared responsibilities.

After gaining an overall sense of the data, I noted down potential codes for ToA components and my analytic memos of what I thought about the data. Those notes included my initial remarks, comments, reflection, or a reminder to myself to check again if my assumption about the participant's belief was correct. These memos helped reinforce my coding by directing my subsequent analysis to the matters that "deserve analytic attention" (Miles et al., 2014, p. 94). They were also valuable for me to revisit and connect them with data within or across the participants' interviews. For example, I noted that an administrator used the phrase "to improve a little bit ... bring a bit more educational values to students" to emphasise his effort to use SET for improvement purposes. Also, some notes involved interpreting the meaning of certain words uttered by the participants. For example, when a participant mentioned that he found SET was a bit "hình thức," I came up with the English equivalence of "formalistic," which means it is more concerned with formal rules than practical matters¹⁹. As I was not sure if it was the right word, I noted it down to remind myself to double-check its English equivalence at the later stage of data analysis and thesis write-up.

As I finished reading each individual transcript, I noted down a short narrative description of the participant's key information, such as their teaching experience or central beliefs about the SET and voting evaluation practices. These summarised descriptions were helpful for my reference during the later stages of individual and group theories of action. Below is an example of an participant's summarised background and account of SET practice:

Participant I3T3 is a female faculty member specialising in business. She was a student at HEI 3 before being recruited as a lecturer there. She considered research and teaching were two equally important roles of a faculty member. She considered students' learning autonomy as the most important determinant of their learning success, and a teacher's job was to inspire their autonomous learning. She was confident with her pedagogical capacity and her understanding of students' psychology. Although her SET scores had always been high, she considered student learning more important than SET (e.g., "students' opinions do not matter as much as their actual learning based on my own observation of their interactions in classes"). She only regarded SET as an additional source of information (e.g., "I don't really care [about SET] if it is above the average level. I just have a look at the paper and leave it there, throw it away"). She denied modifying her teaching to meet students' preferences.

Analysing Policy Documents. I analysed the SET and voting policies document to check if they reflected typical categorisations of evaluation purposes: summative, formative

¹⁹ Source of the English equivalence meaning: <https://www.thefreedictionary.com/formalistic>

and transformational (Buller, 2012; Patton, 1996; Scriven, 1991). The policy document analysis helped me understand the espoused theories of the sampled practices. The institutional evaluation policies were compared with those at the national levels. For example, when comparing the institutional SET (Grand University, 2010, 2014) with those at the national level (MoET, 2010, 2013), I found that the national SET policies indicate summative, formative and transformational purposes while the institutional ones only focus on the first two purposes (see the Policy Context section). The evaluation of policymakers' espoused theories was also compared with the participants' theories of action to identify if the participants' approaches to the SET and voting evaluation matched the policy intentions. These policies were also used to cross-check the participants' accounts of practices to validate and refine the participants' theories of action. For instance, during the analytical description of the participants' theories of action, their use of the SET and voting was checked against the accountability, improvement, and learning orientations, which correspond with the summative, formative, and transformational purposes.

Translating. I adopted *late translation* (Cormier, 2018), which means I analysed data in its Vietnamese origin and only translated the excerpts necessary for my ToA description and write-up. This late translation helped retain the originality of the participants' intended meanings, and it gave more voices and power to the minority participants and languages (Temple & Young, 2004) like those in my study. I also chose to do the translation myself instead of hiring an interpreter for several reasons. First, I considered myself a "linguistic insider researcher," as I am a native speaker and also a previous "insider" in the Vietnamese public HEI system, so I am better able to understand the nuances of the context which may potentially go unnoticed by outside researchers or interpreters (Cormier, 2018, p. 329). As a previous master's student of English and a faculty member of Vietnamese HEIs for nearly 10 years, I have the background knowledge to understand and translate the meaning conveyed by the local participants in my study. For example, one of the QA officers used the phrase "chỉ là một anh gác đền thôi" (literally translated as "only the guard of a temple") to describe her perceived roles and responsibilities. The word "temple" culturally refers to a sacred or holy place where Vietnamese people worship the gods or important figures, and the word "guard" means watching and protecting the sacred place. Taken together, the participant expressed that it was crucial for her to support her HEIs to satisfy all standards and regulations but her role was to monitor but not necessarily interfere in the process of quality improvement. Without a proper understanding of the culture and the language in the context, it would not have been easy for me to uncover the nuances in the participants' responses regarding their perceptions and experience of the practices.

One challenge during my translation was finding the equivalence or similar expressions from Vietnamese to English. To overcome the dilemmas of unavailability of

equivalence in Vietnamese and English, I took a dynamic perspective that sees translation not as something fixed or universally right but as a two-way conversation between the original text and the target language (Sutrisno et al., 2014). I paid attention to both *conceptual equivalence* – “the comparability of concepts or ideas between two languages” and *lexical equivalence* – “exact similarity of lexical meaning across languages” during my translation (Sutrisno et al., 2014, p. 1339). For example, one participant mentioned her tendency to be “cả nể”, which influenced her to prioritise other colleagues’ wishes over hers. The word “cả nể” can be translated into English as “compliant”²⁰ or willing to do what other people want. However, I found that the choice of “compliant” is inadequate as it does not show the manner of the compliant action. From the participant’s account, her “cả nể” tendency was linked to one common habit of the Vietnamese people to be unconditionally submissive to the common expectation of their groups or community. This habit caused the participant to perceive the need to please other colleagues by withdrawing her name from the nomination list to give more opportunities for others to be voted for outstanding titles. After considering the word “cả nể” in the cultural context and based on one of the participant’s additional explanations, I chose the word “people-pleasing” and “people-pleaser”²¹ to denote the behaviour or the person who tries to please others at the expense of their own needs. As the participant commented, a foreign researcher would not understand that this was not her personal choice but “how the [Vietnamese] system” worked. I believed that by linking “people-pleasing” with “cả nể,” I equally addressed both the need for conceptual and lexical equivalence (Sutrisno et al., 2014) while maintaining the meaning intended by the participant.

I kept a record of my translation process in an Excel file containing two Vietnamese and English equivalence columns, the uncertain terms, and the sources I used to make my translation decisions. Although back-translation or parallel translation, which involves two or more translators (Sutrisno et al., 2014), is preferred, I could only carry out a single translation, primarily by myself, due to time constraints and the difficulty of hiring a professional and relevant interpreter. To increase the accuracy of my translated text, I sent some of my translated excerpts (see Appendix C) to some of the participants and some other bilingual academics for feedback and suggested revision (Cormier, 2018; Li, 2011). Positing myself in the nonpositivist worldview, I did not see myself or other translators as “objective instruments” but as those subjectively influenced by our relative positions and familiarity with the research contexts (Temple & Young, 2004, p. 163). Therefore, I am open to meaning negotiation or other ways of interpreting the same instances. I adopted the “bracketing” technique (Shklarov,

²⁰ Sources: <https://en.glosbe.com/vi/en/c%E1%BA%A3%20n%E1%BB%83>, and <https://dictionary.cambridge.org/dictionary/english/compliant>

²¹ Source: <https://www.merriam-webster.com/dictionary/people%20pleaser#:~:text=Brett%20Stewart%20and%20Corey%20Irwin.and%20enjoys%20making%20others%20happy.%20%E2%80%94>

2007, p. 535) in which I included original Vietnamese terms in brackets to signal that these words might be open to alternative negotiation of meaning.

Individual ToA Construction

The construction of individual theories of action involved analysing each interview of the four QA officers, 18 administrators, and 21 faculty members. The ToA construction of all the participants used an identical process of first describing the actions from the participants' interviews regarding their approaches to the SET and voting. Then I worked backwards in both the interviews and evaluation policy documents to see the constraints and the consequences of these actions. The individual ToA construction was based on the key analytic steps (see Table 4) of first identifying actions, then the constraints and consequences that go with the actions. The overall coding rule was to identify segments of the interview transcripts or the SET and voting policy documents that have any reference related to the participants' actions, constraints and consequences, as described in Table 4. I coded these ToA components with the aid of NVivo 12 – a computer software that enabled me to systematically label and summarise the quotes in the interviews and documents that match with the coding rules (see Table 4). This coding process supported my iterative process of analysing all possible ToA components and seeking evidence to confirm and disconfirm the ToA component linkage to reach the most credible data interpretation. The following sections will present further steps to analyse each ToA component and use NVivo to support the individual ToA construction.

Analysing Actions. My individual ToA construction started by analysing actions, the “anchor” points for the later search for the corresponding constraints and consequences (Robinson & Lai, 2006, p. 145). I identify all extracts related to what the participants chose to do (*actions*) or not to do (*inactions*) (Robinson & Lai, 2006) during the SET and voting evaluation processes. The participants' actions and inactions constitute their approaches to the SET and voting evaluation, defined as how they engaged with the evaluation practices by responding to these evaluation policies or interacting with other agents. For example, during the action analysis, I found that an administrator made general comments on faculty performance in response to the voting policy expectation of his role as a moderator of the voting evaluation meeting. In contrast, a faculty member chose not to follow her students' request for her slides because she thought it was not beneficial for the student learning (see Table 4).

Table 4
Key Analytic Steps

Analytic steps	Coding rule	Example quotes in interviews or policy documents that match with the coding rule	Illustrations of how quotes are summarised
<i>Identifying actions:</i> The participants' approaches to the SET and voting evaluation	What the participants said or did	"I only stated general feedback in meetings with the whole faculty."	Making general comments on faculty performance during voting meetings
	What the participants did not say or do	"Some of my students complained that I should send them the slides, but I would not do it."	Not following students' feedback
<i>Identifying constraints:</i> What explained the participants' approaches to the SET and voting evaluation	SET and voting evaluation policy intentions and requirements	"If a faculty member's SET score fell below 3 (out of 5) ... forwarding those results to his/her direct supervisors."	A SET policy requirement about notification of faculty members' SET underperformance
	Institutional conditions, resources, and capacity	"Because the Ministry of Education and Training control the teaching content so our institution and faculty members cannot modify it."	Limited autonomy that hinders teaching and learning improvement
	Individual beliefs, perceptions, assumptions, and attitudes	"It [SET] is generally a means to observe [faculty teaching], and if anything unusual happens, we will solve it."	Perceived purpose of SET as monitoring and solving faculty teaching problems
	Individual values	"When we have such [SET] evaluation criteria, we can use them to review our teaching."	Appreciation of SET criteria
	Institutional and faculty norms	"Foreigners would not know how the system worked like this, but we tend to be 'cả nể' [people-pleasing]"	People-pleasing or harmony-oriented norm
<i>Identifying consequences:</i> The impact of the participants' approaches	Intended consequences	"We know that we will always be judged ... watched, and scrutinised, so we have a better sense of self-discipline."	Increased sense of responsibility
	Unintended consequences	"I do not know which direction to go... It does not help me improve my teaching."	Limited impact on faculty teaching improvement

Note. ToA: theory of action. SET: student evaluation of teaching. The description of ToA components adapted from *Practitioner Research for Educators: A Guide to Improving Classrooms and Schools*, pp. 21–26, by V. M. Robinson and M. K. Lai, 2006, Corwin Press. Copyright 2006 by Corwin Press.

Analysing Constraints. The ToA construction continued with the constraint analysis, which involved looking for all the possible constraints, such as institutional conditions or individual beliefs, that could explain specific actions taken by the participants. The goal of the analysis was thus to attain “a consistent and logical explanation of particular actions” (Robinson & Lai, 2006, p. 156). For example, the participants’ dominant compliance with the SET policy requirements was explained by the SET policy requirements, their limited autonomy and perceived purpose of SET monitoring and solving faculty teaching problems. I identified the potential constraints by checking the possible sources of participants’ own explanations, other participants’ accounts of practice, and the documents that participants identified as influencing their constraints (Robinson & Lai, 2006). As one key set of constraints includes “the goals, beliefs, attitudes and values” (Robinson & Lai, 2006, p. 22) of the implementing agents, I deliberately looked for phrases that denote what the participant may “find,” “think,” “feel,” “value” or “believe.” I also examined key SET and voting policy documents referred to by the participants or those specifying the evaluation purposes and stakeholders’ expected roles as they are “legal, regulatory and resource issues” that may have influenced the participants’ actions (Robinson & Lai, 2006, p. 22). For example, my search throughout the 26-page voting policy document of the Grand University (2015) reveals 218 instances of the terms “thi đua” (emulation) and 199 words related to “khen thưởng” (compliment and reward). Notably, among these 199 words, 174 words directly refer to the “compliment and reward,” 17 words related to complimentary certificates, and eight words were associated with proposing staff actions to be complimented (e.g., ‘khen’, ‘cấp trình khen’). The emphasis on compliments and rewards reflects the policymakers’ value of using external motivation to promote faculty performance.

During the constraint analysis, I read and reread the participants’ interviews and related policy documents multiple times, paying closer attention to those revealing explanations for actions and trying to keep a low level of inference as much as possible (Sinnema et al., 2021). I used an iterative process for determining constraints that could explain the actions by checking back and forth between the reported constraints and information gained from other sources (Robinson & Lai, 2006), such as accounts of participants from other groups or other HEIs, or policy documents. For example, in the example of constructed individual theories of action (see Appendix D), Participant I2A1 expressed his belief that evaluation should be a means of motivating faculty members. However, he was constrained by a lack of an institutional mechanism for acting on his belief. The participant’s value of evaluation to motivate faculty members was first treated as an espoused theory until I continued coding and checking his account of practice. At the subsequent data analysis, I found two confirming pieces of evidence for the participant’s espoused theory. First, the voting policy indicates that only a limited number of candidates

would be selected for outstanding titles and rewards (Grand University, 2015). Second, Administrator I2Q1 initiated a faculty-level award of 200,000 VND (nearly 10 USD) for a publication to motivate his faculty members.

Analysing Consequences. My consequence analysis involved identifying both intended and unintended consequences resulting from the participants' actions (Robinson & Lai, 2006). I extensively searched the participants' interview responses for the impacts of their approaches on individuals and institutions, especially on faculty members' learning and improvement. I identified the intended *consequences* of the participants' approaches by checking if their use of the SET and voting evaluation resulted in increased accountability, learning, and improvement. For example, the participants' comments on using evaluation to satisfy their administrative duties, managerial roles or requirements can be linked to the potential accountability impact of the faculty evaluation practices. I also looked for other consequences that the participants themselves identified using keywords such as "impact," "influence," "consequence," "make," "use," "improve," or "decrease." Besides the participants' use of the SET and voting evaluation, as intended by the evaluation policies, the participants also mentioned various feelings and behaviours resulting from their engagement in the evaluation practices. For example, I found one participant expressed her discomfort whenever attending her faculty evaluation meetings, even when she was recognised with outstanding titles. I categorised this as unintended evaluation consequences as the evaluation policies did not explicitly intend them. One important process during consequence analysis involved establishing and checking the causal link between the actions and the reported consequences. Following the suggestion by Robinson and Lai (2006), I read and reread the data to confirm or disconfirm a given consequence by checking if it was more or less likely than other possibilities.

Overall, I constantly constructed and restructured the individual theories of action until a satisfactory ToA table was created (see Appendix D for an example of an individual ToA table in its first iteration). The ToA analysis also involved a peer review which means having my two supervisors and other doctoral students access some of the raw data and my preliminary interpretation to determine the plausibility of my constructed theories.

Using NVivo 12. My individual ToA construction was supported by NVivo 12, a computer-assisted qualitative data analysis software that allowed me to systematically code and retrieve data (Merriam & Tisdell, 2016). I imported all the interview and document data into NVivo 12 software to analyse the participants' individual theories of action based on the coding rules (see Table 4). NVivo 12 enabled me to code the ToA components by assigning a relevant action, constraint or consequence to each data piece following the participants' interview transcript sequence and related policy documents in one single platform. I condensed codes of similar characteristics into the same categories under the predetermined

PBM components. For example, faculty members' concerns about lenient teaching and learning and the failure to recognise dedicated teachers were identified under the same category of faculty members' limited motivation to use SET within the unintended consequence component (see Figure 8).

Figure 8

Identifying Categories of Consequence Component in NVivo 12

Name	Files	References
SET CONSEQUENCES ON FACULTY MEMBERS	14	32
INTENDED consequences	8	13
Impact on faculty members' ACTIONS	2	2
Positive impact on teaching self-reflections and self-adjustment	2	2
Impact on faculty members' MOTIVATION	6	11
A low weight of SET used for faculty evaluation is OK	1	1
Higher SET results brought a sense of ACCOMPLISHMENT	1	1
Motivate teachers to STRIVE	1	1
Low SET scores force teachers to be MINDFUL about teaching	2	4
SET combined with reward having some motivational impact	2	4
UNINTENDED consequences	11	19
Faculty members' limited MOTIVATION to use SET	3	4
Concern about lenient teaching and grading	1	2
The use of SET for students' selection of teachers unfeasible	1	1
Teaching rewards fail to recognise many dedicated teachers	1	1
Faculty members' limited changes due to SET	10	15
Inconsistent improvement in teaching	3	6
SET data having not too much impact unless scores are low	8	9

Note. SET: student evaluation of teaching

Although using NVivo 12 was relatively time-consuming, it enabled me to systematically construct the individual and group theories of action by analytically naming and grouping ToA codes and categories. I could also easily search text or regroup ToA components with just one click. NVivo 12 was helpful for my clustering data which means “grouping and conceptualising objects that have similar patterns or characteristics” (Miles et

al., 2014, p. 279), which enabled me to easily construct the group theories of action about the SET and voting evaluation in the subsequent stage of data analysis.

Group ToA Construction

A group ToA construction involved summarising actions across individuals sharing the same roles (i.e., QA officers, administrators, and faculty members), then cross-checking the constraints and consequences of each group's actions regarding the SET and voting evaluation. During this process, I identified across-case patterns of the PBM components by going back and forth across the individual participants' theories of action regarding the SET and voting evaluation. I sometimes referred to the original interview transcripts and evaluation documents to add evidence and to verify the newly generated theories of action. Altogether, I constructed five group theories of action for the SET and voting evaluation: three for the SET and two for the voting evaluation (QA officers did not participate in the voting evaluation) (see Appendix E for an example of a group ToA).

Summarising Actions. I started analysing the group ToA by summarising the common actions across individual groups of QA officers, administrators, and faculty members. This involved "counting actions" (Robinson & Lai, 2006, p. 152), which means I counted the number of times an action was mentioned by each participant group. Counting may be a good way to avoid my potential bias and check the robustness of my insights (Miles et al., 2014). I counted actions by looking at the categories of action components on NVivo 12 to check the number of participants who shared the same actions or the number of references a participant made to action. For example, during the first iterative analysis, I classified the administrators' actions regarding SET into three groups: checking SET results, solving SET problems and inaction. I noticed that four administrators mentioned they compared SET scores across individuals and faculties when checking SET scores. Thus, I noted down these participants' name codes (e.g., I6A1, I4A1, I4A3, I4A5) next to the action in the group so that I could return to their account during my thesis write-up.

Notably, in my constructed group ToA, the participants' actions were not mutually exclusive. For example, an administrator may undertake three actions: checking SET scores, solving SET problems in the case of faculty members' SET underperformance and taking no further action (i.e., inaction) when there was no SET underperformance. The participants' actions were analysed iteratively from descriptive to more analytical. For example, during the first iteration, the administrators' actions were categorised as checking SET results, solving SET problems and inaction (see Appendix E). After being checked against the faculty evaluation purposes and orientations (see Figure 1), the administrators' actions were recategorised as complying, problem-solving and disengaging. The analysis of the participants' actions against the faculty evaluation purposes and orientations allowed me to

see if there were similar or different approaches to the evaluation orientations among the three groups of participants.

Identifying Constraints. Identifying the constraint sets of each group involved exploring how “organisational values, culture, and other conditions lead to particular patterns of activity” (Robinson & Lai, 2006, pp. 162–163) about the SET and voting evaluation (see Appendix E for an example of a group ToA). I first worked backwards through individual participants’ theories of action to check what constraints explained their patterned actions. When there were similarities in the constraints for certain actions, I listed them and condensed them into categories if possible. For example, I found similar constraints that explained the administrators’ inactions regarding SET were administrators’ “concerns about students’ preference for lenient studying that influenced their SET ratings” (I4A1), “belief that SET scores are not real indicators of teaching quality” (I2A4), “perceived problems with accuracy and reliability” (I4A1), and “perceived SET scores being meaningless” (I4A2). Hence, I grouped these similar constraints into one category of constraints named “perception” and a code named “Administrators were concerned about the values of SET for teaching improvement” (see Appendix E). In contrast, when various constraints explained the patterned actions, I included these different constraints in the constructed group ToA. For example, when the administrators’ checking SET results action was influenced by both the SET evaluation policy expectation and by the participants’ values of SET, I retained both constraints in the constructed group ToA.

Identifying Consequences. My approach to analysing the consequences for the group was like that for the constraints. I aggregated consequences from individual theories of action into group-level ones. I also extensively compared the stated purposes of the SET and voting evaluation and the participants’ expected use of these evaluations to identify if any intended or unintended consequences were missing. Like the consequence analysis of individual ToA construction, I read the data set multiple times to confirm and disconfirm the causal links between the actions and reported consequences (Robinson & Lai, 2006). For example, one administrator espoused that the voting evaluation was impactful in grading, honouring and motivating faculty performance. However, two of the faculty members perceived that this was not the case. It turned out that the administrator was unaware that her prioritising some evaluation criteria over others made the faculty members uncomfortable about the voting processes and results. The participating faculty members perceived the impact of the voting evaluation as not as positive as that perceived by the participating administrator. Thus, I counted the administrator’s theory as espoused theory (Robinson & Lai, 2006) because her explanation was less plausible than that of her faculty members.

Notably, I applied the “relative frequency” principle for my constraint and consequence analysis, which means that my data selection was not based on the frequency but on the

relevance or the “contribution it makes to the explanation of the problem” (Robinson, 1993, pp. 125-126). This is because “some behaviours which occurred frequently were not interesting, and others which occurred very infrequently were highly interesting” (Robinson, 1993, p. 126). Thus, I mainly used counting to “verify a hunch or hypothesis” and to keep myself “analytically honest” (Miles et al., 2014, p. 282). However, I do not present the number of the participants’ responses to specific constraints and consequences during the ToA description write-up to avoid the readers’ false assumption about the importance of constraints based on their frequency rather than their quality of explaining practices. Like the individual ToA construction process, the group ToA was constantly constructed, peer reviewed and restructured until each group’s ToA was satisfactorily completed. The groups’ theories of action were constantly reviewed and refined until the write-up of the thesis was completed.

Ethical Considerations

The University of Auckland UAHPEC approved this research, and I adhered to the committee’s requirements. I made several ethical decisions to address key challenges during the research projects.

Addressing Dual Roles

In this research, I hold a dual role of both an insider as a former faculty member in a Vietnamese public HEI and an outsider as a doctoral researcher at a university overseas. My dual role as both an insider and outsider has pros and cons. As an insider, my preunderstanding, prior experience and relationship to the Vietnamese higher education context might put me at the potential risk of bias or subjectively jumping to a hasty conclusion (Sikes & Potts, 2008). To reduce potential subjectivity and bias throughout my research stages, I required awareness of my role duality (Brannick & Coghlan, 2007). I believe that educational research is value-laden (Robinson, 1993), which means that my values influenced my choice of research topics, methodology and data interpretation. Instead of withholding my values, I carefully reflected on them together with my potential subjectivity that influenced my research choice of topic, design, and methodology. My reflexivity has been explicitly presented throughout different chapters of this thesis, from my positionality in the introduction to the ways I ensured validity, presented later in this chapter. Through my exposal of my reflexivity, readers can make an informed decision concerning the trustworthiness of the research (Merriam & Tisdell, 2016).

One of the challenges of this research was balancing my dual responsibility to adapt to the local Vietnamese context while maintaining the UAHPEC standards. For example, some participants were unfamiliar with the PIS (see Appendix F) and consent form (see Appendix G), which allowed them to understand the nature of the research and ask questions before

deciding to participate in this research project. I chose to adopt “situated ethics” (McNess et al., 2015, p. 34), which means making moral decisions relevant to the Vietnamese context. For example, my first email invitation attempt received a relatively low number of email responses from the potential participants; thus, I then wrote a reminder email and invited them to nominate a convenient time for a quick phone call. I found that Vietnamese participants did not usually respond actively to the research invitation because I was unknown to them. Thus, the suggestion for phone calls was beneficial for both the participants and me to discuss the research purpose, the interview duration and procedure, the nondisruptive manner of conducting the research, how the results would be reported, and the benefits the participants and sites would gain from the study. My open opportunity to let potential participants talk and discuss via a phone call allowed them to make an informed decision as to whether they wanted to participate in my study or not. This opportunity worked for the Vietnamese participants in my study while still adhering to the UAHPEC regulations. Participants’ permission was confirmed by signing and returning the consent form (see Appendix G). Participation in this research project was completely voluntary. I emphasised that the participants could withdraw at any time without explanation. During the individual interviews, participants could request the voice-recording device to be turned off at any time. Participants would also be allowed to review and edit the transcripts of their interviews upon request.

Minimising Risks

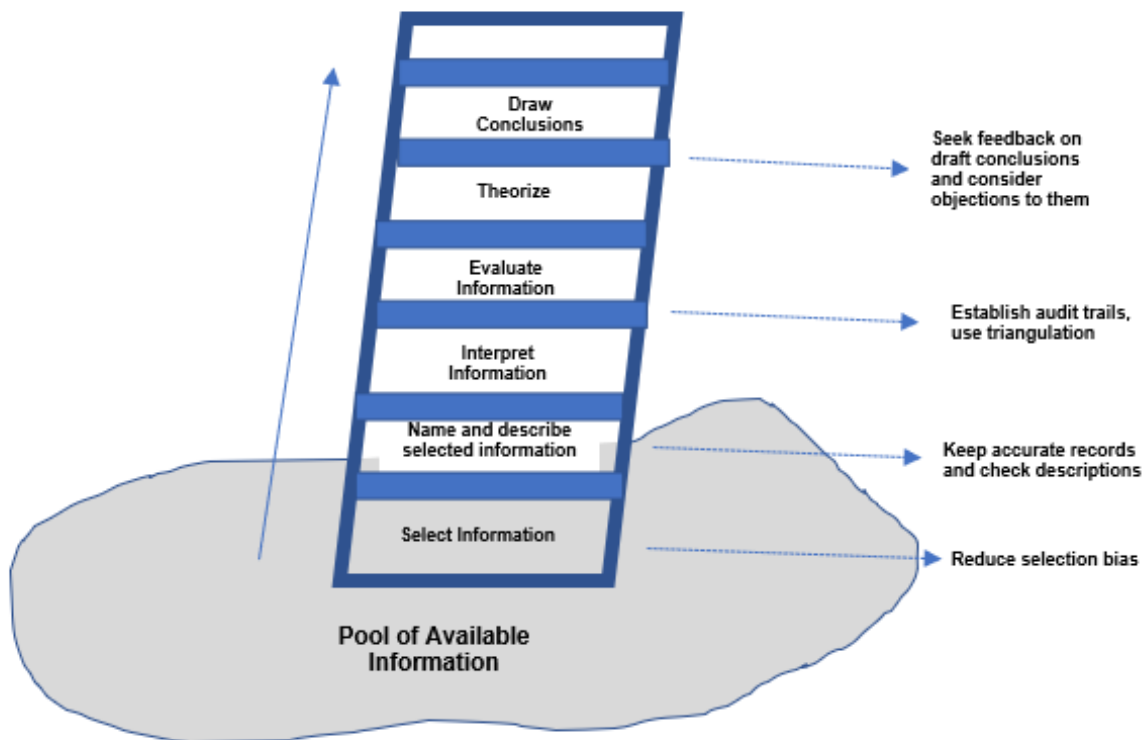
Following the UAHPEC ethical demands, I protected the participants’ safety, sociocultural sensitivities in several ways. First, I informed participants about their rights, for example, not to answer any questions that made them uncomfortable or to withdraw from the study at any time. As Turner (2010) pointed out, the ethical dilemma of undertaking fieldwork in China, Laos or Vietnam might be serious because the research sometimes needs to compromise when publishing data. Participants’ responses, including critical comments against the Vietnamese Communist Party government, are unsafe to publish. Such acts can result in charges of being “against the government,” and the respondents might risk losing their jobs at their public institutions. Thus, I shared with the participants the various measures I would make to ensure the confidentiality of their identities and opinions, such as using randomly assigned codes or hiding personal information. The clear communication of my awareness of and affirmation to reduce the potential risks enabled my participants to freely share their matters of concern. I emphasised my role as a previous “insider” rather than an “objective” outsider researcher, and this made the participants feel more assured to share their opinions with me (T. Q. Nguyen, 2015, p. 35). Most importantly, I communicated clearly with the participants that my goal was to understand practices from the implementing agents’ perspectives, represent all voices equally, and not critique any perspectives. My impartial role

and neutral stance allowed me to reduce my potential bias while gaining the participants' trust (Prasetya, 2020).

Additionally, I took several strategies to address cultural and psychological issues when doing fieldwork in Vietnam. For example, Vietnamese participants are generally unfamiliar with the ethics approval procedures. They also tend to depend on relationships, are concerned about face-saving and hide their thoughts and feelings (T. Q. Nguyen, 2015). Having lived in Vietnam for over 30 years, I was confident in adopting culturally appropriate and socially acceptable norms when conducting my research in Vietnam. First, I called the Vietnamese participants by their first names, not their last names. There are also different pronouns to address participants depending on their age and gender. Second, I compensated the interviewees for their time (Turner, 2010) by offering each participant a small koha (gift), equivalent to 10 USD, for their interviews that lasted between 45 to 90 minutes. Third, Vietnamese participants tend to be afraid of "losing face," so I tried to be sensitive to their "face" to make them feel comfortable. I downplayed the main researcher role and portrayed myself as a young PhD scholar on the way to learning and gaining experience through this first big project of her life. My eagerness to learn and my nonthreatening manner allowed me to have some initial conversations with potential interviewees in a relaxed and comfortable way. This is quite like the strategies taken by T. Q. Nguyen (2015) when she identified herself as a faculty member, an insider or one taking an equal stance to approach the participants in a more friendly and informal manner. I also tried to understand participants' interests, concerns, and styles during my informal conversations. Showing such knowledge about participants and "stroking egos" whenever possible proved helpful when participants opened up more readily (Turner, 2010, p. 128).

Increasing Validity

A rigorous qualitative study provides adequate evidence for its processes for readers to consider the trustworthiness of the study findings (Merriam & Tisdell, 2016). The rigour of my PBM research is determined by the validity of my constructed theories of action about the SET and voting. In PBM research, the validity of an interpretation is determined by "showing how it is more plausible" than others, whereas that of a theory is made by judging the basis of the "implicit theory from which they are derived" (Robinson & Lai, 2006, p. 54). Throughout the project, I established and judged different types of claims based on a *ladder of inference* (Robinson & Lai, 2006), a process of theorising and drawing conclusions based on various implicit steps of selecting, naming, describing, interpreting, and evaluating data (see Figure 9).

Figure 9*A Ladder of Inference*

Note. From *Practitioner Research for Educators: A Guide to Improving Classrooms and Schools*, (p. 58), by V. M. Robinson and M. K. Lai, 2006, Corwin Press. Copyright 2006 by Corwin Press.

In alignment with the ladder of inference, I applied several strategies as follows to increase the validity of each procedure taken to draw the conclusion grounded in data.

Reducing Selection Bias. One challenge of selecting information was *sampling bias* or a lack of representatives from various groups of participants. Thus, I chose to include different groups of participants (i.e., QA officers, administrators, and faculty members) differing in ages, disciplines and working institutions so that faculty evaluation practices could be seen from various perspectives. As stated earlier, I had the previous experience of being a faculty member in a public Vietnamese HEI myself, so I tended to look at faculty evaluation from bottom-up perspectives. The inclusion of other participant groups, such as administrators and QA officers, provided me with views from the management position. This minimised *confirmation bias* or my “tendency to give more weight to data and data sources that confirm rather than disconfirm” (Robinson & Lai, 2006 p. 59) my previous assumptions about faculty evaluation practices in the Vietnamese public HEI context.

Keeping Records and Checking Descriptions. During the data collection and analysis, I constantly checked my understanding of the data and kept a record of my data-analysis thinking process. I made notes of my interpretations or emerging conclusions which were then checked explicitly with participants in the following interviews. I would pose such questions as “some participants found that ... what do you think of ...?” This explicit communication and discussion of my emerging findings allowed me to find arguments and evidence, for or against, provided by other participants (Robinson & Lai, 2006). Using learning conversation principles (Robinson & Lai, 2006), I explicitly stated my assumptions about the participants’ accounts and invited them to comment on my statements. My constant probing, summarising, and requesting confirmation from the participants during the interviews allowed my constructed theories of action to describe the participants’ accounts of practice accurately. I used a tape recorder and transcribed the interviews to ensure that the interview data was reliable. During my data analysis, I used NVivo 12 to record, code, and retrieve interview and document data. NVivo 12 enabled me to easily climb down the rung of my ladder of inference by cross-checking my constructed theories with my interpretations linked to original data presented on the same platform.

Establishing Audit Trails and Triangulating. My audit trail includes a variety of notes such as field notes, checklists, or analysis memos (Robinson & Lai, 2006) to record my thinking process during my data collection and analysis. For example, I noted my challenges during data analysis regarding selecting a relevant word to capture the participants’ actions as follows:

17 August 2020

I have been challenged by finding a suitable verb to describe a particular action: “Usually I do not get [SET scores] under 8 [out of 10], so I just let it be” (I1T1). At first sight, I attribute the verb “dismiss” to the participant’s “cứ để vậy thôi” (just let it be) account of practice. To dismiss is defined by the Online Cambridge Dictionary as “to decide that something or someone is not important and not worth considering.” However, in other sections of the participant’s interview response, I recognised that she valued student feedback but did not have adequate SET data to make decisions regarding her teaching improvement. I found the verb “disengage” better captured the essence of the participant’s let-it-be action since the participant decided to stop using SET for improvement. After the participant had checked that scores were acceptable, she stopped thinking about or using SET data.

Such a memo helped me keep track of my emerging interpretations of the data, especially the basis on which my conclusions were made. This allows others to easily access and check the validity of my interpretations and conclusions (Robinson & Lai, 2006).

Seeking Feedback on Preliminary Conclusions. One way to increase the validity of PBM research findings is to seek feedback on early conclusions and be open to alternative ways of interpreting data and making conclusions (Robinson & Lai, 2006). During my data analysis, I encountered some challenges in gaining feedback from the participants due to their busy working schedules and their lack of familiarity with the English language. Despite those challenges, I managed to get participant feedback on a couple of the transcripts by email exchanges. I also had some of my translated excerpts (see Appendix C) peer reviewed by five Vietnamese doctoral candidates in education. I also sought peer review (Merriam & Tisdell, 2016) of my preliminary findings of the participants' theories of action about the SET and voting by constantly discussing them with my supervisors, my critical friends or through various conferences in New Zealand, the USA and Vietnam. Feedback from these exchanges helped me self-critique and refine my constructed theories of action about faculty evaluation in the study context.

Ensuring Transferability

The transferability of the study findings is justified based on the premises suggested by Lincoln and Guba (1985) that the merit or the potential applicability of research findings depends more on the potential appliers than on the investigator. This idea emphasises the role of the researcher to equip possible appliers with adequate information or "sufficient descriptive data" (p. 298) to enable future transferability of the research findings. The transferability of the study was also enhanced through maximum variations, which enabled "the possibility of a greater range of application by readers or consumers of the research" (Merriam & Tisdell, 2016, p. 257). The maximum variation in this study involved sampling institutions of various sizes and academic disciplines. There were also variations in three groups of implementing agents with different voices and interests to be heard. Including various participating HEIs and individuals allows this research to be more transferable to other situations with similar contexts (Merriam & Tisdell, 2016).

Chapter 4: SET Theories of Action

This chapter presents findings of the participants' theories of action regarding SET, a practice involving collecting, analysing, reporting, and using student ratings and comments about faculty members' teaching activities. This chapter starts with the SET ToA of four QA officers, followed by those of 18 administrators and 20 faculty members.

QA Officers' SET ToA

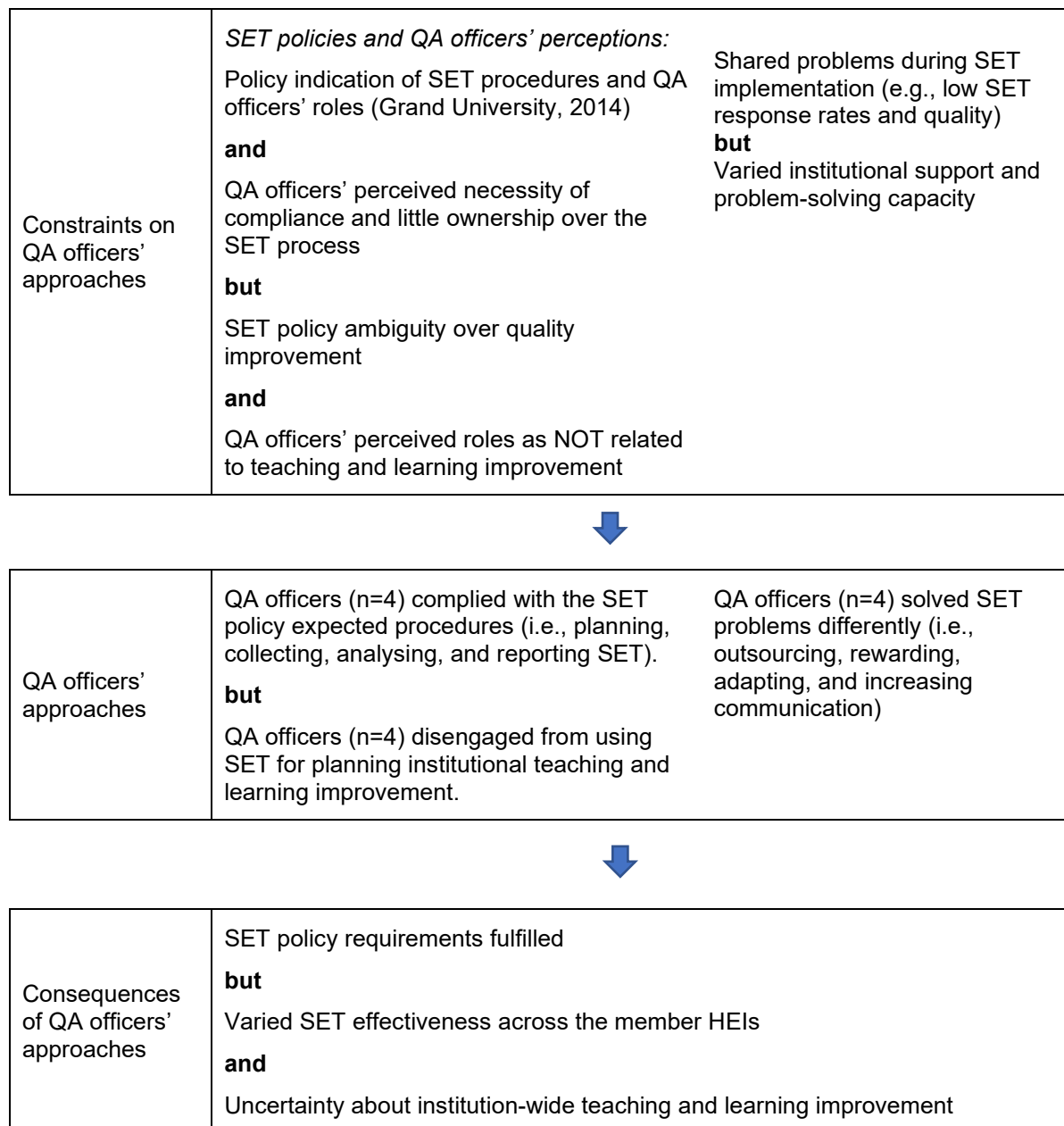
The following sections describe the QA officers' SET ToA: their main approaches to SET, together with the constraints and consequences of these approaches (see Figure 10). Overall, all the QA officers from the four participating HEIs complied with the SET policy requirements, but they did not use SET to improve teaching and learning at the institutional level. While the QA officers shared similar SET implementation problems, they solved problems differently. There were various constraints and consequences of the QA officers' approaches to SET.

QA Officers' Approaches to SET

The QA officers took three main approaches to SET: complying, disengaging, and solving problems during SET processes.

Complying Approach. The Grand University's (2014) SET policy indicates that the member HEIs' QA officers must follow the institutional policy and the rectors' direction to plan and execute SET collection, analysis, and reports for all courses. All four QA officers complied with the SET expected procedures by planning, collecting, analysing, and reporting SET data (see Figure 10).

First, all the four QA officers created annual plans for SET implementation under the rectors' approval or direction in slightly different manners. While the QA officers at HEI 2 were more autonomous in initiating SET tasks, those at HEI 1 waited for the QA directions to be "sent down" by the rectors (I1Q1). QA Officer I1Q2 usually made SET annual plans for their rectors to approve. In contrast, QA Officer I1Q1 followed the directions of her rectors, who "set out policy intentions, [concerning] the aspects, scopes and uses of evaluation, as well as ways to deal with SET results." QA officers at HEI 1 must satisfy SET requirements at two levels of governance: those assigned by their rector board and those required by the Grand University (2014). That explains why SET implementation at HEI 1 tended to be "different every year" depending on the different emphasis from two levels of governance each year (I1Q1).

Figure 10*Summary of QA Officers' SET ToA*

Note. ToA: theory of action. QA: Quality assurance. SET: Student evaluation of teaching. HEIs: higher education institutions. *n* indicates the number of participants who reported taking specific actions. The numbers of participants taking various approaches are not mutually exclusive.

All four QA officers coordinated the SET collection and analysis either within their own institutions (I2Q1, I3Q1, I4Q1) or with an external agency (I1Q1). QA Officer I3Q1 and her colleagues collected SET surveys by themselves to monitor students' responses closely. On the other hand, QA officers at Institutions 2 and 4 assigned another department to support SET form distribution and collection. QA Officer I4Q1 created SET online forms on Google Docs and sent the links to academic affairs officers, who then forwarded the links to students.

In contrast, QA Officer I2Q1 worked with their academic affairs officers to manually distribute and collect SET forms:

We distribute SET forms to M1 & M2 [general courses] students. We assign the Office for Academic Affairs to distribute SET forms to students of M3, M4 & M5 [specialised] courses. We go to each class to give them the SET forms. (I2Q1)

Next, all the QA officers sent summaries of SET scores and comments to administrators and individual faculty members. Three out of the four QA officers mostly adopted the SET form provided by the Grand University (see Appendix H), so they shared a relatively similar SET analysis and report approach, which involved aggregating the average rating scores and compiling written feedback from students. The QA officers' SET result reports (see Appendix B) were used to inform faculty members about the rating scores and standard deviation of individual statements about different aspects of faculty members' teaching performance. To comply with the SET requirements, the QA officers also kept the confidentiality of SET results by only giving administrators a compilation of all faculty members' SET results, whereas individual faculty members knew "only about their own SET results" (I3Q1).

Disengaging Approach. Despite complying with the QA policy requirements, all four participating QA officers did not use SET for planning institutional teaching and learning improvements. These QA officers were unwilling to take further improvement actions because they were not officially required to do so. Besides, the QA officers found SET only one portion of the many QA tasks they had to fulfil annually. There were too many accreditation and QA tasks, so they did not have enough time and resources to use SET for improvement. They reasoned that they did not have the expertise for pedagogical improvement. Thus, the QA officers generally engaged in providing SET information but not in making institutional plans for teaching and learning improvement. While they could observe changes in SET scores, they could not evaluate changes in teaching:

We do not know and cannot tell [you] about teaching and learning changes because we lack proper expertise in the subjects and programmes. QA centre cannot evaluate changes in teaching because SET forms cannot reflect that. We only provide additional information [about teaching]. (I2Q1)

Problem-Solving Approach. During SET implementation, there were some problems, such as a generally low SET response rate and quality. QA officers from the three HEIs reported various strategies to solve the problems.

Outsourcing and Rewarding. Two of the strategies taken by QA officers at HEI 1 to solve their low SET response quantity and quality problems were to outsource the SET

collection and analysis to an external survey agency and use SET results to reward faculty members. First, QA officers at HEI 1 found that they could not increase the SET response rate by themselves, so they hired an agency to support the QA officers to “deal with the SET process more effectively” (I1Q1). The survey agency then made a SET collection and reporting plan for the rector and QA officers at HEI 1 to approve. The agency was responsible for most SET tasks during SET implementation, including:

- Construction of the software system for SET online collection, for example, importing feedback forms, designing interfaces, generating invitation or reminder emails to students
- Collection of SET responses from students, sending students reminder emails once every 2 days
- Analysis and summary of course feedback results to HEI 1 in a package that includes raw data and reports

The survey agency provided the QA officer at HEI 1 with three main SET result reports: the rating scores, the written comments, and faculty members’ SET scores ranked from the highest to the lowest. The faculty members’ SET ranking lists were then used to reward and solve problems with faculty teaching performance.

To improve the potential effectiveness of SET, QA officers at HEI 1 used the SET score ranking list for rewarding faculty members based on the categorisation of faculty members into three groups: the highest scoring, the middle-scoring, and the lowest scoring one. They proposed a financial incentive scheme in which the highest score faculty members would receive a significant bonus:

We proposed the financial incentives [for highest score faculty members] to motivate them to contribute [to our institution]. For example, in our previous semester, we applied a policy that the highest scored faculty member(s) received 10 million VND (approximately 430 USD). It is such a significant decision [that involves a vast number of financial bonuses]. (I1Q1)

After the rector approved the SET incentive scheme, the QA officers at HEI 1 executed it immediately as they believed that those who received the incentives must “feel very motivated.” However, SET results were only used for granting financial bonuses but not for “promotional purposes” such as “appointing new positions.” In the SET policy (Grand University, 2010), SET was first used for improvement and later for emulation and reward at an unspecified time. The use of SET for financial rewards at HEI 1 reflects that the institution was willing to move to the second phase of using SET for emulation and reward.

Adapting SET Forms and Reports. Another strategy taken by QA officers from HEI 3 to solve the problems of low SET response quantity and quality was to develop their in-house SET collection and report forms. After 1 year of using the Grand University's SET forms, QA officers at HEI 3 decided to conduct R&D (research and development) to create one "package" of SET collection, analysis, and report by themselves (I3Q1). The in-house SET package development involved several steps.

First, the QA officers from the HEI experimented with varying scales of SET scores during the R&D of SET design. At first, they tried using a four-point Likert scale in their SET survey. However, they found that with that scale, "faculty members found it difficult [to interpret] whether two-point is acceptable or not" (I3Q1). Thus, they decided to use the five-point Likert scale because "with the five-point scale with one being the lowest and five being the highest score, three-point can be [easily considered] to be the average." The QA officers cared more about identifying what works than merely complying with the Grand University's guidance and chose the rating scales that worked at their institution. The survey provided by the Grand University contains statements about both teaching activities and learning assessments. However, QA officers at HEI 3 decided to "exclude questions about learning assessment" and "use another method to evaluate faculty members' methods of assessing students" (I3Q1).

QA officers at HEI 3 also treated their SET data differently during their SET design, collection, and analysis. While the SET designers at the Grand University (2014) considered SET responses as absolute data, QA Officer I3Q1 regarded it as the "raw data" only. In the Grand University's form, SET results are calculated as the average scores of students' choice in response to individual statements about faculty members (see Appendix B for a SET report sample). However, the QA Officer I3Q1 found that way of data summing potentially biased if they had not applied statistics knowledge in treating such data:

It [SET data] is only the raw data – it is not something we summarise straight away. We must use our expertise to screen out, categorise and evaluate the data under certain principles... It [SET conduct] must meet all scientific research requirements.

Due to the potential subjectivity of the Likert scale that measures students' feelings, the QA officers at HEI 3 took active measures to "eliminate potential biases and prejudices" from students as much as possible. They also took cautious steps to properly collect SET data to reduce the level of students' bias in SET survey responses:

At the two-thirds point [of each course], we distributed the SET surveys to students in their classrooms. We collected right at that time so that students were not affected by the external environment, and they were not able to spend too much time judging based on prejudices about the faculty members. (I3Q1)

Although “students have the right to express their feelings and opinion,” the QA officer at the HEI thought that not all students had the proper attitude (toward SET responses) and that students were “not capable of evaluating faculty members” (I3Q1). Therefore, QA officers at HEI 3 decided to make some “trap” questions in their SET question items to eliminate conflicts in student responses:

If a student is serious about responding to SET questions, their answers are not self-conflicting. For example, one question asks if the student is satisfied with the teaching method; the students might tick randomly 4 or 5. However, another question is whether the method helps them develop necessary skills, and if the student ticks 1 for it, the two answers are self-conflicting, just like that. If students are attentive to the SET questions, they will answer them thoughtfully. If not, they will fall into the trap of giving conflicting answers. (I3Q1)

The inclusion of trap questions in the in-house SET forms helped QA officers at HEI 3 spot which SET students seriously consider their responses. It also allowed the QA officer to eliminate self-contradictory student responses, which increased the reliability of SET data: “after collecting student responses, we started to read and screen out the conflicting answers because the survey items can be cross checked against one another” (I3Q1). Consequently, while other QA officers (I1Q1, I4Q1) complained about the Grand University or their inability to increase student response rates and reliability levels, QA Officer I3Q1 was confident because they applied principles similar to conducting scientific research in conducting and analysing their SET surveys. In reporting SET scores, QA officers at HEI 3 chose to generate the average mean that reflects the degree to which the faculty member satisfied learning requirements. They also provided the faculty members with SET score interpretations reflecting students’ satisfaction levels. For example, the mean score of 4.50 means “93.8% of participating students are satisfied.” In contrast, 4.86 means that the faculty member “has met the learning requirements of all students participating in the survey” (I1Q3).

Increasing Communication. The last strategy taken by QA officers from HEIs 3 and 4 to solve the problem of low SET response quantity and quality was to increase communication about SET in several ways. One reason for the HEIs’ low SET response quantity was that students, especially the freshmen, were not familiar with SET ideas and processes. QA Officer I3Q1 found it necessary to communicate and guide students with basic meaning, processes, and principles of SET:

At our institution, it is a norm that newly enrolled students will be provided with a clear explanation about the number of surveys, ways to respond to the surveys and how the surveys are related to them. Some students are more mindful about it than others, depending on each student’s attitude, and they might not be aware that these surveys

are for evaluation or assessment purposes. However, students need to read and know about such core principles [related to students' survey responses].

Unlike other QA officers, the QA officer found her institution's communication with students about SET effective. Due to the effective communication with students, the participant found that students in her institution usually answered frankly about faculty members' teaching in their SET responses.

Another problem for low SET response quantity was the lack of shared understanding about the purpose of SET among students and faculty members. Thus QA Officer I4Q1 tried to solve the problem of low SET response rates and quality by reminding students that the central purpose of the SET collection was to support teaching improvement. In a public report of SET results of HEI 4 (2018), QA Officer I4Q1 highlighted the need for faculties and departments to closely collaborate with the QA centres to propagandise students about the benefits of SET practice" (p. 8). The propaganda must encourage students to read the SET question items more carefully, rate accurately and give more detailed comments to their teaching faculty. QA Officer I4Q1 also reported changes to follow up on students' SET responses to let them know their voices mattered. He believed that publishing how SET was used to improve practice would promote more active engagement of students and faculty members during SET implementation. As QA Officer I4Q1 asserted, "if students and faculty members know that SET results can be used as a useful means for creating changes in teaching and learning, they will be more willing to participate in SET practice."

Constraints on QA officers' Approaches

There are several constraints that explain or influence the QA officers' approaches to SET. Overall, the SET policy (Grand University, 2014) and the QA officers' perceptions predominantly influenced their complying and disengaging approaches, whereas varied institutional support and problem-solving capacity explained their problem-solving approach (see Figure 10).

SET Policies and QA Officers' Perceptions. The SET policies strongly influenced the QA officers' complying and disengaging approaches in several ways. First, the SET policy (Grand University, 2014) rigidly specifies that they must conduct SET for all faculty members and all courses, which made the QA officers in the member HEIs extensively conduct SET. As QA Officer I1Q1 revealed, SET was on a significantly larger scale than other evaluations, with "100% lecturers to be evaluated, and 100% students being allowed to evaluate their lecturers." Including 100% of courses to be evaluated in the annual plan was a vital indicator of member institutions fulfilling QA activities in their annual report sent to the Grand University. Due to the rigid requirement, the QA officers had all courses surveyed even though they found the surveys were too much or overlapping:

Our surveys are too much... too extensive. For example, we had to conduct the SET survey for a course on Principles of Marxist-Leninist Philosophy in Semester 1, and then the same survey was required for the same faculty members in Semester 2.... The course feedback for several hundreds of courses over about 100 student responses each was too much. (I2Q1)

Second, the QA officers had little ownership over some SET processes. Although they were the direct implementing agent of SET, the QA officers at HEI 2 could not determine the number and types of courses they conducted SET. QA Officer I2Q1 found a better way to improve the quality of SET at her institution was to “conduct course evaluation for M1 [general courses] once every year, giving more time to improve the evaluation and improvement of the M3, M4 & M5 [specialised courses].” However, the QA officers’ proposed change was unfeasible unless the Grand University approved their proposal. QA Officer I2Q1 was quite disappointed when she found her change request was primarily ignored, and the SET requirement for 100% courses to be surveyed remained unchanged. This failed attempt to propose a more relevant SET practice gave QA officers little power to make decisions and be change agents at a lower institutional level.

Third, even when there was some policy allowance for the QA staff to make some adaptations, the actual requirements from the Grand University’s QA Institute still prevented those from the member HEIs from doing so. It is indicated in the SET policy document that staff at member institutions can “use the core content in the student course evaluation sample, possibly adding [new] content that is suitable to the institutions’ contexts” (Grand University, 2014, section 2.3). In theory, this policy indicates that QA officers could adapt their SET forms to be relevant to their institutional contexts. However, the QA officers’ adaptation was constrained by the actual requirements for SET reports to follow precisely the sample provided by the Grand University (2014, s. 2.2). In this required report sample, QA officers needed to insert scores and the standard deviation for each predetermined SET statement (see Appendix B). Thus, even when QA officers modified the SET forms, they still needed to produce the required report sample. As QA Officer I4Q1 stated, he received complaints from the QA officers at the Grand University for having modified the SET form provided by the Grand University:

We used to put some questions which do not follow the university’s sample; however, after they [QA Institute] complained that there is inconsistency in our reports and that of the whole university, causing them difficulty in reporting to the public.

After receiving complaints from the staff at the Grand University, some QA officers avoided modifying SET reports for fear of future inconveniences. QA Officer I4Q1 reasoned that it was acceptable to cast aside individual or institutional needs to guarantee the unity of

the SET practice across the whole Grand University: “to ensure the wholeness, I think everyone needs to sacrifice a few of our individual needs” (I4Q1). Hence, he resorted to the standard ways offered by the Grand University to ensure consistency in SET practice. Similarly, QA Officer I2Q1 affirmed that changes in the SET survey form were not possible due to the Grand University's fixed mechanism: “it [the evaluation] has the university-wide mechanism that cannot be changed. It is like a chain with one linked to another.”

Fourth, while the Grand University's SET policy requirements reinforce compliance, its ambiguity over quality improvement caused the QA officers to disengage from using SET to plan institutional teaching and learning improvement. The SET policy document (Grand University, 2014) requires rectors and QA officers to propose solutions for quality improvement at their institutions (section 2.3, pp. 25, 28). According to the Grand University's (2010, 2014) guidelines, various stakeholders must use SET as a basis for planning to improve training quality:

SET is expected to be a basis for institutions and staff to show accountability and responsibility and strengthen their quality culture. It helps construct and cultivate a quality culture at the university (Grand University, 2010, p. 1).

However, what is quality and how it is linked to the improvement of teaching and learning is not defined in the document. Thus, the QA officers did not see their roles directly related to teaching and learning improvement, and they were unwilling to take further actions if the SET policy did not require them to do so. When being asked about QA officers' attempt to use SET for teaching improvement, the QA officers referred to the university guidelines UG5077 (2014) that required “no extra forms of collecting feedback under these guidelines” (I2Q1). Instead of linking their roles to support teaching and learning improvement, the QA officers saw them as SET information providers. As QA Officer I3Q1 stated, SET was “only one channel of information” and “only a source of warning for faculty members.” She thought a QA officer was “only a ‘guard of the temple’ (chỉ là một anh gác đền thôi) who makes sure the institution meets all standards and regulations related to teaching and learning” (I3Q1).

Also, the QA officers did not directly relate their roles to using SET for improvement because they did not have the expertise to solve academic issues. Thus, these participants thought that using SET for teaching and learning improvement should primarily be the role of administrators and faculty members. These QA officers reasoned that administrators had a better understanding of professional knowledge. They found that administrators had the authority to interfere with faculty members' “disciplined-related” or “expertise-related matters” like SET feedback. They emphasised that faculty deans need to be responsible for thinking through how they interfered, through which channels and whether they decided to assign classes in the following semesters for the faculty members with low SET scores:

Faculty members' professional evaluation or assignment is a task of faculty [administrators] who understand the faculty members ... The faculty [deans and members] are responsible for assigning courses faculty members to undertake, so they need to evaluate and understand their professional capacity. (I3Q1)

Faculty [administrators and faculty members] must be responsible for expertise-related matters. They can conduct peer evaluation of teaching or other methods to evaluate teaching. As QA officers, we cannot interfere in faculties' matters. (I2Q1)

For full-time faculty members who get low scores, I think their faculties [deans] and departments [heads] should find ways to talk to them, train them or assign them proper subjects. Faculty deans or department heads need to organise some kinds of discipline-related meetings to discuss it. (I1Q1)

Varied Institutional Support and Problem-Solving Capacity. Despite encountering similar problems, the QA officers had varied institutional support for SET and problem-solving capacity, resulting in variation in problem-solving approaches.

In terms of the institutional support for SET, there were several reasons for the difference in the participating HEI rectors' attention and investment in SET and QA. First, like other QA practices, SET was a relatively new concept in Vietnamese HEIs. While the HEIs' rectors generally recognised the importance of SET, they did not equally invest in improving SET. Potential students and parents did not widely recognise SET or QA at the societal level in their choice of universities. Thus, some Vietnamese HEI rectors focused on "more immediate issues" such as improving the infrastructure rather than SET (I1Q1). Also, rectors might not be willing to invest in SET due to financial and human resource constraints. According to QA Officer I1Q1, her institution was fortunate to enjoy "exceptionally high financial independent status" and higher tuition fees than others, so some of her proposals for using SET were more easily finalised and put into practice. In contrast, QA Officer I2Q2 was overloaded with SET and other QA tasks due to a staff shortage. She commented that the Grand University demanded "too many surveys to be conducted" and that she and her colleagues did not have time to propose any other improvement plans.

In terms of varied problem-solving capabilities, three out of four participating QA officers perceived themselves as helpless to solve the problem of low SET response quantity and quality, whereas only one of the four felt confident about her SET implementation. Most of the participating QA officers found that the students' results and comments did not always accurately reflect teaching effectiveness: "sometimes a faculty member might get low scores for having conflicts with students, so it does not mean that her teaching is bad" (I1Q1). This was because many students did not enthusiastically participate in the SET and did not respond appropriately to SET. A common observation was that many students gave half-hearted

responses to SET by actions such as having a friend complete their SET forms on their behalf or by ticking all statements randomly at a certain score. Unfortunately, most participating QA officers found it challenging to resolve the problem:

Our SET response rate is not as high as expected as the situation in other institutions. We had to remind students again and again... The quality of the responses is another thing to consider. Sometimes we need strategies to coerce students to complete the form; however, we cannot ensure they fill it accurately – it is tough for us to control it. (I4Q1)

[We] cannot guarantee the responses from all students who were required to evaluate. That [serious attitude to SET] is the characteristic of students at our institution. Even though they receive our reminder email seven or eight times, they still do not do it. We cannot do anything [about it]. (I1Q1)

While three out of four QA officers expressed their difficulty, or even helplessness, in solving problems with the SET low response and lack of reliability, one of them (I3Q1) felt confident and chose a different approach to solving SET problems. The QA officers at HEI 3 set a higher standard for SET results compared to other institutions: “our institution’s acceptable SET score is at least 4-point. 4-point means students “agree” with the statement – at least students agree with it [the teaching]” (I3Q1). While a 3-point SET score was considered an average and acceptable level at the other three institutions, HEI 3 set a four-point score as its minimum requirement for faculty members: “when a student ticks three-point for one statement, it means that he/she is still hesitant about the teaching (I3Q1). As the Director I3Q1 explained, it is “obvious” that faculty members are accountable for providing students adequate knowledge and instruction for mastering discipline-related skills. Therefore, it is “not acceptable” at HEI 3 that a student must be hesitant about “whether the faculty member’s teaching is appropriate” or whether “the faculty member is adequately responsible for their teaching job.” In other words, faculty members must be responsible for ensuring their SET scores rated by their students were above the average standards.

Consequences of QA Officers’ Approaches

While all the QA officers’ approaches to SET helped their HEIs satisfy SET policy requirements, many could not resolve the existing SET problems. These participants could monitor faculty members’ SET scores, but they were uncertain about the institutional use of SET for teaching and learning improvement.

SET Policy Requirements Fulfilled. The QA officers’ compliance with the SET policy requirements (Grand University, 2010, 2014; MoET, 2010, 2013) signified their HEIs’ fulfilment of their SET roles assigned by the Grand University and external accredited agencies.

Completing the required SET implementation contributed to the member HEIs “fulfilling all the requirements in terms of QA reports” annually set by the Grand University (I4Q1). Some QA officers considered compliance as a positive quality of an institution, as compliance meant having the ability to satisfy the SET requirements:

Our QA strength, in general, is that we implemented SET under the Grand University’s regulations. If the parent university puts forward any requirements, we will follow them strictly. That means we will not miss any annual reports that the Grand University requires us to submit. (I1Q1)

Varied SET Effectiveness Across the Member HEIs. Due to various problem-solving strategies, the member HEIs’ QA officers revealed different levels of SET effectiveness. The QA officers in HEI 3 were confident that they effectively addressed the problems of SET low reliability and validity. Their efforts to develop in-house SET collection and report forms helped QA officers at HEI 3 meet their institutional goals towards “fairness and transparency” in faculty members’ SET scores (I3Q1). However, in other HEIs, the problems of SET reliability and validity remained unsolved. QA officers in these institutions continued to modify their SET processes, but it was uncertain if they could then solve the problems of the limited quantity and quality of SET. One QA officer thought that her QA staff would try using the online SET collection to ensure that “all students have to respond to the form and be more honest with their responses.” However, the adoption of online SET collection in other member HEIs did “not guarantee to solve” the problems of SET low response rate and quality (I4Q1).

Uncertainty about Institution-Wide Teaching and Learning Improvement. The participating QA officers could monitor individual faculty members’ SET scores, but they were uncertain about their use of SET for teaching and learning improvement at the institutional level. A couple of QA officers observed that it was challenging to address the problem of faculty members’ SET underperformance in some cases. As reported by the QA officers, SET results were more significant for young faculty members who had short-term contracts than those with permanent contracts. Consequently, increasing SET effectiveness was not always possible, especially for certain faculty members:

Some faculty members, especially the tenured ones with permanent contracts, may not care much about SET results. The institution is sometimes called the “blockhouse” or a concrete shelter [lô cốt]. They are called “blockhouses” because they cannot be removed, and they had lots of contributions since the establishment of the institution. Their capacity [might fall behind the current standards of requirement], but they contributed a lot in the institution's early days. (I3Q1)

Also, the QA officers found that they did not have “proper expertise in the training subjects and areas” and that their institutions “have not yet developed an institutionally-shared agreement” about using SET results for improvement (I2Q1). Consequently, from QA officers’ perspectives, using SET results for improvement depends mainly on faculty-led or individual faculty members’ attempts. However, due to unresolved problems related to SET, it was questionable if administrators and faculty members would consider SET as a reliable and valid source of feedback for teaching and learning. The following sections about administrators’ and faculty members’ SET theories of action will provide more in-depth answers to the question.

Administrators' SET ToA

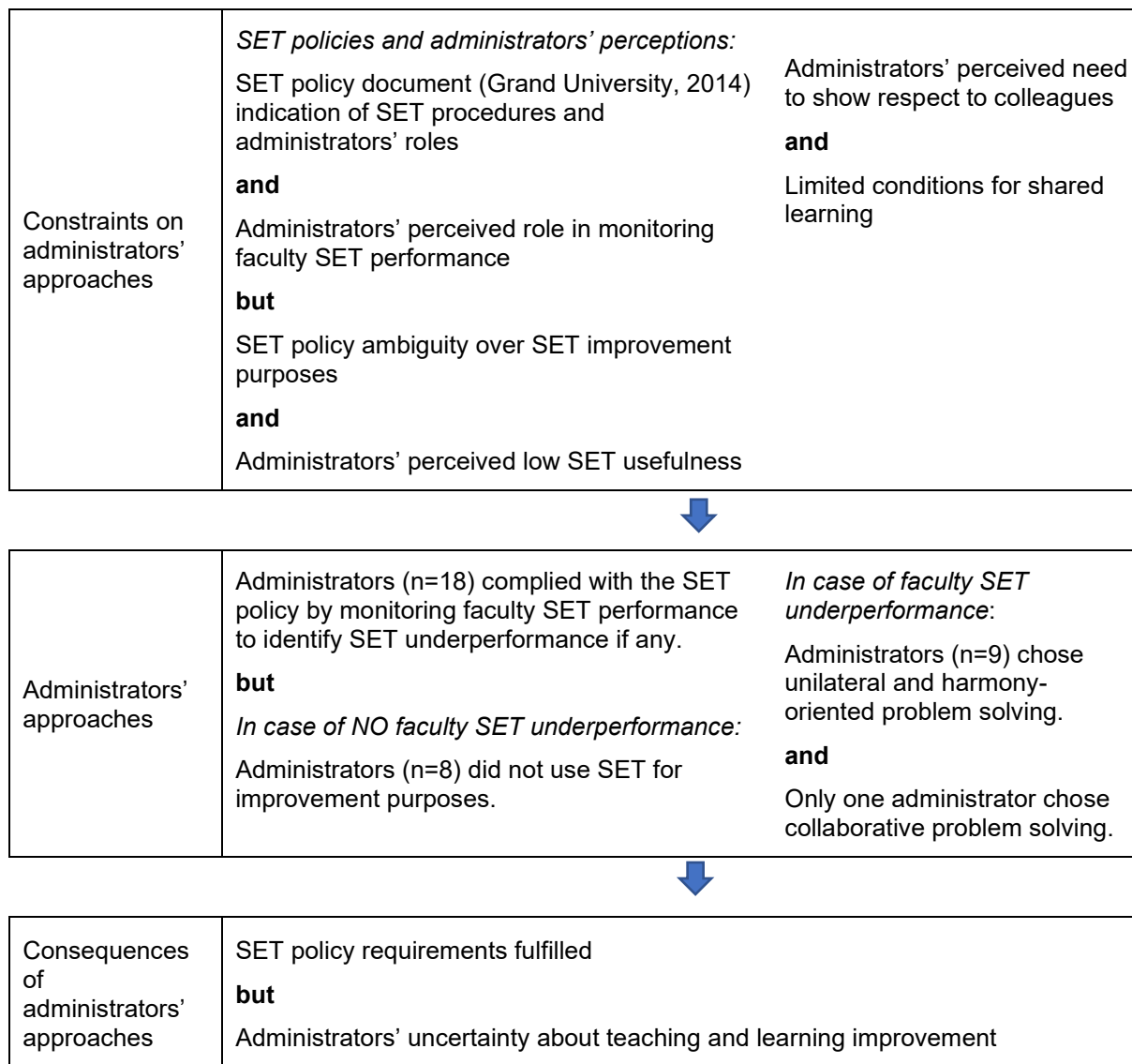
The following sections describe the administrators’ SET ToA: their approaches to SET, the constraints, and the consequences of these approaches (see Figure 11). Overall, all the administrators complied with the SET policy requirements; however, many did not use SET for teaching and learning improvement when there were no problems with faculty members’ SET underperformance. When there were SET-underperformance problems, the administrators mainly chose a unilateral and harmony-oriented approach to problem solving.

Administrators' Approaches to SET

As Figure 11 illustrates, the administrators’ principal approaches were complying, disengaging and problem solving. The following sections describe each of the approaches taken by the administrators.

Complying Approach. All administrators complied with the policy indication of SET procedures and administrators’ roles (Grand University, 2014). They mainly complied by monitoring faculty members’ SET performance to identify and solve problems with SET underperformance. The administrators’ monitoring SET performance involved reviewing or comparing faculty members’ SET ratings against their institutional benchmarks or cut-off scores and negative comments if any. When the HEIs’ QA officers sent SET results to individual faculty members (in sealed envelopes), they also sent to the administrators (usually the dean) a separate pile of individual members’ results:

After each semester, [QA officer] sent [SET results] to individual faculty members and a compilation [of SET results] to our faculty management board [deans or vice-deans]. That is the information our institution sends [us] to know our faculty members’ teaching activities. (I2A3)

Figure 11*Summary of Administrators' SET ToA*

Note. ToA: theory of action. SET: student evaluation of teaching. SET underperformance refers to the case of faculty members getting SET ratings below a benchmark or a cut-off score (e.g., under 3 out of 5) or receiving negative comments written in the SET responses. *n* indicates the number of participants who reported taking specific actions. The numbers of participants taking various approaches are not mutually exclusive.

After receiving the SET results from the QA officers, the administrators usually reviewed and checked if individual faculty members' SET ratings were up to their acceptable institutional benchmark. For example, one administrator reported that:

As a faculty leader, I only looked at the column about the overall SET performance of the faculty. The maximum score is 5. If a faculty member gets 3, that is not too bad... But if he has only 2 or 2.5, which is 50% of the maximum score according to the benchmark, we need to have some form of discussion. (I2A2)

The participants reported relatively different benchmark or cut-off scores, for example, “50% of the maximum score” (I2A2, I4A5) or “7 (out of 10)” (I1A2). This also reflects slight variations in four institutions’ cut-off scores compared to the upper institutional benchmark of “below 3 (out of 5)” (Grand University, 2014, p. 26). Besides, some administrators reported examining students’ written comments to spot faculty members’ disciplinary issues:

We read the students’ written comments because we need to scan them to see which faculty member received hash comments from students. For example, some faculty members almost did not teach students anything but always required students to do self-study and examinations from this channel. Also, by reading them [students’ comments], we learned that some faculty members skipped two classes but did not provide make-up classes. (I4A1)

In case QA officers provided an average of SET ratings across departments or faculties of the institution, administrators compared the SET ratings across different groups to form a relative judgement of SET performance across other groups. For example, one administrator commented: “I compare our faculty average scores with that of our institution. I found that our average usually is higher, meaning students usually highly appreciate our faculty's faculty members” (I4A3).

Disengaging Approach. Although the administrators complied with the SET policy requirements, they disengaged from using SET when there was no problem with SET underperformance. Eight administrators reported that they did not use SET for improvement or any purposes unless there were faculty SET-underperformance problems. This means that the administrators only used SET to monitor unusual circumstances such as faculty members’ SET ratings under the benchmarks; otherwise, they no longer needed it. As one administrator commented: “I *only* looked to see if anything faculty members’ scores fell below 7” (I1A2). The administrator used the word “only” to emphasise that her main tasks were to monitor faculty members’ SET performance rather than improve their performance. She reasoned that using SET for improvement was mainly the faculty members’ responsibility: “faculty members need to look at it [underrated SET statements] to modify it [teaching] *by themselves*, and it is important for their self-reflection and modification” (I2A1).

Some administrators found that SET underperformance was quite unusual at their institutions. Due to the rarity of having faculty members with SET problems in some faculties, some administrators admitted that they did not necessarily have anything to do with SET data, such as warning, talking to the faculty members or using SET for improvement:

I do not have much work to do... I do not have anything to warn our faculty members.
(I2A1)

To be honest, I do not think this is a big problem because four out of five means very good, and there is no reason to have a private talk with these faculty members. (I4A5)

Problem-Solving Approach. When there were problems with faculty members' SET underperformance, the administrators adopted different problem-solving strategies. While nine administrators reported unilateral and harmony-oriented ones, only one chose a shared problem-solving strategy.

Unilateral and Harmony-Oriented Problem Solving. Nine administrators reported a unilateral and harmony-oriented approach to solving the SET problems. The unilateral approach means various strategies to solve problems were used independently by the administrators without involving faculty members and students in the decision making. The harmony-oriented approach means diplomatic strategies such as avoiding, pacifying or stabilising to avoid potential conflicts or control a difficult situation without upsetting others (see Table 5).

Table 5
Examples of Administrators' Pacifying Approaches

Strategies	Goals	Examples
Avoiding	to prevent potential conflicts	<p>"inform but do not interfere with the faculty members."</p> <p>"try not to criticise or give judgmental feedback to the faculty members."</p> <p>"ask the underperformed faculty members to postpone teaching for a semester."</p> <p>"change the faculty members to another class."</p> <p>"avoid meeting students' demands [for more lenient study]."</p>
Pacifying	to control a difficult situation without upsetting others.	<p>"remind the teacher gently."</p> <p>"ask a group leader to talk to the faculty members with SET underperformance."</p> <p>"state general reminders in a faculty meeting, for example."</p> <p>"some faculty members need to pay attention to teaching."</p>
Stabilising	to find and fix problems between teachers and students	<p>"allow the faculty member to explain their SET results."</p> <p>"check what criteria the faculty member was underrated."</p> <p>"checking students about their feedback."</p> <p>"ask them some questions like 'what do you think?', or 'what change would you make?'"</p>

Note. SET: student evaluation of teaching

First, some administrators avoided solving SET underperformance in to prevent potential conflicts. One way to prevent future conflicts between faculty members and students

was to postpone teaching specific courses or change classes for underperforming faculty members. Also, some administrators tried to stay away from criticising or interfering with faculty teaching. One administrator avoided discussing SET's negative results with the underperforming faculty members and postponed their teaching for a semester. He reasoned that the faculty member's SET underperformance was justifiable:

For example, a faculty member was underperformed at SET, but I know for sure that she has good professional capacity. I know that because my students are in the business discipline, so they don't like it when a faculty member gives in-depth knowledge about information technology. That is why I did not criticise the faculty member. After that, I told the faculty member that she could postpone teaching the subject for a semester, and I almost did not mention anything about it. (I1A1)

Second, some administrators chose to pacify faculty members by using diplomatic actions to prevent the potential unpleasantness of negative feelings. One example was administrators addressing SET underperformance by giving harmless comments in a faculty meeting. The innocuous comments acted as a general warning for all faculty members while causing no harm to specific faculty members:

So I usually only stated general feedback in meetings with faculty members. For example, I would say, "one general characteristic of our faculty is that some of our faculty members talk too much and have few activities to foster students' critical thinking skills." (I2A4)

This example shows that the administrator avoided mentioning the specific faculty member who had low scores and only stated the general feedback during a whole faculty meeting. Even when some administrators decided to have a private talk with SET-underperforming faculty, they only provided gentle reminders instead of in-depth discussions with the individual faculty members: "I remind faculty members who receive below 4 in two conservative years something like "you need to consider if there is any problem." (I4A1)

Third, some administrators chose to stabilise SET underperformance by finding and fixing the problems between faculty members and students. This action generally started with administrators questioning faculty members or students to get more information about the incidents of SET underperformance. Then the administrators tried to fix the faculty SET underperformance by "discussing criteria where faculty members' scores were low" to "finding common ground" between faculty members and students (I4A3, I2A3, I4A4). One administrator commented:

I listen to both sides [students and faculty members] or verify the feedback. Sometimes students say this, but we will find out if it is true when we teach them. For example,

some students may say that faculty members' delivery made it difficult to understand or that a faculty member was too strict. Once, I had to change class to another faculty member; she told me she could not teach the same class over two semesters. I also listened to the students to find valid reasons for necessary changes. (I4A4)

Although the administrators asked for opinions for their faculty members and students, they usually talked separately with the two groups and made their decisions unilaterally:

I exchanged with the faculty member to find out what the problem was, and then there was an exchange with students to see how students thought so that I found out how I could help solve the problem of faculty members and students. (I2A3)

Overall, the administrators mostly solved faculty SET-underperformance problems by avoiding, pacifying, and stabilising actions. While these actions may help solve the problem immediately, they did not directly engage faculty and students in making decisions related to SET problem-solving.

Collaborative Problem Solving. Among 10 administrators who mentioned problem solving, only one used a collaborative approach which involved administrators working together with faculty members and students to solve the problem of SET underperformance. This joint strategy started with the administrator having an individual meeting with the faculty members:

We consider the reasons for student feedback. We would call and check whether it is true or not... If the faculty member has a problem, our leadership team will meet, and we will allow faculty members to explain if they feel the same way. After that, we would find the reasons for them to improve. (I1A3)

Not only was the administrator involved in discussing SET issues with faculty members, but she also conducted the class observation of teaching to help faculty members find evidence and possible strategies for teaching improvement:

We would first observe the class before organising a class meeting with the students. The purpose of the class observation of teaching and meeting with students is to communicate between faculty members and students more effectively... So in most cases, after the class observation of teaching, we would discuss if the faculty member has any problem in terms of teaching. (I1A3)

Such an approach to SET problem solving was joint because it also involved two-way communication between administrators, faculty members and students. Also, this approach inquired into problems with student learning rather than superficial problems between faculty members and students. However, it was a rare approach since only one among 18

administrators shared the practice of using SET data for a follow-up inquiry into teaching and learning.

Constraints on Administrators' Approaches

Overall, the SET policy's (Grand University, 2014) expected procedures and roles and administrators' perceptions influenced their complying and disengaging approaches to SET. Besides, many administrators' need to show respect to colleagues and their faculties' limited conditions for shared learning hindered their collaborative approaches to problem-solving.

SET Policy and Administrators' Perceptions. One key constraint that drove the administrators' compliance with SET by monitoring SET underperformance was their perception of management roles. One reason for the administrators' perceived role was that the annual faculty evaluation policy (Grand University, 2014) requires them to approve faculty members' self-evaluation reports. Thus, monitoring how well faculty members scored in SET helped them fulfil their role as a moderator and an evaluator during annual evaluations. Another possible reason for the administrators' perceived roles was that the SET policy indicates that a faculty member's "direct supervisor" must solve problems with SET underperformance (Grand University, 2014, p. 26). As administrators were direct supervisors of faculty members, they considered solving faculty SET underperformance as their role. As one administrator stated, in the case of faculty members underperforming at SET or having negative comments from students, he would discuss within the faculty board of management (i.e., faculty deans, vice deans) to consider the best way to solve the problem:

If there is a teacher who has [SET scores under expectation] or some negative [comments from students], we [faculty board of management] will discuss it if we should discuss it directly with that faculty member. We may assign one to be directly involved in talking to that teacher – only the two of them know [about SET underperformance] and let [the underperforming faculty member] solve the problem. (I2A2)

SET Policy Ambiguity. One key constraint that caused administrators to disengage from using SET for improvement was the SET policy ambiguity. First, the key SET policy document (Grand University, 2014) did not indicate the role of administrators in using SET for faculty learning and improvement. The appendix of the SET policy document (Grand University, 2014) indicates SET results should be sent to administrators only in the case that faculty members receive SET scores below the predetermined benchmark over a relatively long span of time (2 years). It is stated in an appendix of the SET policy document that:

If a faculty member's SET score falls below 3 (out of 5) for the same course in 2 consecutive years, the *leader of a unit* [thủ trưởng đơn vị] should *consider* [xem xét]

forwarding those results to the *direct supervisor* [cấp trên quản lý trực tiếp của giảng viên] [emphasis added] of the faculty members. (Grand University, 2014, p. 26)

In the policy indication above, the role of administrators is mentioned relatively briefly within a few sentences in the appendices, not on the main SET policy sections. Besides the nonidentification of the administrators' roles in the main procedure section, the SET policy document also uses vague language expressions. For example, it implicitly rather explicitly indicates that "the leader of a unit" and the "direct supervisor" refers to the HEIs' rectors and administrators. The ideas about the faculty members' SET underperformance "for the same course in two consecutive years" seem impractical since there may be many changes in the faculty members and their teaching within 2 years.

Furthermore, the Grand University (2014) SET guidelines do not specify how administrators must solve faculty members' SET underperformance. Under the instruction on how to deal with SET results, supervisors (i.e., administrators) must "discuss the modification of course content" or "discuss the modification of assessment content or methods" when scores of course evaluation related to these question items are below 3 (out of 5) in 2 or 3 consecutive years (p. 26). However, regarding problems with faculty members' SET underperformance, the policy document suggests that the rector "consider forwarding" the results to the administrators. The use of "consider" makes the statement tentative, which indicates that it is up to the rectors and administrators to decide what to do with the SET underperformance. On the one hand, the policy document may imply that administrators could apply the same strategy of discussing the modification of teaching activities similarly to solving course-content or learning-assessment problems.

On the other hand, it conveys an explicit message on the same page of the document that "the SET results have not yet been used as a basis for rewards or sanctions on faculty members." These instructions reflect the Grand University's policy intention to use SET as mainly a tool for an institution's reflections or warnings to faculty members, rather than making any critical personnel decisions related to individual faculty members. The hesitance to use SET for personnel decision making is understandable since SET is a relatively new practice in the Vietnamese HEI context. Due to the Vietnamese teacher-centred tradition, SET policymakers at the upper institutional level (Grand University, 2014) possibly use hedging or tentative language to avoid potential resistance from SET-implementing agents at the lower institutional level.

Administrators' Perceived Low SET Usefulness. Another key constraint explaining why the administrators disengaged from using SET for improvement was their perception that SET had low reliability. The administrators raised concerns over the reliability of SET data because of "inappropriate SET items," "low response rates," and especially "students' lack of

a serious attitude towards SET responding,” and students who “tick randomly on the SET sheet” (I1A2, I2A1, I2A2, I2A3, I4A and I4A5). Some administrators (I1A2, I1A3, I2A3) also raised critical concerns over the QA officers’ procedures of collecting feedback. Administrator I2A3 observed that only 10% of his students read the SET form carefully and wrote their comments on their teachers’ performance, and he found the rest of the students played it safe by ticking the scores above the average for their teachers. Further explanation was reported as follows:

QA officers in our institution have a minimal number of officers while there are numerous classes to be in charge, so they had to ask the administrative assistants in each faculty to help them collect SET responses. We only have one full-time administrative assistant who must take charge of a large number of classes, so he had to ask for the student monitor of each class to help. However, students were not aware of their rights and responsibilities related to SET, so there were cases when students ticked randomly on the SET sheets to get it done or when several students ticked for those who were absent. Due to that way of collecting SET responses, the SET results cannot be 100% accurate. (I2A3)

Besides, administrators commonly shared a trend that their teachers’ average scores were from “not too high or not too low” to “generally very high” or “seldom seen any negative feedback from students” (I4A2, I4A5, I5A1). Because most faculty members’ SET scores were rated within the acceptable level, many administrators found SET scores meaningless or no use differentiating teaching effectiveness:

The majority of SET scores are over 4... we have not got any low SET feedback cases. (I2A1)

We had a discussion with these faculty members, but it was scarce – it happened once since I started this position nearly 3 years ago, there are only one, two maximum three cases like that. (I2A2)

As faculty [we/ I] observe for the past 4 or 5 years, we have not got any faculty members who need private discussion. (I2A3)

We have not got any faculty members in our faculty who received scores below average. (I2A4)

It [faculty members’ SET average score] lies within the acceptable level, for example, from 3.4 to 4 out of 4, so it does not inform anything. (I4A2)

Due to the perceived limited SET usefulness, many administrators regarded SET as a “formalistic” or “bureaucratic” process (I2A1, I2A3, I5A1). Thus, they disregarded the SET unless there was faculty SET underperformance.

Perceived Need to Respect Colleagues. One constraint that influenced many administrators’ problem-solving approach, especially their harmony-oriented strategies to faculty SET underperformance, was their perceived expectation of respect for colleagues. While being leaders, administrators also positioned themselves as colleagues among teachers or faculty members. They were aware that faculty members, especially the senior members, tended to “have their own opinion about their teaching” (I4A1) or sometimes be “very conservative” and “reluctant to change” (I3A1). Whether or not the respect for senior teachers is typical for Vietnamese or Asian culture, the administrators admitted that they could “mainly observe SET scores” but not “interfere much” with individual teachers’ SET scores. Also, due to the expected certain level of autonomy over higher education teaching, teachers are expected to “look at it [their SET results] *by themselves* [emphasis added], and it is important for their self-reflection and modification” (I1A2). These perceptions of higher education teacher autonomy were supported by the administrators’ perceived lack of “right and responsibility” (I2A3) to deal with or utilise SET for accountability and improvement purposes. All these perceptions taken together have cemented a light approach to faculty SET underperformance among administrators. Hence, they chose not to act as managers or SET evaluators but as protective colleagues when dealing with faculty SET underperformance.

Limited Conditions for Shared Learning. Many administrators found unfavourable conditions for shared learning, such as their limited autonomy to enact their roles of promoting teaching and learning improvement. The administrators’ light attention to SET underperformance was congruent with their perceived lack of authority to deal with faculty members’ problems. The administrators must use SET results to improve teaching quality, but they did not have the “power, autonomy and resources” available for them to do so (I2A1). The administrators’ perceived lack of authority shows a potentially significant gap in the institutional conditions for quality improvement. There was a gap in document policies that generically set expectations for administrators and institutional leaders to use SET data to improve the training quality while giving them no specific guidelines, incentives, or improvement mechanisms. Some administrators perceived that they did not have the right and responsibility to use SET for incentives and rewards; therefore, they disengaged from using SET for teaching improvement:

So generally, it is a means to observe, and if something unusual happens, we will solve it. Otherwise, the results are almost used for no particular purpose in terms of standard quality assurance. When we evaluate or grade faculty members, we do not use it. It is

not like a business that can use such scores to pay salary or rewards; we do not have such rights and responsibilities. (I2A3)

One possible reason for the administrators' light attention to solving SET underperformance was their perception that faculty had autonomy over their teaching content usually controlled by a test bank. As the prescribed teaching content was relatively heavy and fixed, making it hard for faculty members to modify the content or to improve interaction with students:

How can faculty members increase their interaction with students? That is quite hard. Why hard? Because the MoET controls the challenging teaching content, our institution and faculty cannot modify it... Regarding the specialised course, we can alter no more than 20% of the prescribed content, so all the modifications are restricted to the subject title and a small part of the content (I2A4).

Also, a few administrators affirmed that their faculty teachers would have difficulty removing any content because of the heavy course content within a restricted time frame and the controlling mechanism. Thus, they found teaching and learning improvements were generally challenging, which hindered the effective use of SET. As the administrator commented:

Our institution controls the teaching content by using a test bank to prevent faculty members from reducing or adding course content by themselves. The effective delivery of the teaching content depends on individual teachers' pedagogical skills and the professional capacity to organise a lesson that covers all the contents and provides independent learning activities for students – that is such a problem. (I2A4)

In contrast to the major unilateral approach to problem solving, one out of 18 administrators (I1A3) chose a joint approach in which she directly worked with faculty members and students to inquire into problems with teaching and learning. An enabler of the administrator's joint approach to solving SET underperformance was the favourable faculty conditions for shared professional learning and teaching among faculty members, such as an institution-wide value for student feedback and the faculty culture of shared learning.

The institution-wide value for student feedback was evident at HEI 1, where all three administrators (I1A1, I1A2, I1A3) were working. As Administrator I1A1 pointed out, students in his institution were more demanding of their teachers because they generally paid much higher tuition fees than other public HEIs. Thus, faculty members in his institutions needed to "compromise" or "satisfy students' demands" (I1A1). HEI 1 was also the only one that used SET results for granting financial rewards to faculty members (I1Q1). The use of SET results for financial bonuses made SET become "more important than ever before" (I1A1).

However, while three administrators mentioned institutional value for SET, only one of them (I1A3) reported a joint approach to solving faculty SET underperformance. One possible reason for this distinction was the faculty norm of shared learning and teaching. In this faculty, there were frequent peer evaluations of teaching to support faculty members who were young or who had started teaching a new subject: “even when there are no problems with SET results, we still hold frequent evaluations of teaching and class meeting with students” (I1A3). Besides, this faculty also had various means to collect students’ reflections about teaching continuously:

In general, students can share their feedback [to us] via various means. For example, students can send notes to us at the end [of each class], or sometimes students file a proposal [to complain or change teachers] after 1–2 classes with them. The students in our faculty are highly democratic, so many teachers have got such things [candid feedback from students].

This shared approach to SET problem solving involves the administrator individually meeting the SET underperforming faculty members to listen to their viewpoints about the SET results: “I would allow teachers to explain what they think about student feedback before working out ways to improve the teaching and learning” (I3A1). Due to the faculty culture, Administrator I1A3 often integrated observation of teaching and student meeting when she solved with SET results:

When there is an [negative] opinion from students, we would first observe teaching before organising a class meeting with the students. The purpose of combining the class observation of teaching and meeting with students is to communicate between teachers and students more effectively.

Administrator I1A3 found that students did not typically reflect on teachers’ professional capacity, but the main problem was the communication between teachers and students. She observed that when students did not want to talk directly with the teachers to modify “this and that,” the distance or the gap between the teachers and students grew gradually, and students became very frustrated. That is why the administrator found it necessary to observe a class and talk to students after the class observation to find out if the teacher had any problems in terms of teaching. The feedback given to teachers for improvement would be based on the evidence from the observation and the students.

Consequences of Administrators’ Approaches

While the administrators’ complying and problem-solving approaches helped them fulfil their SET requirements, most were uncertain about faculty members’ use of SET for teaching and learning improvement.

SET Policy Requirements Fulfilled. The first consequence is that administrators used SET to monitor faculty performance. By reviewing individual faculty members' SET scores, administrators could identify possible problems of faculty members not meeting the predetermined SET benchmarks or receiving negative comments from students. Administrators generally reported that they fulfilled the SET requirements. Many administrators showed their task fulfilment by giving some sort of warning to the faculty members who were reported to have some problems with disciplines, such as "skipping class regularly" or "leaving class early" (I2A2, I4A1). These administrators thought that SET improved faculty members' accountability or increased their sense of responsibility [tinh thần trách nhiệm] towards teaching. One administrator affirmed that he discussed the problems with the SET low-rating members, which helped the members understand students' preferences.

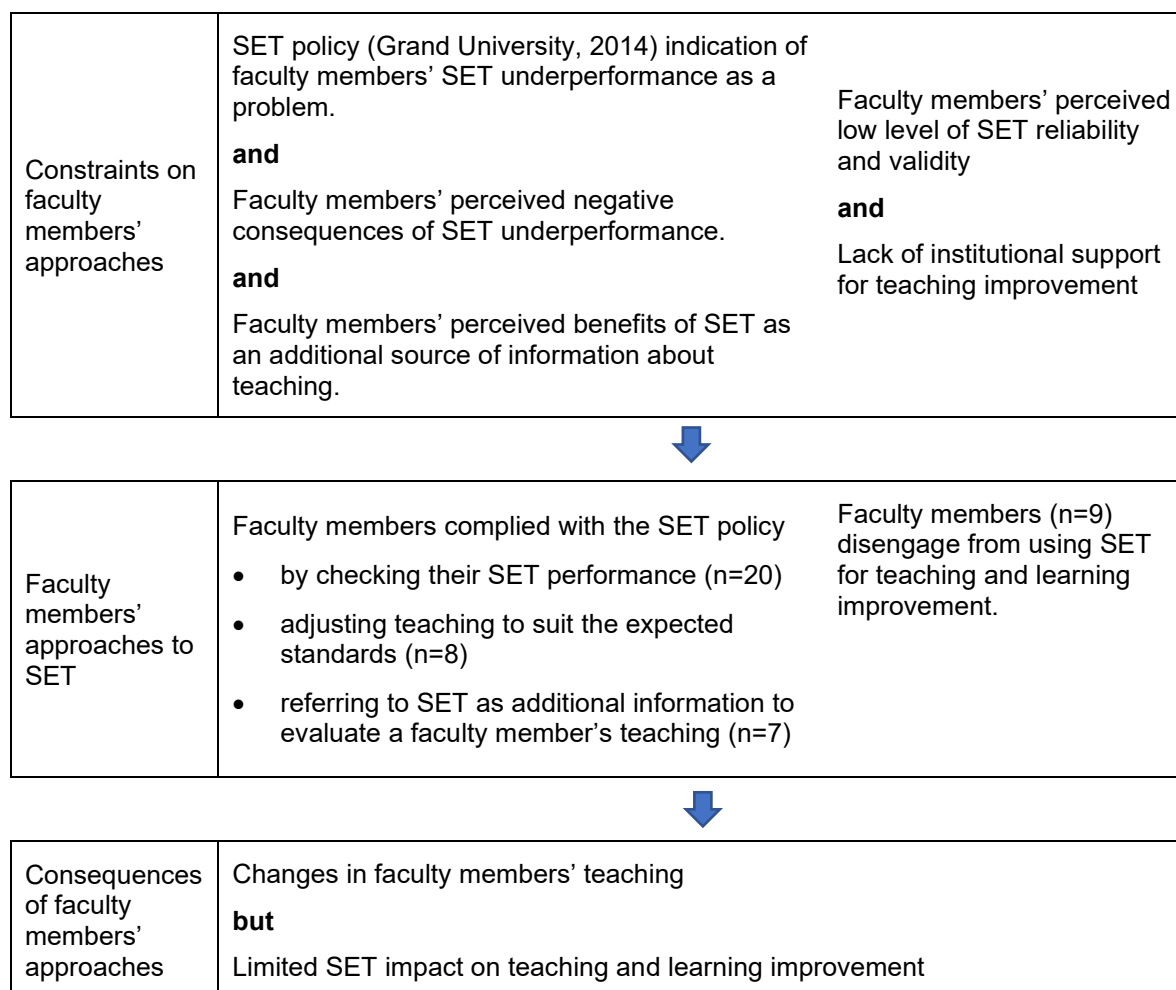
Administrators' Uncertainty About Teaching and Learning Improvement. While the strategies taken by the administrators appeared to solve the SET-underperformance problems, they did not necessarily help improve teaching and learning. The underperforming faculty members might then change their behaviours to be more disciplined with their class teaching or suit their students' preferences. However, such changes would be superficial because they were unrelated to student learning. Since most administrators did not use SET for further inquiry into student learning or make faculty-level teaching improvement plans, they were not sure about faculty members' use of SET for improvement. Instead, teaching improvement was dependent on individual faculty members' motivation and capacity: "the effective delivery of the teaching content depends on individual faculty members' pedagogical skills and the professional capacity to organise a lesson that covers all the contents and provides independent learning activities for students – that is such a problem!" (I2A4).

Faculty Members' SET ToA

The following sections present the faculty members' SET ToA: their approaches to SET, the constraints, and the consequences of these approaches (see Figure 12). Overall, all the faculty members complied with the SET policy (Grand University, 2014). Many faculty members used SET to adjust their teaching, but they disengaged from using SET for teaching and learning improvement.

Faculty Members' Approaches to SET

As Figure 12 illustrates, the administrators' main approaches were complying and disengaging. The following sections describe each of the approaches to SET taken by the faculty members.

Figure 12*Summary of Faculty Members' SET ToA*

Note. ToA: theory of action. SET: student evaluation of teaching. SET underperformance refers to the case of faculty members getting SET ratings below a benchmark or a cut-off score (e.g., under 3 out of 5) or receiving negative comments written in the SET responses. *n* indicates the number of participants who reported taking specific actions. The numbers of participants taking various approaches are not mutually exclusive.

Complying Approach. All the faculty members complied with the SET policy by checking and monitoring their SET performance (Grand University, 2015). Many faculty members used SET to adjust their teaching performance, and a couple of them used SET to evaluate other faculty members' teaching.

First, the faculty members checked their SET performance to identify if their overall SET scores were above the benchmark or the cut-off scores established at their institutions. For example, two of the faculty members commented:

It is okay if the result is like eight, nine or ten. (I1T1)

I care about the results overall. How would students rate me, and if it is ok... I usually only get the total results, a rating out of five. (I3T1)

Following the SET guidelines (Grand University 2015) that indicate faculty members' use of SET for teaching adjustment, eight faculty members based on SET results to modify their teaching to suit the SET expected standards. Some faculty members mentioned analysing SET ratings of individual criteria to draw out their strengths and weaknesses perceived by students. They looked carefully at the students' ratings for individual question items to identify which one the students rated high or low. In other words, faculty members used SET scores to identify students' preferences over certain aspects of their teaching or recognise their "strengths and weaknesses" through the lens of their students. For example, one faculty member described:

For example, there were 10 statements – I looked at the statements that got the lowest scores. Then concerning the lowest score, I considered if I had probably not done well in those statements, for example. I will look at the statements I got the highest rating scores and those with the lowest scores to see my strengths and weaknesses. I think it is a good channel for us to know how good our teaching is and to what degree our students were satisfied with us. (I1T6)

Similarly, many faculty members also checked students' written comments carefully. They appreciate written comments though such feedback was quite limited in quantity. Faculty members considered such written comments to reflect "students' thoughtfulness" and "indicate the area of improvement" for faculty members (I1T2). Faculty reflections on their students' perceptions of their teaching motivated them "strive to do the teaching job well" (I1T4, I3T2). Thus, these faculty members adjusted their teaching to adhere to SET criteria or student evaluation:

When we have such evaluation criteria, we can use them to review our own teaching, like "well, I have met these specific standards, or I need to do this or that to make my course more complete!" Based on the criteria list, I evaluated my own teaching and directed my teaching to adhere to the criteria requirements. (I2T2)

Notably, some faculty members reported that they only made a minor adjustment to adhere to SET criteria or to suit students' preference for learning:

For example, if a student says you need to increase your practical knowledge, or you need to create an opportunity for students to discuss more, I will base on the comments ... just adjust my teaching in the subsequent sessions. (I3T2)

Secondly, a couple of faculty members used SET results as a basis for evaluating other members. One faculty member reported that her administrator usually publicly praised those who had outstanding SET scores during the annual evaluation meeting. Thus, she used higher SET scores as a reference to evaluate and vote for faculty members for outstanding

titles of excellence. Another senior faculty member stated that his faculty assigned him to support junior faculty members. Thus, his administrator used formal SET and informal student feedback to recognise faculty members that needed help. By combining various sources, the senior faculty member and his administrator could provide support for the less experienced faculty members:

A female (junior) teacher said that the students did not understand, then our department head assigned that person to continue attending other people's classes and ask others (including me) to guide. I instructed her people on how to prepare lesson plans or techniques to deliver knowledge and manage a class. (I2T3)

Disengaging Approach. Nine faculty members reported their disengagement from using SET. Three faculty members indicated that they did not use SET to inquire into student learning, and six of them shared that they disregarded SET if they found no problems with SET results. The faculty members' disengaging approach was also linked to several inactions.

First, some faculty members refused to adjust their teaching following SET and they chose other means to inform their teaching. These faculty members only considered SET as an additional source of information about teaching. They thought the central goal of SET results was to check students' degree of satisfaction with them or identify possible problems with their teaching. For example, one faculty member commented:

SET data is just another source of information for reference... but I will not be affected much by those results unless it's too low; I will not be thinking about why it is like that. (I1T1)

People sent it via email, but nothing happened, so I did not pay much attention. If there is a problem, people will have to tell me, but if there is no problem, then well ... I do not use it at all. (I3T4)

I did not do anything [with SET results], just to know if students were satisfied with me. (I4T1)

Secondly, even when some faculty members received negative SET scores over a specific aspect of teaching, some faculty members did not base on SET to adjust their teaching because SET criteria did not match their belief about effective teaching. For example, one faculty member got an overall negative rating, especially concerning using technological equipment such as PowerPoint slides in teaching. He did not think that such use was conducive to teaching the subject and so decided to disregard the SET information:

I usually get the average score, not very good nor below-average level. Why so? One of the reasons for my average (not too high) score is that I do not often use modern

technology to support my teaching (which is one criterion in the evaluation list). Due to philosophy being a subject related to abstract reasonings, which limits the use of graphs and models, I rarely used it [PowerPoint slides]. This is why I got a lower score for that criterion. (I2T3)

Thirdly, some faculty members did not support an adjustment of teaching for the sake of maximising SET scores. For example, one faculty member believed that students would give her a higher rating if she were a little more relaxed or lenient. However, she refused to resort to the strategy of seeking to win students' hearts through leniency. She found it more important to be fair to all students, guaranteeing a certain standard of teaching and necessary discipline:

If I am a little more relaxed or lenient, students will rate me [higher]. But I do not care about that. I just do the things everyone is asked to do, and I will try to guarantee a standard. For example, one requirement is to have a strict attendance record because students who skip so many classes cannot take the test. If I am not strict about attendance records in my classes, isn't it very unfair to the other strict classes? If it were not for me to do things that were meant to win students' hearts, I would not care. I just did the right thing that needed to be done. (I4T5)

Other faculty members also expressed their disapproval of adjusting their teaching for fear of students' low ratings. They refused to adhere to SET criteria because they knew what worked best for their students, and they felt satisfied with their own commitment to the teaching job:

Well, it is not true that I adjust nor frantically make it better for the fear that they underestimate me. It's not that I will promote the aspects of my teaching that are judged useful by others. Such [ratings] do not mean anything to me. (I3T3)

Actually, students' comments are not too important to me... because no matter what good or bad comments or whatever, I still teach them with the same quality and degree of commitment. It means that I have tried my best to fulfil my responsibilities, and then I feel satisfied with what I have done. (I4T5)

Lastly, some faculty members chose other means or strategies to inform their teaching instead of using SET data. These faculty members adopted one strategy to communicate directly with students during class time. This helped them gather information about teaching and learning immediately, and thus they could modify their teaching in a timely way:

No, it [SET data] has no information, and I've never had [written] comments, so I don't know. Usually, when I come to class, I always talk with students, and if something goes

wrong, I adjust immediately... In my class, I always had conversations from the beginning [of the course]. During the learning process, I always asked [my students] if there was any problem. (I7T1)

Some faculty members determined student learning by observing the ways students responded to their questions or problems. The ways students engaged in the discussion, their genuine excitement during class or the level of critical thinking their answers displayed gave the faculty members a more in-depth evaluation of their student learning. One faculty member admitted that “student engagement and thinking ability is what motivates me most to evaluate my teaching” (I2T1). Some faculty members also evaluated teaching by monitoring students’ final exam scores. When faculty members used rubrics for assessing student learning, they felt more confident in grasping the degree of effectiveness their teaching had on their student outcomes:

I just used it [SET] for reference, but what I evaluate for myself, I personally based on the degree of student [learning]. At the end of the term, we assess whether the students meet the subject's requirements because each subject has its own [assessment] rubrics. (I2T1)

Constraints on Faculty Members’ Approaches

Several constraints explaining the faculty members’ complying and disengaging SET were the SET policy indication and faculty members’ perceptions of SET. The faculty members’ complying approach was explained by the SET policy indication and the participants’ perceived consequences of SET underperformance and SET benefits. In contrast, the faculty members’ disengaging approach was explained by their perception of SET’s low reliability, validity, and their institutional lack of support for teaching improvement (see Figure 12).

SET Policy Indication and Faculty Members’ Perceived Negative Consequences of SET Underperformance. As indicated in the SET policy document (Grand University, 2014), faculty members were expected to have their SET scores above a specific benchmark. Students were also given some blank space to write additional comments on their faculty members. Concerning these policy requirements, faculty perceived that they might face some consequences if their SET results were not acceptable. First, if faculty members scored less than the benchmark, they may have needed to have “some kind of discussion” with their administrators. They may also have been deemed to be not completing their teaching tasks well. If faculty members scored well in SET, they not only had a better sense of self but, in some cases, they could also have been rewarded with titles of excellence or a financial bonus.

As mentioned in the previous section, though the Grand University suggested 3 out of 5 as the benchmark, participating institutions had different benchmarks for acceptable ratings on SET. For example, HEI 2 set “50% of the maximum score” as the benchmark, while HEI 1 chose “7 out of 10” as the benchmark (I1A2, I2A2). No matter what benchmark each institution chose, the faculty member participants generally perceived a vital goal of achieving the minimum benchmark scores. Thus, all of the faculty members checked the overall SET rating scores first to identify if they had any problems with SET.

Faculty Members’ Perceived Benefits of SET. Despite concerns about SET reliability and validity, half of the faculty members expressed a particular appreciation for SET. Some of them found SET criteria to be relatively specific and, in some ways, relevant to their teaching activities.

First, SET criteria were considered a relatively reasonable standard for faculty members’ evaluation of teaching. One of the faculty members emphasised that this introduction of criteria for teaching was helpful since, previously, there had not been a shared teaching standard among faculty members. Before the introduction of SET, most faculty members had to evaluate their teaching based on their accumulated individual teaching experience or through their peer observation of teaching. However, since the introduction of SET practice, a list of predetermined criteria for evaluating teaching was available to faculty members. One faculty member attributed the meaningfulness of SET to its established criteria for his self-reflection on teaching:

We do not have predetermined teaching standards sent to us, so we were dependent on our personal experience in teaching (for self-evaluation). When we have such evaluation criteria, we can use them to review our teaching like “well, I have met these specific standards, or I need to do this or that to make my course more complete!” (I2T2)

Similarly, some other faculty members (e.g., I1T6, I3T2) attributed a certain level of usefulness to SET due to its specific list of teaching activities related to faculty members’ technological application, professional knowledge, instructional methods, or content delivery. Besides, a few faculty members noticed a level of authenticity and meaningfulness of their students’ written comments because they could “point out the problems that I need to work on,” and they reflected “the seriousness of the SET respondents” (I3T2).

Faculty Members’ Perceived Low Level of SET Validity and Reliability. The faculty members’ disengagement from SET was explained by their perceived issues with the SET validity and reliability. SET validity and reliability refer to the accuracy and consistency of SET as a measurement of the faculty member or teaching effectiveness; in other words, effective faculty members are rated with high SET scores and vice versa. Many faculty members

wondered if SET results could be relevant, accurate and consistent predictors of their teaching effectiveness. One faculty member commented:

“It is funny to see that within an intake, I taught the same courses and same levels for two different classes... However, one class rated me an overall of 8 while another rated me 9.5 or 10... I do not understand what the difference is because I taught two classes with enthusiasm and nothing different... I do not know precisely if it [SET] was accurate. (I1T2)

Firstly, except for HEI 3, where QA officers designed their own SET forms, the rest of the participating institutions adopted a unified version of the SET survey. Although these member institutions were different in majors, disciplines, teaching and learning focus, they shared the same SET forms across institutions and faculties. Some faculty members thought that the use of one-size-fits-all forms caused SET criteria to be irrelevant in some cases. For example, one faculty member of philosophy found that the criteria related to the use of technological equipment were not conducive to his teaching the subject because the use of equipment may hinder students' understanding of abstract ideas:

Criteria like “[faculty members]’s effective use of teaching equipment” or using technology to illustrate teaching content might not be useful for learning, especially under the subject of philosophy where technologically designed models might limit students' authentic understanding of the subject matter... Because these SET criteria are not relevant, so I do not care much about them. (I2T3)

Similarly, other faculty members found some SET question items inappropriate, such as one about faculty members' adherence to the course outline, as students did not have adequate understanding to evaluate such criteria. They thought that students could only evaluate what is obvious to them, such as “if the faculty members are punctual or comprehensible” (I4T2). As SET was found to be inadequate to reflect the essence of teaching, many faculty members disregarded SET as an accurate tool that provided genuine data about teaching. One faculty member suggested SET evaluation tools should not be collectively used in all institutions but SET criteria should be specifically designed and discipline-oriented (I2T1).

Moreover, some faculty members were concerned about the improper SET collection procedures and students' attitudes to SET responses. For example, one faculty member found that QA officers did not provide students with specific instructions on how to respond to SET during SET form distribution. The staff who collect SET surveys were claimed to be “superficial” (phiên phiến) and to “throw a SET pile” to the students and ask them to return them upon completion (I1T1). A similar attitude was observed among administrative officers

who oversaw collecting SET responses at another institution: “I heard some administrative officers [who are in charge of collecting SET responses] said “pew... it’s so lucky that we do not have to collect student evaluations this term” (I2T1). The faculty member found that the staff in charge only did what they were assigned to do to “get the work done” (làm cho xong, làm hết trách nhiệm). The faculty member’s account seems to show that some staff commonly viewed SET as a time-consuming administrative task that did not practically help solve any problems.

Corresponding to the neglectful attitude of QA officers in the SET-survey collection were faculty members’ concerns over students’ lack of serious attitudes to SET responses. Students were observed by faculty members in various institutions to “tick randomly on SET forms,” “respond in a rushed manner,” or be “too lazy to provide written feedback” (I1T2, I2T2, I2T3 and I3T1). For example, one faculty member commented:

That students' random and fake rating (điền bừa điền ảo) is quite common... Random evaluation does not mean that students tick all the good ratings, but they tick in a way that shows differentiation in responses... There are too many criteria, and during class hours, there are so much knowledge students need to cram; that’s why they may tick randomly in response to SET. (I1T2)

Besides, many faculty members found their students unable to make fair judgements of teaching. They believed students usually based their ratings on their subjective opinions and emotions about faculty members. These faculty members found students generally preferred lenient and easy-going faculty members rather than those who were strict, even though the latter had a more profound knowledge of the subject:

I think student evaluation is mostly sentimental (connected to emotions rather than reason)... Some teaching staff deliver the lessons with a deep level of professional knowledge, but they do not know how to make it an interesting, humorous, or lively atmosphere in the classroom, so the faculty members do not get good feedback from their students. Students may prefer other faculty members who can tell jokes to entertain them, even though they have a relatively superficial knowledge of the subject. (I2T2)

I used to be a student at this university, so I already know [how students generally responded to SET]. Sometimes I evaluate it very quickly, and if I like the faculty members, I rate them 4 or 5 (the highest rate) ... As far as I'm concerned, students usually register for classes of lenient and easy-going faculty members who typically give higher learning results. Some faculty members teach very well but demanding, so students do not like them. (I3T1)

If I am a little more relaxed or lenient, students will vote for me [with a higher rating]. (I4T5)

Another reason why students' ratings might not be reliable was the perceived low quantity and quality of student responses in some cases. Sometimes the students' response rate was "too low to be convincing" (I4T1). Some faculty members also believed that indifferent-to-learning students tended to respond irresponsibly to the SET survey. For example, one faculty member found that:

Students who frequently missed classes and did not read the course material ended up not understanding the lectures, regardless of how interesting they may be. Such students found learning very difficult. Thus, when evaluating lecturers, they ticked randomly on the given choices or judged the lecturers ineffective. That is what makes the value of SET practice reduced. (I2T3)

Another faculty member found similar issues with irresponsible students responding to SET ratings. He thought SET results would give an inaccurate teaching evaluation when students were not fair in judging faculty members. He commented that:

Democracy depends on the level of understanding the people have... The problem is the students [generally have a low] sense of responsibility for evaluation. [I wonder if] they are fair or impartial, or they are emotional, giving inaccurate ratings for the faculty members they do not like. (I2T2)

Some faculty members did not trust SET scores because of students' perceived low level of responsibility and their tendency to give subjective SET ratings based on their feelings about the faculty members. For example, one faculty member commented: "I usually get quite high SET rating scores, but I do not believe in it very much. Maybe students like me, or perhaps they are afraid that I will continue teaching next year" (I2T1). Another faculty member also commented on her lack of trust in her given SET scores: "there must be something wrong when the whole class of students all rated me with one or two [the lowest score out of five] - I am pretty sure that they misunderstood the rating order" (I1T1).

Lack of Institutional Support for Teaching Improvement. One constraint that explains the faculty members' disengagement from using SET data for teaching improvement is their perceived lack of a supportive environment. Although more than half of the faculty members (n=13) valued teaching more than research, many found their institution did not share the same priority over teaching. For example, in the faculty evaluation system, teaching is not equally important as research. Faculty members did not see an institutional emphasis on teaching: "the achievement in doing research is more important than the others because it is clearly stated in the document" (I1T7). The lack of institutional recognition of teaching

reduced faculty members' motivation to improve their teaching. Besides the lack of institutional support, the institutionally shared teaching values also influenced faculty members' attitudes to SET. One of the faculty members commented:

If you really want an institution that is really committed to the development of teaching and learning, the faculty members should be the priority. But I don't think that we... I think the main difficulty for me is that I don't have an environment that recognises my effort to improve teaching and learning. I really care about the students' learning and achievement, but not many people have the same value. (I1T3)

Another hindrance to using SET for teaching improvements originated from the lack of culture for continuous learning among the faculty members. Faculty members found that many of their colleagues consider themselves "the master of knowledge," not "continuous learners." One faculty member commented that using SET for teaching improvement only works for those who are "young" and "have the right attitude and enthusiasm in teaching and learning," but not for the older group of faculty members who were "highly conservative, idle, experience-oriented, and have difficulty in English." Besides, the conditions of low income combined with the heavy workload and outdated teaching curriculum placed a great deal of pressure on faculty members:

I think faculty members' continuous learning is still sporadic in our university context. It is tough for faculty members to do so, except for some faculty members who are really passionate about the job... Many teaching staff do not really care about the evaluation criteria that tell faculty members what to prepare for a class. (I2T2)

Consequences of Faculty Members' Approaches

Overall, the faculty members' complying and disengaging approaches to SET had a limited impact on faculty members' teaching improvements. Even when faculty members used SET for making changes in their teaching, this did not necessarily improve student learning.

Changes in Faculty Members' Teaching. SET resulted in changes in some faculty members' teaching. Some faculty members found SET as an additional source of information to modify their teaching. While SET scores were not adequate for predicting teaching effectiveness, they were an additional source of information about teaching. For some faculty members, positive SET scores and comments reflected students' overall satisfaction and appreciation, and they found students' appreciation a "special gift" for their devotion to teaching (I1T5). For other faculty members, SET ratings for individual question items about faculty member instructional performance somehow reveal faculty members' strengths and weaknesses in the eyes of their students. Some faculty members found that the exceptionally high or low scores, to some extent, informed them of the degree of students' satisfaction over

specific aspects of their teaching, which might be helpful for them to respond to students' demands.

Limited SET impact on Teaching and Learning Improvement. However, SET did not have much impact on other faculty members. Thus, they disregarded SET when they had no problems with SET results. To some extent, the SET results informed faculty members about the degree of students' preference for their teaching. However, SET results were sometimes ambiguous, so faculty members did not make sense of the data. For example, some faculty members commented:

Last year, a student wrote some notes [for me] ... I really wanted to meet with the student to ask about their comment on my "being outdated" – Do they mean that the course is outdated, or I am obsolete compared to other faculty members? (I1T4)

Some students just wrote, "everything is fine." [It is generic], so I do not understand exactly what they liked about my class so that I can maintain [my strengths]. (I4T1)

Some faculty members commented that SET was inadequate for instructional improvement. SET results were mostly generic numbers without specific in-depth information about how to improve. Unless further information or training were provided, faculty members would not know how to improve their teaching:

[SET scores] are just general numbers, without in-depth or detailed information. I know that I must work hard to improve the class atmosphere, but I do not know which direction to go, learning and playing or helping students get high scores. It does not help me improve my teaching. (I1T1)

Furthermore, SET results were reported to come too late to enable changes. To bring about actual changes, faculty members emphasised that they needed to receive timely feedback from students:

Also, another issue is that the feedback came to me too late, so I could not change any when teaching a new class. (I1T2)

My expectation, if any, is to get timely feedback from students. For example, if students have any feedback about the lecture content, I would like to know as soon as possible to explain it more clearly to students so that they do not misunderstand. (I2T3)

Overall, SET had various and inconsistent impacts on faculty teaching improvement. Besides, faculty use of SET was limited in modifying some aspects of their teaching to suit students' expectations. This modification tended to focus more on faculty content delivery or teaching techniques than on student learning. Thus, faculty members' use of SET for teaching modification did not necessarily improve student learning.

Concluding remarks

Overall, the three stakeholders (i.e., QA officers, administrators, and faculty members) took various approaches to SET: complying, solving problems, and disengaging from using SET for learning and improvement. These findings confirm that the three purposes of using SET for accountability, improvement, and learning (i.e., transformation) were evident in the participants' approaches to SET. First, the most dominant ones were the accountability-oriented approaches (Lonsdale, 1998) to SET, which included complying, monitoring, and checking. The participants took these approaches to satisfy the accreditation, managerial and administrative purposes of SET from the accreditation bodies and the parent university policy requirement (Grand University, 2014). These approaches led to the fulfilment of SET requirements from external bodies. Second, the improvement-oriented approaches (Lonsdale, 1998) to SET, which included solving the SET problem, were present. Some participants tried to increase the usefulness of SET responses, solve problems of SET underperformance, and change their teaching following SET criteria and requirements. While these approaches increased attention to SET, they did not necessarily improve faculty members' teaching and student learning. Last, the learning-oriented approaches (Dahler-Larsen, 2009) to SET, which included joint and direct problem solving, were the least common ones. This approach was only evident in one case where the administrator jointly engaged faculty members and students in an inquiry into learning and teaching. This approach resulted in both teaching and learning improvements. However, this approach was limited to only one case where the whole faculty members and administrators had shared effort to improve teaching and student learning.

Chapter 5: Voting Evaluation Theories of Action

This chapter presents findings of the participants' theories of action concerning voting evaluation, a practice that involves faculty members voting among themselves for titles of excellence at the departmental and faculty levels. This chapter starts with the voting evaluation ToA of 18 administrators, followed by that of the 20 faculty members.

Administrators' Voting Evaluation ToA

The following sections describe the administrators' voting evaluation ToA, which is the administrators' approaches, the constraints and the consequences of these approaches (see Figure 13). Overall, all administrators complied with the evaluation roles and procedures outlined in the Grand University (2015) policy, but many disengaged from using the evaluation for individual faculty learning and improvement. The administrators had various strategies to solve the voting evaluation implementation problems.

Administrators' Approaches to Voting Evaluation

The following sections describe three key approaches to voting evaluation taken by the administrators: complying, disengaging and problem-solving approaches as shown in Figure 13.

Complying Approach. Administrators generally complied with the institutional voting evaluation policy document (Grand University, 2015), which indicates that administrators must follow the basic voting evaluation procedures as follows:



- Organising a whole faculty evaluation meeting
- Moderating faculty members to evaluate selves and others during the meeting
- Having faculty members nominate and vote emulation titles for faculty and individuals by votes of confidence (e.g., by ticking paper ballots.)
- Sending reports and proposals for faculty and individual achievements to the member HEIs' Emulation Council. (article 31)

All administrators (n=18) reported following the voting procedures as required by the Grand University's (2015) voting policy. From 30 minutes to 2 hours, a usual departmental voting evaluation involves the administrators' "coordinating and moderating" the meeting evaluation of all faculty members (I4A3). The meeting generally started with the administrator summarising what the department had achieved and the area of improvement. The administrators then shared a summary of the whole faculty/ department/ individual performance based on the faculty members' submitted self-evaluation forms. The administrators also moderated the meeting where their faculty members shared their

comments and voted among themselves for the titles of excellence at different levels. Then the administrator would summarise, document and forward the voting results to upper levels of management for further consideration.

Figure 13

Summary of Administrators' Voting Evaluation ToA

<p>Constraints on Administrators' Approaches</p>	<p><i>Policy constraints:</i> Voting evaluation policy (Grand University, 2015) indicates administrators' roles to organise, moderate evaluation meetings, and report evaluation results.</p> <p>and</p> <p>Voting evaluation policies support administrators in enacting their management roles.</p> <p>but</p> <p>Policy's lack of focus on discussion on individual performance quality improvement</p>	<p><i>Conflicting expectations from various stakeholders, including</i> One-size-fits-all voting evaluation criteria, quotas, and processes</p> <p>but</p> <p>Administrators' need for more accurate evaluation results & more effective evaluation forms and uses</p> <p>and</p> <p>Harmony-oriented norms (i.e., avoiding conflicts or showing respect to others)</p>
		
<p>Administrators' Approaches</p>	<p>Administrators (n=18) complied with the voting procedures as outlined by the evaluation policies</p> <p>but</p> <p>Administrators (n=7) did not use the voting evaluation to improve individual faculty members' learning and teaching.</p>	<p><i>In response to multiple stakeholders' needs,</i> Administrators (n=10) took various problem-solving strategies, e.g., adapting forms, compromising results, or improvising additional rewards</p>
		
<p>Consequences of Administrators' Approaches</p>	<p>Voting valuation policy requirements fulfilled</p> <p>but</p> <p>Limited impact on faculty members' learning and improvement</p> <p>and</p> <p>Administrators' conflicting feelings and attitudes towards voting evaluation</p>	

Note. ToA: theory of action. *n* indicates the number of participants who reported taking specific actions. The numbers of participants taking various approaches are not mutually exclusive.

Many administrators paid particular attention to supporting their coordination and moderation of the whole faculty voting evaluation meetings by monitoring their faculty members' performance throughout the school year. Overall, administrators used both formal

and informal sources of information to monitor faculty performance and compliance to their institutional goals, rules, and tasks. As moderators of faculty voting evaluation meetings, administrators needed to hold as much information about faculty performance as possible. Before the voting evaluation meeting, the administrators coordinated within and outside their faculties to collect information about faculty performance and disciplinary issues. During the voting evaluation meeting, they drew on their gathered sources to provide additional information about faculty performance so that faculty members could use it as a reference to evaluate others. One administrator commented:

We base on many different sources... Not only student evaluation of teaching is taken into consideration, but we also use inspectors' reports. Besides, we consider other factors, such as faculty members' levels of willingness or attention paid to assigned tasks... or how faculty members fulfilled their required teaching hours or how they met specific course requirements. (I4A2)

Some administrators used various sources of information to exclude those who did not meet the basic requirements of the teaching jobs. They also excluded some faculty members from the nomination list, mainly for not duly completing tasks, such as not completing professional learning courses, not completing course content, being late for class or exam invigilation, or having inadequate teaching or research hours.

Among various sources of evidence were those from inspectors, students, colleagues, and faculty members. One key source of information the administrators depended on for faculty performance was their institutional inspectors' reports. A board of inspectors were made up of administrative staff and faculty members within each institution, and their role mainly involved checking if faculty members followed disciplinary requirements such as being punctual at work or if their teaching followed the prescribed course content. However, the frequency of inspectors' checks varied across institutions. For example, inspectors from HEI 4 conducted random checks of faculty members' disciplinary conformity, while those from HEI 2 conducted regular checks. One administrator described that their inspection board regularly checked if individual faculty members were punctual in class. The inspection board usually recorded it in their notebook if any faculty member was 15 minutes late or more. Besides those from the inspectors, administrators also had reports about faculty member performance from various functional departments in their institutions. For example, the scientific research department provided the administrators with information about faculty members' engagement in the research-sharing activities, and the academic affairs department sent records of faculty members' teaching hours.

Another source of data that the administrators referred to when considering faculty members' performance during the voting evaluation was faculty members' SET scores and

students' formal and informal feedback. One administrator claimed that she learned through students' feedback that one faculty member was cancelling classes without offering replacement classes. She commented:

Without student feedback, I would not have known that one faculty member in my faculty skipped classes without conducting the make-up classes. I cannot spend my whole day just monitoring the faculty members fulfilling their class teaching hours. That is why I found student feedback quite a good channel of information. (I4A1)

Administrators also collected feedback about faculty member performance through colleagues. Peer feedback about faculty performance was communicated to administrators in different ways. Some administrators assigned leaders to report the level of disciplinary misconduct of individual faculty members. If faculty members forgot their invigilating tasks or frequently made mistakes in designing forms, they would receive reminders from the administrators. In several cases, when administrators were unsure about individual faculty members' work attitude or performance, they could also talk to some leaders for additional comments on the faculty members.

Disengaging Approach. Over one-third of the participants (n=7) reported their disengagement from using evaluation for faculty teaching and learning improvement. This meant that these participants only focused on following policy requirements to consider faculty performance as a whole, discussing generic feedback on the whole faculty or departments, rather than individual feedback or teaching and learning improvement, during evaluation meetings. The administrators tended to focus more on the faculty members' achievements and productivity rather than the quality of individual faculty members' work: "We did not usually comment on individual performance because there were not much to share. All the faculty achievements in the members' self-achievement report forms" (I4A3).

Problem-Solving Approach. One key challenge for the administrators was to address multiple stakeholders' needs during the voting evaluation implementation. While the voting evaluation policies (Grand University, 2015; Law on Emulation and Commendation 2003) demanded that administrators comply with the established criteria and procedures, many did not find these criteria and processes useful. They individually expected more effective evaluation forms, more accurate and fairer evaluation results, and more effective use of evaluation to motivate faculty members' increased productivity. Some administrators also found a need to moderate voting evaluation meetings in ways that maintained harmony across faculties and institutions. In response to the multiple stakeholders' needs during the voting evaluation, administrators (n=10) chose various strategies: adapting evaluation forms, compromising evaluation results, and improvising faculty rewards.

Adapting Voting Evaluation Forms. While many administrators referred to the Grand University's generic evaluation criteria during the meeting, a few administrators created their faculty's specific self-evaluation report forms based on the institutional voting evaluation policy guidelines. The institutional voting policy (Grand University, 2015) only lists seven criteria for evaluating faculty performance, mainly focusing on faculty members' timely completion of teaching, supervision, and research hours. These one-size-fits-all evaluation criteria fail to recognise multiple activities and services that faculty members from different member HEIs participate in during a year. Thus, a few administrators created faculty-tailored forms to specify and recognise more accurately the number of professional activities and institutional services in which faculty members participated during a year.

For example, in one faculty-tailored evaluation form provided by a participant (I4A6), there were eight categories of evaluation criteria with more than 25 groups of activities and over 70 statements that specified the faculty members' activities in a school year. The form also explicitly listed additional tasks, giving reward scores to faculty members involved in teaching and management activities outside the standard requirements or supporting student activities. The faculty-tailored evaluation form was distributed to individual faculty members to calculate their own scores by adding or deducting their scores on specific categories based on a given score for each activity. The administrator and faculty members must carefully consider each group of evaluation criteria, their maximum scores, and the score metrics to determine the scores they deserve to receive. Notably, the first four criteria were closely related to the general criteria for voting evaluation established by the Grand University (2015), such as participating in institutional activities such as the Vietnamese Communist Youth Union and Trade Union.

Indeed, many expectations were placed on faculty members to complete tasks and participate in institutional activities. The faculty-tailored forms helped the administrators recognise faculty members' participation in various activities they were required to perform every year. However, the adaption of the voting evaluation form was not commonly reported among the administrator participants. In some other faculties, administrators (e.g., I2A1, I4A4) created their own list of criteria for evaluating faculty performance, but their criteria were not publicly presented and shared across their faculties.

Compromising Voting Evaluation Results. The quota for selecting outstanding titles of excellence during the voting meeting was limited (around 15%), and the selected members needed to have more than two-thirds of the votes. This caused difficulty for administrators and faculty members to select the most deserving members among themselves. Thus, during the voting evaluation meetings, some administrators chose various ways to compromise evaluation results by negotiating evaluation criteria, withdrawing from nominated outstanding titles, or influencing the nomination list.

First, some administrators chose to negotiate the evaluation criteria and nomination lists with their faculty members. Together they shortlisted the nominated members before everyone voted to ensure they had a major agreement among themselves regarding who would be voted for. As one administrator stated, “the voting is by essence just recording the name that everyone had agreed over” (I1A1). This action was to limit the selected number before voting so that whoever won the votes had a higher chance of being selected by the higher level of management:

I usually try to negotiate [with faculty members] so that the number of selected candidates and the nomination list are relatively equivalent; otherwise, the number of votes for the chosen candidate would be scattered. For example, once we were able to select only two outstanding faculty members, but three people were nominated with 70%, 80% and 30% each – so the first among the list did not get selected [for having the percentage of votes lower than two thirds of the total number of votes] ... so I try to avoid that circumstance happening. (I2A1)

Second, some administrators chose to withdraw themselves from the nominated list. Except for a few exceptions, many administrators were still categorised among the faculty members list and included in the voting evaluation. As they were also moderators of the evaluation meeting, some administrators withdrew their names from the nomination list to allow other members to be voted for instead. For example, one administrator decided to do so even though he thought that he deserved the title of excellence:

In our department of about six or seven people, we would first share our self-report of our own achievements. Each of us would choose a title that we think is deserving of ourselves. I would say, for example, that I had received the outstanding title of excellence many times... and the quota was limited. I would give it up this time so that other faculty members could be selected. If people did not have any comments, that was dealt with! (I1A1)

Similarly, compromising actions were reported when administrators directed their faculty members to prioritise voting for some faculty members in several cases. For example, Administrator I2A2 commented that the emulation and reward evaluation took much time and was quite a complicated procedure to get it done correctly; therefore, he and his faculty members just simplified the process. They set up an implicit faculty-level order of priority for voting that had become a “tradition” for quite a while (I2A2). The first group of faculty members that were on the priority list was the to-be-retired senior lecturers due to the Vietnamese tradition *kính lão đắc thọ* (respecting the old and you will achieve longevity), which means respecting senior members. The administrator found his faculty members accepted this

tradition because they wanted to show their gratitude to the senior lecturers who had contributed to the faculty for so many years:

When senior lecturers are about to retire, our faculty members want to celebrate their contribution to training our students. These [senior] faculty members also had mentored young staff [including me] since we graduated and were trained to become faculty members here. They helped the faculty train young staff to [become] [Communist] Party members and mentored those who shared the same research disciplines with them. (I2A2)

Improvising Additional Rewards at Faculty Level. As the institutional reward for outstanding faculty members was limited with only around 15% of the total faculty members being selected for outstanding titles through the voting processes (Grand University, 2015), some administrators improvised faculty-level excellence titles and rewards to recognise the deserving faculty members who missed the voting evaluation rewards. For example, Administrator I2A1 wanted to use evaluation as an effective administrative mechanism to motivate his faculty members to attain professional advancement, such as teaching more effectively and writing more high-quality articles; however, he did not find any institutional mechanism for doing so. Thus, the administrator improvised a faculty-level reward scheme for dedicated faculty members. He initiated a faculty research fund to reward faculty members for their publications. He asked an assistant to support him monitor his faculty members' research and article-writing process, rewarding members with 200,000 VND (nearly 9 USD) and 500,000 VND (over 20 USD) for an international publication. Similarly, some administrators decided to create their own Emulation Fighter title at the faculty level in addition to those formally awarded by the Grand University (I4A3, I4A4, I4A5). In the latest employee meeting at the faculty level, Administrator I4A4 publicly commended and awarded some faculty members a very beautiful certificate of merit and a material reward of 500,000 VND. She thought that it would be best if the faculty could appreciate the contribution of these deserving-but-not-recognised faculty members.

Constraints on Administrators' Approaches

Several sets of constraints explain the administrators' complying, problem-solving and disengaging approaches to voting evaluation. Besides the institutional voting policies (Grand University, 2015) being overarching constraints for the administrators' complying and disengaging approaches, multiple stakeholders' expectations influenced how the administrators solved the problems during the voting evaluation implementation (see Figure 13).

Policy Constraints. One key constraint that explains the administrators' complying and disengaging approach is the voting policy document policies (Grand University, 2015). The policy indicates that administrators must organise evaluation meetings and report evaluation results (article 31). Aligning with the policy requirements, the administrators had faculty members complete their evaluation reports and organised evaluation meetings where faculty members commented and voted on titles of excellence for each other. The voting policy document (Grand University, 2015) also indicates that administrators comment on the achievement reports of individuals and the whole faculty. Thus, the administrators collected information about faculty performance and disciplinary compliance from various sources, as described earlier. Such disciplinary compliance is important because it reflects faculty members' quality of political ideology or people's conformity to the communist state's political ideology. Specifically, it is clearly stated in the Grand University's voting evaluation guidelines (2015) that one out of four common criteria for the basic titles of Advance Labourer is related to their compliance with organisational regulations, state laws and the Vietnamese Communist Party's policies (article 11, pp. 5–6).

Another constraint explaining the administrators' compliance is that the voting evaluation policy supports their enactment of managerial roles. There are general criteria for considering the different titles of excellence, one of which is that all staff, including faculty members, are required to adhere to the rules, regulations, and laws established by their departments, faculties, institutions and those by the government and the communist-ruling party (article 11.1.b). As coordinators and moderators of the evaluation meetings, some administrators found that the voting policies supported their managerial roles for several reasons. First, some administrators found that the use of informal and formal data sources, or the "cross-checking mechanism," together with the quantification of faculty members' performance, evaluated members relatively accurately (I2A4). As part of the voting process, the administrators' coordination of faculty performance from various data sources helped them monitor faculty members' compliance with disciplinary requirements, such as being punctual for classes and meetings. Second, the voting evaluation was helpful for the administrators in their organisational management, especially in identifying and solve problems with faculty members' performance. Voting evaluation results were especially "significant" for contract faculty members. If these members did not complete all tasks for 2 consecutive years, they could not extend their contract (I4A3). Third, the voting evaluation was also beneficial for outstanding individuals who received certificates of merit from high levels of authority. When faculty members were rewarded with outstanding titles, it was an "honour" that they could carry with them wherever they worked (I1A1).

On the other hand, the policy constraint that explains the administrators' disengagement from faculty teaching and learning improvement is the lack of focus on quality

improvement of the voting evaluation policies. At the national level, the voting evaluation aims at mobilising human resources towards organisational and institutional performance and productivity (Law on Emulation and Commendation 2003). As the voting evaluation is a policy directed by the Vietnamese Communist Party leader, the leaders at Grand University and its member HEIs must follow the national voting policy. Thus, the Grand University (2015) emphasised its voting evaluation process on monitoring and reinforcing faculty members' compliance with institutional performance goals rather than supporting individual faculty members' professional learning and teaching improvement. During evaluation meetings, the administrators focused on listing and counting performance rather than on discussion quality and typically did not give much feedback on teaching, learning, and research quality. Compliance with rules on faculty work was reinforced by the administrators incorporating the evaluation of faculty members' work with surveillance of faculty members' work by various functional departments besides the academic departments. The way policy requirements were written and the coordination between the administrators with staff from other functional departments such as the inspectors' board or the personnel offices made the administrators' evaluation roles more administrative than professionally driven.

Conflicting Expectations From Various Stakeholders. Another key constraint that explained the administrators' problem-solving approach was their need to address multiple expectations from the voting evaluation stakeholders. First, the voting evaluation policymakers at the Grand University instructed the administrators to follow the same evaluation criteria, whereas the administrators and faculty members needed more specific evaluation forms to reflect faculty members' participation in professional activities and institutional services more accurately. The voting evaluation policy document, which prescribed one-size-fits-all evaluation criteria, did not cover multiple aspects of worklists and the workload of faculty members across the institutions and faculties. The Grand University only provided generic evaluation criteria, whereas institutions and faculties had a range of fixed and unpredictable tasks annually assigned to faculty members. Generally, the administrators faced addressing faculty members' expectations that their participation and contribution to collective events would be fully recognised. Meanwhile, faculty members' tasks were enormous, ranging from faculty members' completion of teaching and research hours, participation in professional activities, and participation in various campaigns, which were significant in numbers and workload.

Some administrators improvised additional rewards outside the voting evaluation ones because they wanted evaluation to be more motivational to faculty members. The Grand University has imposed voting evaluation processes, and the rewarding mechanism made it difficult for the administrators to motivate their faculty members to boost their performance and contribution to teaching and learning improvement. They perceived that the current voting

evaluation system did not give them much authority over recruiting, incentivising, or sanctioning faculty members based on their quality of work. Nevertheless, some administrators expected that the voting evaluation could be used to leverage the quality of faculty members who had a direct influence on student learning and employability. They believed the recognition and material rewards through the voting evaluation would motivate faculty members to outperform as “*một đồng tiền thưởng bằng mấy lần tiền lương* (one-coin reward is equal to several times the normal wage) (I4A3). Some administrators found that many faculty members deserved to be better recognised and appreciated for contributing to the institution due to heavy and difficult managerial and institutional demands. However, as the quota of title recognition through voting was limited to a certain percentage (around 15%) and directed by the top-down management, so the administrators’ desire for promoting deserving members through the voting process was not possible:

Unfortunately, for example, there were many new and difficult jobs for the whole faculty in some years. We want to propose outstanding titles for many faculty members, but these titles are always limited from the top-down, with no more than a certain percentage of staff selected. It is difficult to praise all the deserving faculty members for their excellent contributions during an entire school year. (I4A3)

Due to the administrators’ lack of authority over the evaluation process, they could mainly only adapt the evaluation forms and criteria or create faculty-led rewards for their faculty members. As one administrator put it:

I want to build a working environment that motivates faculty members, but it’s hard to do so in the public sector. I oversee motivating faculty members to work, but I don’t have the authority to recruit, give financial incentives, or sanction faculty members. Those bring power to administrators like me to motivate the staff. It is hard because I do not have any tools to require my staff to work. (I2A1).

In this case, the administrator’s value of using evaluation for motivation conflicted with his perceived lack of authority over the process. He linked this lack of power to his institutional lack of financial and personnel autonomy. He predicted that if the institutions and staff had a higher level of autonomy, the faculty evaluation could be directed more towards quality than the current focus on inspection and monitoring faculty discipline. As the current evaluation system did not recognise the quality of faculty teaching, several administrators found ways such as rewarding faculty members’ good supervision of students, as described in the previous section.

The voting evaluation quota and processes were found to be contradictory to the Vietnamese participants’ harmony-oriented norms, which caused difficulties for the

administrators to moderate fair evaluation meetings while maintaining harmony among faculty members. One administrator found that her perceived need to give opportunities to others was rooted in the Vietnamese culture of people-pleasing and being relationship-oriented (trọng tình cảm nể) and egalitarianism (cái gì cũng quân bình). The people-pleasing and relationship-oriented norm refer to the administrator's perceived expectation to please others to maintain harmonious relationships among faculty members. Egalitarianism means that one should not always be too outstanding and receive all the benefits but should be willing to withdraw to allow others to be recognised. The administrator illustrated with her own example:

It's like if you get it [the outstanding title and material reward] this year, the following year [title and reward] should be given to others. It is not competitive like when I work in foreign companies. If I am rewarded there [in foreign companies], I do not need to be like that [stepping back to give opportunities to others]. (I4A1)

The limited quota drove faculty members (including administrators in many cases) to compete against each other to be selected among the few (around 15%) outstanding members. Also, the voting evaluation process requires the administrators and faculty members to share their own achievements, comment on others, and vote for the most outstanding members among themselves. Due to the collective nature of faculty evaluation meetings, administrators and faculty members did not have enough time and available information to commend and give thorough feedback on each faculty member's performance. The competitiveness and the whole faculty evaluation process also placed all the administrators and faculty members in an awkward position. Faculty members and administrators had to play dual roles of appraisers and appraisees in meetings where they shared a conflict of interest. If a faculty member voted for himself or herself, he or she would take the opportunity away from other members.

This conflict of interest was odd with the Vietnamese participants' harmony-oriented norm, which involved showing respect and avoiding conflict. Due to the Vietnamese culture of face-saving, administrators found it hard to give negative or constructive feedback to faculty members for fear of them losing face (I2A1). Therefore, during the evaluation meeting, administrators only stated generic feedback on the whole faculty or departments. The harmony-oriented norm is also linked to the perceived need to prioritise others over oneself. One participant found that relationship orientation was a norm as it appeared to be well accepted by other faculty members. She shared her experience when she withdrew from the nomination list because she had received the outstanding title the year before. Her faculty members all accepted it even though she outperformed the other faculty members:

I found that I got the title of Emulation Fighter at the grassroots level in the previous year, so I decided to step back to allow others [to be selected]. When I proposed that

idea, everyone accepted and found it quite normal. That year I got one international and five domestic publications, whereas the other member [had no publication, and she] only had one presentation at a national conference. Foreigners would not know how the system worked like this, but we [Vietnamese people] tend to be people-pleasing [cả nể].

When asked if the limited material reward was one reason for allowing the outstanding title for other people, the administrator found it not necessarily the case. In her faculty, the voting titles were tied to additional income, so the total amount of material reward was quite significant. Due to the significance of the additional income, the administrator felt the need to share it with others, no matter who performed better:

When I was a new faculty member, I did not care [about the material reward]. It's because we only got 200,000 VND [nearly 10 USD] of material reward for the Advanced Labourer title and 300,000 VND [over 10 USD] for the Emulation Fighter at the Grassroots level. Thus, it did not matter which titles we received at that time... But for my institution, in the example I showed you, the person who achieved the outstanding title would get 16 million VND [about 650 USD]. In contrast, those with less outstanding titles would get only 8 million VND [about 325 USD]. It was quite a "big plate," which makes us think that we should not always "eat it all, " so we should give others the opportunity to others... We would not mind if the material reward were like that, or only 500, 000 or 700, 000 VND [20 to 30 USD].

Consequences of Administrators' Approaches

Several consequences resulted from the administrators' complying, disengaging and problem-solving approaches to the voting evaluation.

Voting Evaluation Policy Requirements Fulfilled. The first consequence is that administrators' roles of monitoring faculty performance were fulfilled, and in addition, they believed that it would help increase members' sense of responsibility and rule compliance. In other words, the administrators' complying approaches to the voting evaluation helped reinforce faculty members' adherence to their institutional rules, goals and values. The inspection boards' reports were deemed important to individual faculty members because such reports were associated with individual faculty members' extra income and opportunities for being considered for the titles of emulation during the voting evaluation meeting. For example, at HEI 2, faculty members' additional income was calculated based partly on the inspectors' reports. If a faculty member was recorded with disciplinary problems such as being late to class, their additional monthly income would then be deducted. Those on the inspectors' blacklist would also run the risk of not being voted for outstanding titles in the voting evaluation:

Those who get recorded will be deducted the extra income of the month. That means, for example, the additional income is 4 million 1 month, then I will be removed 2 million, only 2 million, that example is deducted, the first is subtracted from finance... Also, it's a form of quality control, right? Going to class without punctuality is related to quality and attitude. When considering the annual emulation, you know that at the end of the year, we will consider emulation and commendation, emulation soldiers, advanced labour, etc ... For all kinds of titles, anyone who has such a late arrival is not considered. (I2A2)

Limited Impact on Faculty Members' Learning and Improvement. A few administrators found that comments from other colleagues during the whole faculty meeting for voting evaluation helped some junior or inexperienced faculty members; other than that, many administrators found the peer comments were not meaningful for faculty members' learning and improvement. Unfortunately, many administrators had to prioritise nonprofessional criteria over professional ones, causing faculty evaluation to become somewhat political rather than professional-oriented. Although faculty members' professional competence is crucial, one administrator emphasised that other nonprofessional qualities, such as the *quan hệ với đồng nghiệp* (ability to maintain good relationships with other colleagues) or *phẩm chất chính trị* (having the proper awareness of the Vietnamese Communist Party's political ideology), were more important:

We normally find that the evaluation criteria can be divided into two parallel facets: one facet is professional knowledge of the recruited faculty members, while the other is even more important facet being faculty morality or their quality of political ideology because this is how thing goes here at our faculty. (I2A4)

Although several administrators did not fully agree with the emphasis on nonprofessional evaluation criteria, they needed to comply because of the heavily weighted policy constraints and the need to maintain organisational stability. Consequently, the voting evaluation would "not make sense to individuals" (I2A1) as it did not create much motivation for advancing professional learning for a broad mass of faculty members.

Administrators' Conflicting Feelings and Attitudes Towards Voting Evaluation. Interview data reveal administrators' varied feelings and attitudes towards the voting evaluation at their institutions. First, administrators tended to feel more satisfied with the impact of voting evaluation when they could adapt the evaluation to recognise faculty members' performance and contribution better. For example, Administrator I4A1 found that by looking at the summarised tables that include a list of individual scores generated by faculty-tailored forms, she would easily spot the level of contribution among individuals. With the detailed forms of self-evaluation provided to faculty members for self-reports, the administrator

felt more confident in using the total scores to grade faculty members into different levels because each activity was considered against specific criteria for calculating the scores (I4A1). Administrators were also satisfied when they could improvise some alternatives to reward some deserving-but-not-recognised faculty members due to the limited quotas from the upper management. For example, Administrators I4A4 and I4A5 also found faculty-level recognition of members' contribution through public honouring, praising or commendation as a pleasant "spiritual encouragement" or "good source of motivation."

On the one hand, some administrators were dissatisfied with the current approach to solving problems during their evaluation meetings. For example, Administrator I4A4 found that compromising evaluation results based not on faculty performance but on the relationship between faculty members was unfair and "very crazy" when sometimes the hard-working and dedicated members were not selected. This selection was unfair because these dedicated faculty members had not registered for the outstanding title in advance. Also, the university's quota set a limit that only one-third of the votes could be given to members with managerial roles such as head or vice-head of the department, so many of the administrators were not supposed to be selected no matter how much they had contributed to the faculty task assignment and development. As Administrator I4A4 stated,



So the ratio is very limited, very crazy when in a department of 23 people, only two will be emulated with outstanding titles and only one third of which are group leaders... Sometimes it is very bad that even the head and the vice-head work very hard, but they must give it up for other less active members. Because of that quota, many faculty members who meet all the criteria [for outstanding titles] are not considered. (I4A4)

A couple of administrators expressed their disapproval that, during voting evaluating meetings, some faculty members tried to be loud to dominate others' choice of voting. For example, one faculty member intentionally and loudly reminded others that she had the Emulation Fighter title at the grassroots level in the previous year and that she would be promoted to the title at the university if she was voted 3 years in a row: "it was as if the faculty member was trying to overwhelm others" (I4A4).

Faculty Members' Voting Evaluation ToA

The following sections describe the faculty members' voting evaluation ToA comprising their approaches to the evaluation, the constraints and consequences of these approaches (see Figure 14). Overall, all faculty members complied with the evaluation roles and procedures specified in the Grand University (2015) policy, but many disengaged from using the evaluation for their own learning and improvement. During the voting evaluation meetings, the faculty members used various strategies to evaluate selves and others.

Figure 14*Summary of Faculty Members' Voting Evaluation ToA*

Constraints on Faculty Members' Approaches	<p>Voting policy requirements of faculty members' roles (i.e., completing reports, attending meetings) (Grand University, 2015)</p> <p>and</p> <p>Necessity of compliance</p> <p>but</p> <p>Perceived ineffective voting evaluation criteria and processes (e.g., irrelevant criteria, hasty and bureaucratic process)</p> <p>and</p> <p>Low material and emotional incentives</p>	<p><i>Conflicts between the policy and the norm:</i></p> <p>Collective leadership principle (Grand University, 2015) that espouses everyone is a leader</p> <p>but</p> <p>Collectivist norms (i.e., harmony-oriented and relationship-based decision making) that prioritise groups' interest</p>
		
Faculty Members' Approaches	<p>All faculty members (n=20) complied with the roles and procedures following the voting evaluation policy requirements</p> <p>but</p> <p>Faculty members (n=9) disengaged from the voting evaluation (i.e., participating perfunctorily in the evaluation, or not using voting evaluation for learning and improvement)</p>	<p>Faculty members chose various problem-solving strategies to evaluate others during voting evaluation, including:</p> <ul style="list-style-type: none"> • Choosing safe and modest self-evaluation (n=7) • Following collectivist evaluation of others (e.g., avoiding direct and harsh comments) (n=8) • Creating own evaluation criteria (n=2)
		
Consequences of Faculty Members' Approaches	<p>Voting evaluation requirements fulfilled</p> <p>but</p> <p>Limited impact on faculty members' learning and improvement</p> <p>and</p> <p>Faculty members' conflicting thoughts and emotions about the voting evaluation</p>	

Note. ToA: theory of action. *n* indicates the number of participants who reported taking specific actions. The numbers of participants taking various approaches are not mutually exclusive.

Faculty Members' Approaches to Voting Evaluation

As Figure 14 shows, the faculty members' key approaches to voting evaluation included complying, disengaging and solving problems of evaluating themselves and others.

The following sections describe each of the approaches to the voting evaluation taken by the faculty members.

Complying Approach. Faculty members generally complied with the institutional voting policy document (Grand University, 2015), which specifies their roles during the evaluation procedures mainly by

- Writing a report of achievements
- Participating in the whole faculty evaluation meetings directed by administrators

All faculty members (n=20) complied with the voting evaluation policy requirements. Specifically, they completed their self-achievement reports, attended evaluation meetings, and followed the administrators' directions.

Complete Self-Evaluation Reports. All faculty members completed their self-evaluation reports before or during the evaluation meetings. The specificity in the evaluation forms used across HEIs' faculties varied. The evaluation forms in many faculties are generic within a couple of pages, which generally require faculty members to list their hours of research, teaching, the activities they had participated in, and their previous titles of excellence. However, a few faculties provided their members with up to eight-paged evaluation forms with very detailed categories of faculty members' annual activities.

In the case of generic self-evaluation report forms, faculty members listed and then rated their performance based on their framework, mainly listing the numbers of hours they spent on teaching and doing research annually: "I list how many lessons I have taught in that school year, how many students I instruct, how many exams I have and how many papers I write" (I7T1). However, as they did not have specific evaluation forms, they tended to rate themselves based on roughly subjective estimation of their quality. One of the faculty members commented:

If there were a specific list of points given to each of our performances, for example, different scores were given to publications at different levels; I would quantify my quality of work. However, as usual, I based on my own framework. I usually chose to complete my tasks well. For example, I guarantee the time I go to class, I am well prepared for my teaching, and I am responsible for what I do regarding my assigned tasks like designing course materials or marking exams. So, as I cannot quantify my performance, I evaluate based on my subjective opinion. (I1T1)

In case faculty members were provided with detailed forms for self-evaluation, they calculated their total scores before or during the meeting.

Attending the Whole Faculty Evaluation Meeting. After completing self-evaluation reports, the faculty members shared their self-evaluation before the meeting via emails or

during the meeting by a brief oral presentation. A whole faculty evaluation meeting generally involved faculty members (administrators included in some cases) sharing, commenting, and voting for titles of excellence among each other:

Our usual meeting procedures at the departmental level started with each person reading their review, then we commented on other members' strengths and weaknesses in the past years. Then we vote for titles such as outstanding achievers, good achievers, or so. When the meeting is over, our head of department signs the minutes of the department meetings. (I4T1)

In a few cases, where tailored evaluation forms were provided, faculty members evaluated their self-performance based on the given work lists and rating scale, for example, a Likert scale of 1 to 3 for "average to the very good completion of work" (I4T4). After the whole faculty finished calculating their own scores, the administrators wrote all the members' self-evaluation scores on the board. Based on the written scores, the entire faculty entered the stage where they voted titles of excellence for each other. The meeting ended after everyone submitted their complete self-evaluation form to the administrator.

During the whole faculty evaluation meeting, many faculty members followed their administrators' direction, especially regarding the exclusion and inclusion criteria for evaluating selves and others. For example, one rule that the faculty members followed, despite their potential disagreements, was the exclusion rules during the faculty evaluation meeting. This rule means faculty members exclude all those who did not meet specific requirements, no matter how these members performed at work. For example, faculty members would not be considered for outstanding emulation titles if the inspector had caught them during the semester for being late or missing classes without proper explanation. Also, following administrators' directions, many faculty members focused on celebrating the whole faculty and individual achievements rather than on identifying areas of improvement.

Disengaging Approach. The disengaging approach refers to the faculty members' perfunctory engagement in the voting process or their disregard for evaluating learning and improvement. Nine faculty members shared the disengaging approach to voting evaluation in two main ways.

First, some faculty members reported engaging quite perfunctorily in the voting evaluation, which means that they just did everything quickly to get it done without much care, interest, or attention. These faculty members did not find much influence of the evaluation on their salary, promotion, or professional learning. Thus, they only tried to complete their assigned tasks and did not care much about the titles of excellence: "I do not care too much, no matter what the evaluation results are, it does not matter. I have been trying to teach enough hours, do scientific research enough hours, and it is done!" (I3T1). These participants

only tried to reach the minimum requirements of the faculty members' expected tasks, and they usually evaluated themselves as having completed all tasks during the voting evaluation.

Second, many faculty members disengaged from using the voting evaluation as an opportunity for learning and improvement. These participants did not actively contribute to the self-reflecting, sharing or commenting on each other's performance during the voting evaluation meetings because they did not find it the true meeting purpose. Some participants did not see any importance in evaluation meetings where no one showed critical comments no matter what they thought: "critical feedback is not the faculty norm, and we just get it [evaluation paperwork] done!" (I4T2). Faculty Member I3T1 commented that he raised his voice, but nothing changed. The participant chose to ignore the voting evaluation because it focused more on maximising institutional commercialisation of teaching and learning rather than improving quality.

Problem-Solving Approach. During the voting evaluation, the faculty members met with a key challenge of giving a fair evaluation of themselves and others, which contradicted the relationship-based or harmony-oriented norm. These conflicting requirements were especially difficult as the faculty members play dual roles of evaluators and evaluatees during the faculty evaluation meetings, directed by a principle of collective leadership. Due to the conflicts between the evaluation principle, requirements and norms, faculty members took various strategies to evaluate themselves and others during the evaluation meeting.

Choosing Safe and Modest Evaluation of Selves. Seven faculty members mentioned their safe and modest evaluation of themselves as strategic actions to their expected norm of showing modesty. First, they typically chose a safe level of achievements for their self-reports. Usually, the emulation titles were divided into different ranks, with about 15% quotas for the highest grade and unlimited quotas for the lowest grade. For example, 2 out of 21 faculty members would be selected for the outstanding "emulation fighter" title, whereas all 21 faculty members could be selected for the lowest title of "advanced labourer." Thus, some faculty members chose the modest achievement titles to play it safe. One faculty member commented that she never rated herself with an "excellence" title because she found the word "excellent" "too much," probably like overrating oneself (I1T4). The overrating of self was contradictory to the Vietnamese expected norm of being modest. Another faculty member mentioned that they tended to be humble about their own achievements when orally presenting their self-achievements in front of the whole faculty. For example, talking about the last voting incident, one participant admitted selecting only "modest" task completion, not because he underrated himself but because he needed to show modesty:

Researcher: How did you rate yourself in the latest evaluation meeting?

Participant: Just modest completion.

Researcher: Why do you rate yourself as “just modest” task completion?

Participant: Well, I do not even know what’s “good.”

Researcher: Don’t you have evaluation criteria?

Participant: I complete all the assigned work, but whether the job is good or not is not judged by me but by others. Another reason for my modesty is that everyone is humble.

Researcher: Why humble?

Participant: A popular saying goes that being humble is good. It creates a harmonious atmosphere for everyone. In fact, as I said, part of it is difficult to judge whether what you do is good or not. I just know that I do it fully, complete the work on time, and the quality of the work is to be judged by others. (I7T1)

Another faculty member felt it was morally required to outperform and contribute significantly and profoundly to the faculty development to rate oneself with an “excellence” title. Thus, he only rated himself with an average title because he had not contributed enough to his expectations:

I rate myself as good completion of a task. Of course, I try to do well or perfectly for normal everyday things like teaching. However, in terms of something more profound, my contribution is not much. Many others are the same as well. I consider my faculty as my family, so the responsibility is not only to teach and then receive a salary, but the target is to create the strength of the subject, not only within my university but also spread to a broader scope. Of course, it is necessary to exchange knowledge more with foreign colleagues in this information era. (I2T3)

Being modest was a sound tactic because the participant did not know how other colleagues would then comment on their performance. After the participant read his self-evaluation and ratings, others would state their opinions on his performance, such as he deserved “to be excellent because of having done this and that,” his “quality of work is excellent, but it cannot be completed or completed well” (I7T1). From the faculty members’ accounts, they felt more comfortable being set for the excellent titles by others rather than self-nominating.

Following Collectivist Evaluation of Others. Collectivism refers to the Vietnamese participants’ value of prioritising group over individual interests, which is manifested through various norms such as harmony-oriented and relationship-based decision making. Eight faculty members mentioned several collectivist strategies to evaluate others during the voting faculty evaluation meeting. One strategy was that faculty members avoided underrating or giving direct and harsh comments to others during the evaluation meetings. They provided

feedback to other teaching staff in a very soft manner, and they tried to be as tactful as possible to avoid potential unpleasantness:

We may provide feedback to other teaching staff, but very soft, and we do not relate that feedback to the teaching quality. Some teaching staff might not have very good SET results in terms of teaching content or methodology, but we do not use it to lower their evaluation grade to B or C. The most negative way of the evaluation could be to comment sincerely or in a tactful way but not use it to consider their evaluation grade. In Vietnam, the voting and grading of different levels of evaluation are mostly zigzag – evaluation practices tend to be governed by “affection” more than “reason.” (I2T2)

Another strategy was that faculty members followed collective opinions of individual faculty members when considering ratings for them. As described earlier, evaluating others in a big meeting was challenging, especially when there were no clearly defined rating criteria or when faculty members did not have much insight into other members’ work. In such cases, faculty members had to refer to both formal and informal sources of information about other members’ performance. Thus, some faculty members referred to the opinion of other members, usually the senior or outspoken members who publicly nominated someone they felt was deserving. An example of the comment was, “this year, Member A and Member B can do many things that others can’t, so we must give credit to them for their excellence” (I1T7). Many others referred to their administrators who moderated the evaluation meeting. As one faculty member recalled in her latest evaluation meeting, her administrator took the initiative to suggest the specific selection criteria, and then the rest would narrow down the list of the deserving members based on her suggestion:

Right in the meeting during our lunchtime today, Ms. Administrator gave us the first criterion as “whoever has an article, regardless of in a domestic or international ranking journal, will be shortlisted.” Then we narrowed down a list from 13 to six members. Next, Ms. Administrator called out another criterion as “who participated in our in-house projects as a leader will be shortlisted,” then we narrowed down the list to 2 members... that is basically our two main selection criteria! (I1T1)

Usually, these rounds of comments and nominations ended up with everyone agreeing, and then the results were recorded in the minutes. As one shared, he was familiar and comfortable with the process because it was a hidden principle in his faculty that members took a turn to receive outstanding results. Thus, no one in his faculty would think of questioning an unwritten norm like that. One faculty member expressed that this kind of seemingly majority-based voting was formalistic, and it prevented alternative opinions on selecting the most deserving ones:

Sometimes I feel that the voting evaluation in my faculty is just a formal ritual. For example, if someone says, "I think these three people deserve it," others will be okay. There is almost no time and opportunity for us to say that "I find others better." Most people in my department judge others based on the majority's opinion. (I1T5)

Creating Own Evaluation Criteria. Two of the faculty members, who did not conform to other people's opinions, invented their own evaluation to evaluate other colleagues' performance. For example, one faculty member did not base her evaluation on the listed criteria but on her ongoing observation of other members' work based on their work's quantity, quality, and attitude. Another participant found the current evaluation focused on merely listing rather than sharing and considering the quality of faculty work. She regarded evaluation as recognising one's dedication and enthusiasm to quality improvement rather than counting the number of teaching hours or the research outcomes. Thus, when evaluating others, she did not look at the summary of faculty members' task completion, but she based on her own observation of how others worked, interacted, and contributed during the whole year. Although these evaluation criteria seemed subjective, the participant felt confident because her faculty was small and members were familiar with each other:

I do not really need to look at that summary to see how other faculty members performed. This is because our department is very small, so mostly all of us know very well about each other's work after a whole year we worked and interacted with each other. It means that without looking at the summary, we can make their judgement table such as "this member worked very enthusiastically, and that member did not" I will base partly on the summary of the task completion and partly on the degree of collaboration among members during the whole year to judge how a faculty member is. I think there is a certain degree of subjectivity here, but because we know each other very well, I think our evaluation of others is accurate. (I1T7)

The participant found that she had a clear sense of who was more enthusiastic and devoted than others: "for the same task, some might invest more time while others might spend less time on it. I still appreciate those who spent more time on one task than others" (I1T7). The participant found her own criteria, for example, her preference for enthusiasm over research output, more appropriate because the former was related to attention to quality whereas the latter was just a required task:

Well, when it came to priority, I would choose the faculty's enthusiasm in task completion. Doing research, for example, is one criterion, but research output does not actually show the degree of faculty's enthusiasm. This is because doing research is only a required task that faculty must fulfil, concerning the fixed research hour

requirement – like an obligation one must fulfil. Still, one may not really be enthusiastic about doing research, for example. However, I think essential qualities in teaching should be enthusiasm and willingness to support students. (I1T7)

Constraints on Faculty Members' Approaches

The voting policy indications explained the faculty members' complying approaches, whereas the participants perceived that ineffective voting evaluation criteria, processes and administrator leadership influenced their disengagement approach. The participants' problem-solving approach was explained by the conflict between the collective leadership principle and the collectivist norms during the voting evaluation (see Figure 14).

Voting Policy Indications and Necessity of Compliance. One key constraint for the faculty members' compliance was the policy indication of roles. As described early in this chapter, the voting policy (Grand University, 2015) mainly requires faculty members to undertake three tasks during the evaluation: write a self-achievement report, select a self-evaluation grade, and participate in the evaluation meeting directed by the administrators (article 31). Therefore, as described earlier, all faculty members (20 out of 20) followed the voting procedures directed by their HEI leaders and administrators. The voting policy role indication was supported by the faculty members' perception of the benefits of the voting evaluation and the perceived consequences of noncompliance.

Many faculty members complied with the voting evaluation system due to the perceived benefits of the voting evaluation. As one participant commented, the voting evaluation was "meaningful in various senses" (I3T4). First, the evaluation was important to individual members and the whole organisation. The evaluation served many important purposes, such as considering rewards and maintaining staff discipline. The voting evaluation results influenced individual financial bonuses and the institutional system of personnel decision making. As one participant put it:

There is definitely an influence not only on individuals but on all employees because it is related to salary, bonus, ranking and promotion, job arrangement and things like that. In terms of regulations, those are serious ratings. For example, if you are considered not to complete your tasks, you will not receive a monthly salary bonus. Or, if you are only an advanced worker, you will be rewarded, for example, a coefficient bonus of 1.0. If you are an emulation fighter, go to a coefficient bonus of 1.2 or 1.3. That is, you will have a reward corresponding to all of those. (I7T1)

Second, through evaluation meetings, faculty members could officially receive feedback on their quality of work from their colleagues. The voting evaluation involved self-evaluation and peer evaluation of performance which were feedback channels for some faculty

members to know what they had done so they could correct or improve the quality of their work. Notably, the perceived amount and quality of colleagues' feedback varied, depending on the size of the faculty, the administrators' leadership, and the mutual trust among faculty members. One junior faculty member who had 3 years of working experience shared that he appreciated the valuable comments from his senior colleagues

After I shared my self-report, I received very objective and democratic comments, very impactful. When people comment on my shortcomings, I see a lot of impacts. Every time people give positive comments or praise, I do not say anything. But for what I have not achieved or suggestions - those are valuable contributions, I appreciate them, and I feel that I need to change. (I2T4)

For some junior, less experienced faculty members in supportive faculty environments, feedback from administrators and colleagues helped them realise what they had and had not achieved. Similarly, faculty members benefited from the voting evaluation meetings where everyone actively, respectfully and constructively shared and learnt from each other. Also, for a few faculty members, the general process of voting was "correct and clear" because the results were based "entirely on secret votes," which were entirely up to the faculty choice (I3T2).

Another reason for the faculty members' complying approach was the perceived consequences of noncompliance. Their biggest expectation regarding the voting evaluation was to get the most basic emulation title and not to be on the blacklist of those who violated discipline. In these cases, the necessity of the voting evaluation was linked to their fear of being blacklisted. They did not care much about the emulation titles or the financial bonus connected to the voting evaluation. However, their fear of being on the blacklist was connected to their institution's collective ways of publicly evaluating and announcing individuals' evaluation results. One faculty member described that her institution emailed all members a long list of those who did not fulfil their required tasks every year. For example, in the most recent year, a list of approximately 17 members out of over 100 faculty members was publicly announced as having not completed their tasks. The person's name, department, and reasons why they were not eligible for being considered for any emulation titles were listed, for example, for not accumulating enough research points. Since this public announcement of negative evaluation results was a source of shame, the faculty member did not want her name and department to be on such a blacklist:

Actually, it's not because I'm afraid of being cut from salary and bonus, but I'm afraid to make mistakes, and I simply don't want my name to be shamed on the year-end board. For example, I find that very disgraceful when people say they don't pass the

competition in any department. Why does that have to put everyone's name up through the email of an entire school like that? (I4T5)

Perceived Ineffective Voting Evaluation Criteria and Processes. Despite the necessity of complying with the voting policies, many faculty members disengaged from actively participating and using the evaluation for learning and improvement due to the perceived ineffectiveness of evaluation criteria and processes.

First, the faculty members perceived some problems with the evaluation criteria, which focused more on quantity and productivity than the quality and process of faculty work. The voting evaluation meetings mostly centred around listing, ranking, and voting the faculty members with the highest productivity related to the number of teaching hours, research outcomes, and participation in institutionally assigned activities. One faculty member commented:

The evaluation only served two purposes: first, we tried not to be caught over disciplinary issues by the institution, and second, we received enough salary and bonus, that is it! As a faculty member, I just need to complete enough teaching and research hours - to fulfil my responsibility. That is good enough. (I3T3)

The evaluation criteria in many faculties were also rigid, so some members found they did not match their career aspirations. Due to the non-track evaluation system in the HEIs, all full-time faculty members in a faculty were mainly evaluated under the same criteria, same expectations, and same expected titles and rewards. Some faculty members aspired to dedicate their time to their teaching job, but they could not simply choose it because the evaluation criteria demanded that they fulfil enough research hours within a year. Also, some faculty members found a mismatch between their expected evaluation focus and the actual focus of the evaluation meeting. The voting policy indicates that faculty members report self-achievement, comment and vote for themselves' and others' titles of excellence during evaluation meetings (Grand University, 2015). However, instead of giving comments or feedback on faculty members' professional learning and teaching, many faculty evaluation meetings focused on celebrating achievements. This lack of professional focus on evaluation eroded faculty members' interest in engaging in faculty evaluation meetings.

Some faculty members disengaged from the evaluation because they found the voting evaluation process hasty, bureaucratic, and formalistic. One participant described the process of voting evaluation in her department as "usually too hasty" in which "urgent" emails about the voting evaluation had become a norm:

The department is always in a state of urgent evaluation meeting from the announcement time. We are always in a state of being late to the deadline. Actually,

people are not mentally prepared at all. Six to seven out of 10 meetings are in a hurry. Once, my head of department forgot the deadline. [The voting evaluation] email [announcement] always has the word “urgent.” (I1T5)

The hastiness in the voting evaluation announcement caused some faculty members to hurry to evaluate themselves and others. Some faculty members just made a rough self-report with insufficient information because they could not remember their own and others' achievements and participation in numerous institutional assigned activities during a year. As some faculty members did not have enough time to sit back and recalculate all the achievements and penalties, they tended to select those who were more outstanding and outspoken than others. Also, the whole faculty one-off meeting where every faculty member and administrator came together to evaluate themselves and others provided limited space for individuals to set their own goals, receive individual feedback, or discuss their own goals and professional learning needs. Some faculty members also found that too many evaluations required their participation during a year. Many of these evaluations were “formalistic like performing rituals” (I1T5). Besides the voting evaluation, faculty members had to complete several other self-evaluation reports or attend several ritualistic kinds of meetings during a year. This administrative and bureaucratic paperwork took away lots of their time while not adding much professional value to their job:

There must be three or four of such evaluations during a year, not to mention that the Trade Union also has an evaluation, and then the Youth Union has another evaluation. Suppose you are a Trade Union or Vietnamese Communist-Party member. In that case, you must be evaluated by the union member or Vietnamese Communist-Party committee, which will ask you what contribution you have made to the faculty as a Trade Union or Vietnamese Communist-Party member. (I7T1)

The bureaucratic paperwork in the public HEI is linked to the culture of distrust within the system. Due to the lack of trust, the central authorities require faculty members to fill in many forms and provide lots of evidence when they submit their evaluation report:

The entire operating system is not based on trust. Too many papers need to be “born,” and I need much proof every time I do something. We must photocopy everything for a simple achievement report at the end of the year, especially related to our research output. (I7T1)

Low Material and Emotional Incentives. The voting evaluation process was also perceived to be ineffective due to the low levels of material and emotional incentives. First, some others found the voting evaluation a weak link between the evaluation with the “carrots

and sticks” (rewards and sanctions) (I3T4). For example, one participant claimed that the materialistic incentives were relatively low and too insignificant to take notice of:

Last year, I got the Advanced Labourer title, but I do not remember exactly how much financial reward I got. Mostly everyone got the same amount, an odd number, so no one remembers it. It's about several hundred thousand VND [under 40 USD]. Those who got the Grassroots Emulation Fighter titles would receive over 1 million VND [over 40 USD]. (I4T1)

The faculty members perceived the low material incentives to engage in their institutions' evaluation and reward practice as linked to their low basic salary. One faculty member shared that she got an average of 7 million VND (over 300 USD) per month, but she needed three times as much to cover the basic expenditure of her family. Many faculty members earned extra income from other activities such as undertaking additional research projects or pursuing part-time jobs outside their institutions. Thus, the voting evaluation material rewards were not strong enough for these participants to fully engage in the evaluation and development of their jobs:

If I don't get the emulation title this year and my reward gets cut off, the bonus is not worth that much. I can work part-time outside to make money. The bonus is nothing compared to the cost I paid for my professional learning, such as attending conferences. (I4T5)

Second, some faculty members did not engage actively in the voting evaluation process because they perceived their administrators to be unjustly moderating the evaluation. Administrators play the mediating role between the institutions and their faculty members. Due to the unfixed faculty evaluation system in the participating HEIs, the evaluation criteria and focus changed from time to time. Thus, the administrators acted as those who interpreted or made sense of the policies then instructed their faculty members to follow them. One participant shared that the voting results depended half on the whole faculty and half on the administrators, so the administrators need to be “extremely fair” (cực kỳ là công tâm) (I4T4). Given the lack of a system to monitor faculty work and the hasty nature of the usual voting evaluation, the administrators' role was even more challenging. As stated by a participant, the administrators' critical mindset and EQ (emotional quotient) were important to the success of setting evaluation goals and criteria that were relevant to the specific context:

As an implementing agent, our administrator needs to see which criteria she can flexibly adapt to suit our department's specific context. This means that the critical mindset of the administrator is very important. It would be better if she set goals for us to follow, rather than let us go freely, then finally evaluate us on a random basis. (I1T1).

However, some participants found that the way their administrators monitored faculty performance and moderated the evaluation meetings did not meet their expectations. One faculty member was discontent that her administrators were tactless when explicitly exerting her power over the choice of deserving members:

I remember the last time [in the voting evaluation], my administrator briefly listed those with scientific research achievements. my administrator [stated her list of] self-selected five members with research publications... Finally, my administrator told us, “you have to do it [use your votes of confidence] right. If you don’t do it right, I still have to review it.” She meant it in a way that she decided for herself. (I1T2)

Some thought that their administrators should focus not only on selecting a few outstanding members but also on recognising other members. The way some administrators ignored the less outstanding members made them feel demotivated:

Our voting evaluation is mainly to select very few people – for example, and it often starts with our administrator saying, “we will have the opportunity to vote for two to three faculty members,” which was a very fixed number of selections. Usually, the administrator tells us, “here, this time, these two faculty members got the title of excellence; others, please try next time!” I don’t think this statement can motivate other faculty members and me to believe that we are highly appreciated... It means that other members are not that important. (I1T5)

Due to the collective nature of the faculty meeting, administrators sometimes had little interaction to evaluate other colleagues accurately. One faculty member who worked in a faculty with over 40 members reflected that her administrator sometimes did not know all the institutional tasks that the faculty members undertook during a year. In such a case, the administrator’s evaluation of faculty members was relatively generic: “our administrator can only evaluate our working attitude by her common sense because there are so many activities that she does not work directly with us” (I4T4).

Conflicts Between the Policy and the Norm. Two constraints that explained the faculty members’ various problem-solving strategies were the conflict between the voting evaluation policy requirement and the collectivist norms among the participants. The national and the institutional voting policies (Grand University, 2015; Law on Commendation and Emulation 2003) require *nguyên tắc lãnh đạo tập thể* (collective leadership principle) that espouses everyone is a leader of a community and thus organisational decision making is decided by a majority of the community members. Collective leadership is one typical principle

of the Vietnamese communist *nguyên tắc tập trung dân chủ* (democratic centralism principle)²² where every member's performance must be considered and approved by at least two-thirds of the whole faculty members (administrators included):

Everything needs to be considered by collective opinions and decided on by the majority. After the collective opinion has been selected by voting, we have to comply, even if we do not like it. It is called concentration. Then any opposing views against the agreed collective opinions are considered destructive, deteriorating, or self-transforming. That is the requirement, so everyone has to follow. (I7T1)

During the evaluation meetings, faculty members evaluated both themselves and others based on the democratic centralist principle. Thus, individual faculty members needed to be cautious not to act against the administrators' collective voices or the prominent or outspoken members. Only a limited quota of approximately 15% of faculty members was nominated and voted for outstanding titles. Selected faculty members need to receive more than two-thirds of the votes to be considered in the following selection rounds. Thus, individual faculty members did not have the total freedom to select and vote entirely based on their preferences. If they did so, the voting results would be scattered, and thus the selected members may not have more than two-thirds of the votes. Thus, many departments and faculty had "unwritten rules" to select outstanding members.

One faculty member mentioned the "unwritten rule" where his whole faculty compromised over whom to select based on prioritising other criteria than on selecting who was more deserving. For example, his faculty tended to vote for those who had the outstanding title at the institution in the previous year because one would get upper institutional outstanding titles for having lower institutional outstanding titles in 2 consecutive years (Grand University, 2015). His faculty would then take turns to allocate members to get the upper institutional outstanding levels. In this way, the whole faculty can also upgrade its outstanding collective title to a higher level. The written rule reinforced faculty members' colluding with others' suggestions of members for limited outstanding titles, even when they thought the chosen ones were not deserving. For this reason, the participant thought that it was not worth fighting against those rules when everyone seemed comfortable with them:

²² Democratic centralism involves calling all members to give their opinions on an issue before the final decision is made through a vote. The choice or the person that gets a major share (from about 50%–60% or above) will win. This principle does not encourage individual voices once a collective decision has been made, at which point, everyone must follow it unquestioningly. Those who show opposing views to the collective decision or against the Communist Party directions can be considered as destructive" (phá hoại), "deteriorating" (suy thoái), or "self-transforming" (tự diễn biến/ tự chuyển hóa) with a negative connotation, from <https://tuyengiao.vn/nghien-cuu/ly-luan/tap-trung-dan-chu-nguyen-tac-can-ban-de-xay-dung-dang-trong-sach-vung-manh-134199> and <http://hvlq.vn/tin-tuc/tin-trong-nuoc/nhan-dien-nguy-co-tu-dien-bien-tu-chuyen-hoa-trong-noi-bo-va.html>

I care about evaluation, but it's not worth fighting for! And it is not so necessary that because of all the pay and benefits... if I get it, others will not. Except for exceptional cases of merit, the awards will be nominated first; otherwise, it will be in order. If one person wins one year, another he will be selected the next year. There are principles that everyone finds quite reasonable because only in that way our department title is upgraded. In public organisations, there are always unwritten rules. For example, there is no need to compete for anything, just sequentially... Everyone is willing to support that. That is what ultimately maintains the so-called colleague, the very special co-worker relationship. (I7T1)

However, the fair and accurate evaluation among faculty members contradicted the Vietnamese participants' norms of avoiding conflict and maintaining harmonious relationships among people. When interacting with the Vietnamese collectivist norms, the communist-embedded collective leadership principle resulted in faculty members' less honest and less fair evaluations of themselves and others. The collectivist norms expect individual members to suppress their personal voices to follow the collective decision to maintain organisational unity. Thus, faculty members tended to evaluate based on the perceived consequences of their ratings rather than on the performance of those involved. For example, faculty members usually rated themselves and others leniently as having completed all tasks and for the most basic titles of excellence for two reasons. As the basic titles of excellence were unlimited, there was no conflict of interest among faculty members over these least desirable titles. The faculty members did not want their colleagues' financial bonuses to be deducted for their low ratings. This relationship-over-rule norm was quite common among Vietnamese people:

Usually, everyone would be rated as completing all tasks... Our Vietnamese people are affectionate, and we mainly base our decision on our affection towards others. Some members are sometimes performed very poorly, but we just forgive it... because we would feel pity for them if their bonuses were deducted. (I2T1)

I think it (lenient evaluation) is all the same in Vietnamese organisations and units. In some cases, even the violation of the regulation, after consideration, was then proposed as "light misconduct" and "let bygones be bygones" sort of thing. That is why no one would be graded at the low level, and so all would receive grade A! That is how the evaluation is typically going on here. (I2T2)

Consequences of Faculty Members' Approaches

Overall, faculty members fulfilled their roles as required by the voting evaluation policy (Grand University, 2015). However, the voting evaluation had a limited impact on the faculty

members' professional learning and improvement. Also, the voting evaluation resulted in some ambivalent thoughts and emotions among the faculty members (see Figure 14).

Voting Evaluation Policy Requirements Fulfilled. Faculty members' compliance with the voting requirements fulfilled their roles. The whole faculty evaluation meetings became events where the administrators and all faculty members came together to officially consider the faculty completion of assigned tasks and disciplinary conformity.

The compliance with the voting process was helpful for the faculty members to receive feedback from others, especially in the context of good administrators' leadership, mutual trust, and a supportive environment. The faculty members' fulfilment of their roles during the voting evaluation encouraged them to strive to meet the minimum expectations of the work discipline and the required workload: "we tried not to be caught over disciplinary issues ... to complete enough teaching and research hours – to fulfil my responsibility" (I3T3). Some of the faculty members were motivated to strive their hardest to meet or exceed the requirements of their HEIs. One of them commented:

One [faculty member] must first complete the task, how many hours a year one has to teach, and how many hours of research one spends a year. It also acknowledges your accomplishments in the past year. When I first entered the university, I did not know how I was supposed to be evaluated. Then I found the research point was 1.5 times higher than others, so I was determined to double my publication the next year. (I3T4)

Besides, faculty members also found voting evaluation as a source of warning, which resulted in an increased sense of discipline among themselves: "We know that we will always be judged and seen as being watched and scrutinised, so we have a better sense of self-discipline, for example being more punctual to class or being more mindful about grading timetable" (I4T5). Overall, the participant's compliance with the voting evaluation helped maintain the participating HEIs maintain institutional discipline and stability.

Limited Impact on Faculty Members' Learning and Improvement. Except for a few faculty members who benefited from quite helpful feedback and comments from other colleagues during the evaluation meeting, many faculties did not have a similar positive experience. Due to many faculty members' disengagement, the voting evaluation tended to have a limited impact on their learning and improvement. The faculty members' perfunctory engagement and some collectivism-embedded strategies for solving problems prevented the participants from accurately and fairly evaluating themselves and others. In many cases where there was little professional learning and sharing among faculty members, administrators and faculty members did not have the chance to learn and understand each other's work. In other words, the administrators and faculty members were deprived of the opportunity to reflect on their practice and learn about their effectiveness.

Faculty Members' Conflicting Thoughts and Emotions. Although everyone seemed to accept the voting evaluation practice, many shared conflicting thoughts and emotions about it. While accepting the necessity of the voting evaluation, many faculty members expressed negative emotions, such as confusion, discontent, demotivation, and indifference. For example, one faculty member felt that the voting process of “bầu vo” (raw voting), which meant everyone evaluated themselves and others based on their own subjective opinions, caused him lots of confusion over the evaluation processes and results. Worse, it potentially caused unjust evaluation outcomes that sparked unspoken disagreements and discreet resistance among some faculty members. One faculty member even reported that she felt “invisible” for having never been recognised by her administrators during the voting evaluation. Some other members refused to take more assigned responsibilities or engage in more professional and administrative services, reasoning that they were consistently rated at the lowest titles of excellence. Another faculty member expressed her common feeling of tension during her faculty evaluation meetings, even when she was voted with the outstanding titles. The faculty member was personally uncomfortable about her titles of excellence because she still sensed her colleagues' disappointment for not being recognised during the voting evaluation meetings:

There are times when I go to meetings where I feel like I'm really feeling the negative attitudes [from my colleagues]. For example, when I sit next to Ms. Y [who was not usually voted with outstanding titles], she says the sentences that I feel... how do I feel... I feel sorry for her. She said it in a way that was “then everyone stops complaining about being tired [by faculty work], you've got money and titles!!” ... After all, sometimes I feel like, “why do meetings feel so suffocating that it doesn't create a common opportunity for everyone?” ... Because only a few people hold key positions in the department [were voted] ... then I see, oh my god, sometimes there aren't [voted with outstanding titles].... because the job position has decided it like that. Then if it stays like this, where is the chance for the rest?

Given many faculty members' negative feelings and attitudes towards the voting evaluation, it is not easy to enact the evaluation for learning purposes.

Concluding Remarks

The three participant groups (i.e., QA officers, administrators, and faculty members) took various approaches to the voting evaluation, including complying, problem-solving, and disengaging. These findings confirm the use of the voting evaluation for accountability and suggest more limited use of voting for improvement purposes.

First, voting evaluation for accountability purposes (Lonsdale, 1998) was most prominent among the approaches. The accountability-oriented approaches included administrators and faculty members complying with the policies by monitoring, reporting, and attending faculty evaluation meetings to evaluate the performance of themselves and others. These approaches satisfied the managerial, administrative, and practical purposes by maintaining the HEIs' stability and individual productivity. These approaches also fulfilled the communist ideological purposes of extending the large-scale emulation movement across Vietnamese public organisations. However, the collective leadership (which represents the democratic centralist principle), combined with the collectivist norms, prevented the administrators and faculty members from undertaking fair and transparent evaluation.

Second, the improvement-oriented approaches were evident in a few faculties, and they depended mainly on the willingness and leadership of the responsible administrators and faculty members. The improvement-oriented approaches to evaluation (Lonsdale, 1998) involved maximising both institutional and individual faculty performance. These approaches involved some administrators improvising rewards or adapting evaluation forms at faculty levels. These approaches recognised various aspects of faculty performance and motivated them to reflect on their performance more effectively.

Third, the use of voting evaluation for learning and transformation purposes (Dahler-Larsen, 2009) was absent at the faculty and institutional levels. The administrators and faculty members had limited authority to exercise their ownership over the voting evaluation implementation process. The collective evaluation process, the democratic centralism, and the collectivist norms made it unfeasible for the participants to use evaluation for faculty members' learning and transformation.

Chapter 6: Discussion

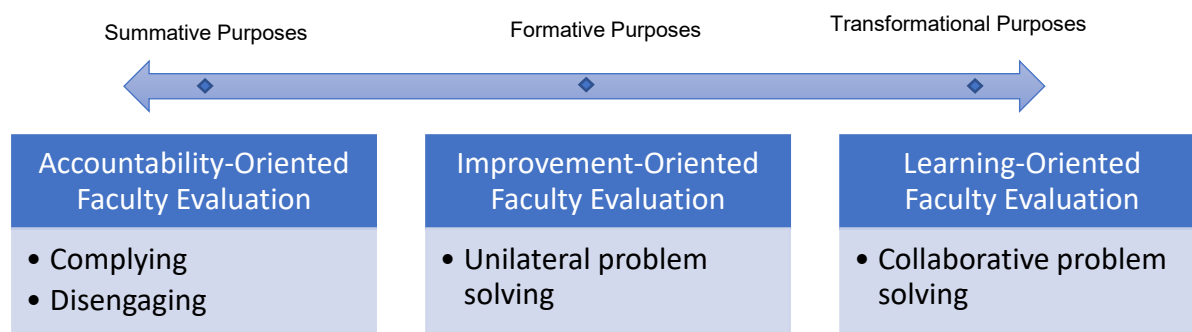
This chapter discusses key actions and constraints of the SET and voting evaluation practices against the Vietnamese and global higher education context of faculty evaluation. It then discusses the improvability of the current approaches to faculty evaluation in the Vietnamese HEIs, followed by some implications for improving the practices.

The Predominance of Accountability-Oriented Approaches

My finding chapters reveal the participants' theories of action about SET and voting evaluation practices. The participants' approaches to these faculty evaluation practices were complying, disengaging and problem solving. While the complying and disengaging approaches can be categorised into the accountability-oriented faculty evaluation (i.e., evaluation for accountability), the problem-solving approaches can be either improvement-oriented or learning-oriented (i.e., evaluation for learning and improvement) (see Figure 15). My study reveals that the faculty evaluation for accountability was more dominant than the evaluation for learning and improvement.

Figure 15

The Participants' Faculty Evaluation Orientations



First, the dominant approaches of complying and disengaging were consistent with an accountability-oriented evaluation which emphasises using external requirements, incentives, rewards, and sanctions to promote established outcomes and productivity (Lonsdale, 1998). All the participants complied with the faculty evaluation policies required by the government, the MoET, the Grand University, and the HEI's board of rectors. When there were no problems with faculty members' underperformance, many participants disregarded or did not use the results for improvement or learning purposes. Overall, these accountability-oriented approaches ensured that the participants satisfied the external requirements during the evaluation process, for example, those related to accreditation or disciplinary management.

Second, there was some evidence that the participants solved problems and used the SET and voting evaluation to improve faculty performance. The participants' problem-solving approaches were primarily unilateral, which means they resolved the problems of practice by themselves rather than by collaborating with other stakeholders. For example, QA officers in one HEI adapted the SET forms and reports to make SET more informative, or some faculty members adjusted their teaching to suit students' demands. The participants' unilateral problem-solving approach was somehow consistent with improvement-oriented evaluation (Lonsdale, 1998) that aimed to optimise institutional and individual performance. However, the participants' efforts towards improvement were ad hoc and fragmented, unlike the systematic efforts to link faculty evaluation with faculty development in more developed higher education systems (e.g., Norsworthy & Sanders, 2021; Roxå et al., 2021).

Third, there was rare evidence of the participants adopting a collaborative problem-solving approach to improving faculty professional learning and teaching. For example, an administrator inquired students and faculty members through class observation, class meetings and giving feedback to those with SET underperformance. This approach was consistent with the learning-oriented approach that emphasises faculty members' increased ownership, collaboration, and adaptation to use evaluation to transform practices (Dahler-Larsen, 2009). This orientation is linked to responsive constructivist evaluation (Guba & Lincoln, 1989), which considers evaluation a dynamic process of various agents co-creating evaluation outcomes relevant to their value system and professional interests. The participants' collaborative problem-solving efforts reflect organisational learning (Argyris & Schön, 1996), in which their values of faculty evaluation shifted from satisfying managerial requirements to a means of collaboration into the real problems of teaching and learning. Unfortunately, the organisational learning evidence in my study was single-loop learning which means changes in behaviours without changes in underlying values at the institution-wide level (Argyris & Schön, 1996). The participants' collaborative problem-solving strategies were limited to individual and faculty levels and not shared and converted to action nor "embedded in the organisational environment" (Argyris & Schön, 1996, p. 16). The lack of shared goals across faculties and institutions limited the extent of double-loop learning or an institution-wide change in the culture of evaluation for learning (Argyris & Schön, 1996).

Overall, the participants' overemphasis on evaluation for managerial accountability limited the impact on faculty members' learning and improvement. Very few participants of my study recognised SET as a source of inquiry for learning and improvement, and few focused on authentic feedback about the quality of individual faculty members' performance in the voting evaluation. The key constraints of the participants' approaches will be discussed in the following sections.

Managerial Accountability: A Better Explanation of Compliance Than Institutional and Individual Passivity

The participants' approaches were oriented more towards managerial accountability (Røiseland et al., 2015), which means satisfying accreditation and managerial bodies through external incentives. However, these approaches did not satisfy professional accountability (Røiseland et al., 2015), which means improving faculty members' learning and improvement based on their internal values and incentives (Røiseland et al., 2015). The participants' focus on managerial values in the Vietnamese HEI participants' approaches to faculty evaluation were in line with those in various HEIs worldwide, where neoliberal ideas and performativity culture focus on technocratic input-output relationships rather than humanistic harmony between HEIs and individuals (Blanco & Metcalfe, 2020; Lyotard, 1979/1984). In the Vietnamese context, MoET staff mainly reinforce HEIs' compliance to reporting performance through bureaucratic processes that are superficial and limited in improving practice (Hoang, 2017; Lam, 2018, as cited in Salmi & Pham, 2019). Furthermore, the rise of the performativity culture is exacerbated by the inherent "diseases" of achievement (Huynh, 2016, p. 198) commonly observed in the Vietnamese education sector. My findings add to the existing literature by drawing attention to potential threats to faculty evaluation professional accountability values in Vietnamese higher education.

More importantly, my findings on the HEIs' managerial accountability (Røiseland et al., 2015) contribute an alternative explanation for accountability in Vietnamese higher education. Previous authors (B. Bui et al., 2017; Hoang, 2017; Quynh, 2017) would describe Vietnamese academics' compliance as *passive accountability*, which means that HEIs and individuals are mainly responsible for complying with the national regulations, fulfilling the assigned task, and reporting to the management boards. H. T. Pham and Nguyen (2019) concluded that accountability in Vietnamese higher education most generally means a power-related activity between those with power and those being evaluated to monitor and control their performance, function, and authority. The Vietnamese HEIs' passive accountability is argued to be a predominant legacy of the Vietnamese government's centralised governance, which involves vertically assigning responsibilities, encouraging submission, and punishing disobedience (Hoang, 2017). The Vietnamese centralised governance has hindered HEI leaders, administrators, and faculty members from flexibly adapting their practice at the faculty level (T. N. Pham & Goyette, 2019; Salmi & Pham, 2019). While the Vietnamese participants in my study took a predominantly compliant approach to faculty evaluation, these approaches are not "completely passive" as others implied because the MoET and Grand University faculty evaluation policy demands legitimised their compliance. Rather than being passively obedient, the participants in my study responded dutifully to the evaluation for practical benefits at both the institutional and individual levels.

At the institutional level, the participating HEIs used faculty evaluation to promote individual achievements in recognised activities such as producing international and national publications. Individual achievements enabled the HEIs to fulfil their managerial accountability and accumulate their collective achievements, leading to increased institutional prestige and potentially more funding or allowance of higher student enrolment quotas. The HEIs' practical purpose for using faculty evaluation reflects what Salmi and Pham (2019) described as a *quid pro quo* relationship between the HEIs and the authorities (e.g., the university, MoET, and Vietnamese government), in which HEI rectors and staff adhere to "specific instructions considered by their line-management authorities to be important" in exchange for their higher autonomy to "manage their institutions without much constraint" (p. 109).

At the individual level, the participants made a practical choice to comply with the evaluation system though they were aware of the pitfalls of the system. Several participants commented that they complied with the evaluation requirements because they did not want to be publicly reprimanded for not having completed their assigned tasks during the evaluation process. The participants' fear demonstrated that the compliant approach was not due to their passivity but to their institution's rule-by-sanction strategies are consistent with the manipulation commonly used by the Vietnamese communist state-centralised regime to maintain its stability and legitimacy (Homutová, 2018).

Overall, the methodology used in this study has revealed that the HEIs and individual participants' practical views of the evaluation were a more robust explanation for compliance than the typical portrayal in the existing literature of compliance as a cultural trait and of Vietnamese academics as having a "culture of obedience" (Truong, 2013, p. 19). My findings support Spillane et al. (2002), who suggested that policy implementation problems should not simply be attributed to any specific patterned behaviours of the agents. Instead, the participants' situated cognition (Spillane et al., 2002), which means their understanding of the situation and context, significantly shaped their policy sense-making and their choices of approaches to faculty evaluation.

Multiple Perspectives Towards Harmony-Oriented Approaches to Problem Solving

My study found that one heavily weighted constraint that strongly influenced many participants' unilateral problem-solving approaches was their so-called norm of *đĩ hòa vi quý* (harmony orientation). The harmony-oriented norm refers to the participants' perceived obligation to act tactfully and diplomatically to maintain their group's cohesion and stability. My findings were consistent with previous studies (M.-P. Nguyen, 2008; Truong et al., 2017), who found that harmony is the utmost goal for determining relationships and communication among Vietnamese people. In my study, the harmony-oriented norm was evident in the participants' following summary of their behavioural patterns:

- Showing respect to colleagues, especially the more senior members
- Trying not to interfere in other members' teaching
- Ignoring other members' weaknesses
- Avoiding commenting negatively or underrating other members
- Being humble or modest about own achievements
- Prioritising others over selves for outstanding title nomination and selection
- Making decisions based on *tin* (heart/relationship/emotional drive) rather than *ly* (head/rules/rational drive)

Truong et al. (2017) also found that Vietnamese teachers tended to suppress their personal opinions to maintain their group cohesion and unity, and they unquestioningly accepted hierarchical authority for seeing themselves as subordinates. My study expands extant literature by demonstrating that these Confucian hierarchical norms were also evident in Vietnamese HEIs among both subordinate (i.e., faculty members) and the mid-level leaders (i.e., administrators). Many Vietnamese participants in my study perceived the obligation to suppress individual needs (e.g., prioritising others over selves) and accept hierarchy (e.g., showing respect to senior faculty members).

While harmony is a commonly recognised norm, scant literature has explored its relations with problem solving in Vietnam, which has a Confucian heritage context (Truong, 2013). In my study, harmony acted as a double-edged sword in faculty evaluation. While the participants' harmonious strategies maintained their organisational stability, the harmony-oriented norms prevented many from giving genuine comments about SET underperformance and fair evaluation of others during voting meetings. Therefore, my findings confirmed that the Vietnamese participants' collectivist Confucian harmony norm conflicts with the communist collective leadership principle where Vietnamese public employees like those in my study vote among themselves for outstanding titles. My constraint analysis provides harmony-oriented norms as a sound explanation for Vietnamese public employees' patterned "diseases" during annual voting processes (*bệnh mùa bỏ phiếu*) (Song, 2021, n.p.).

My findings are in line with previous studies (Carson & Nelson, 1996; Pham, 2018; Truong, 2013), which highlighted that an orientation towards harmony in a collectivist culture in China and Vietnam prevents teachers and students from contributing their diverse, critical, and potentially valuable ideas. My study also provides empirical evidence of some unintended consequences of the Vietnamese harmony-oriented strategies for solving faculty members' underperformance and confirms that the policymakers' failure to recognise the participants' cultural norms leads to less effective evaluation for learning and improvement. Therefore, my findings support several authors (C. P. Nguyen, 2020; P.-M. Nguyen, 2008; Pham & Renshaw,

2013) who advocated education reform in Vietnam to pay more attention to sociocultural and emotional dimensions in initiatives for change.

Furthermore, my study contributes to the existing literature by demonstrating that the concept of the multifaceted concept of harmony is applicable in the Vietnamese context. Harmony in collectivist communication may manifest in two complementary aspects of *yin* (i.e., submissive, ongoing interaction to attain the state of shared interest and mutual respect) and *yang* (i.e., dominance, opposition to fellowship) (Chen, 2011). Some Vietnamese participants in my study demonstrated the *yin* aspect of harmony by *hi sinh* (sacrificing) their own benefits in exchange for group harmony. In contrast, others demonstrated the *yang* aspect by making a joint effort to engage various stakeholders in using evaluation for learning and improvement. Overall, the Vietnamese participants were oriented towards the *yin* aspect of harmony since many adopted a *cả nể* (people-pleasing) – a more submissive and less direct approach when dealing with conflicting situations.

Also, the constraint analysis of my findings reveals a link between two concepts of “harmony orientation” in the non-Western culture with the “conflict avoidance” prevalent in the Western literature. While maintaining harmony does not initially mean avoiding conflicts, many participants in my context equated it with taking a gentle and diplomatic approach to solving faculty evaluation performance issues. My Vietnamese participants’ harmonious approach to problem solving is consistent with what Patuawa et al. (2021) termed *control-focused problem solving*, which commonly involves prioritising relationships over the task or avoiding conflicts to control the process unilaterally. Thus, my Vietnamese participants’ approaches to faculty evaluation were linked more to Model 1 secretive, superficial, and nonconfrontational problem solving rather than Model 2 co-operative and joint decision making in problem solving (Argyris & Schön, 1978, 1996; Cardno & Piggott-Irvine, 1997; Rutherford, 1992). However, while Western scholars (Anderson, 1994, 1997; Argyris & Schön, 1978, 1996) saw conflict avoidance as negatively suppressing conflicts and leaving problems unresolved, such views might not be shared by the Vietnamese participants who found conflict avoidance appropriate in their collectivist Confucian culture. Thus, my study challenges the assumption of existing models of problem solving which prioritise certain norms such as direct confrontation of those involved, as this might potentially upset the harmony-oriented norms taken for granted by the Vietnamese participants. My study calls for future studies to investigate the context-sensitive theory of problem solving in collectivist Confucian countries like Vietnam.

Exceptions to the Rule: Understanding Why Some Individuals Act on Their Own Values

A few participants in my study followed their desire to act according to their own professional, ethical, and educational values. Despite being scarce, these findings were

noteworthy because these participants' normative view and their so-called self-reliant²³ approaches to faculty evaluation counter the negative effects of the Vietnamese communist centralised governance (L. H. Phan & Doan, 2020) and the neoliberal value of performativity (Cherry et al., 2017; Sułkowski et al., 2020). For example, an administrator's collaboratively solving faculty members' SET underperformance reduced individual competitiveness resulting from centralised requirements combined with the neoliberal value for excellence and productivity. Similarly, a few faculty members' inventions of their own evaluation criteria to evaluate other colleagues' performance reflects their discreet breaking free from the Confucian norms of individual conformity to authority or suppression of personal desire (P.-M. Nguyen, 2008).

The participants' self-reliant approach (i.e., acting on their own values that broke free from political and cultural influence) resonates with Buddhist principles on learning and transformation in the Vietnamese context. Under the Buddhist views, learning is ongoing reflexivity, and the moral standard for individual acts should be based on specific contexts (Chu & Vu, 2021; Thich Nhat Hanh, 1987). Though most participants did not claim to follow any religion, their acts and values were consistent with Buddhism, which is considered a significant religious, spiritual or philosophical influence on Vietnamese people (Vuong et al., 2018). As N. T. Nguyen (2019) commented: "while Confucianism is incorporated into the ruler's ideology system and used to define the past, Buddhism may have shaped Vietnamese in popularity regardless of their social classes. It cemented Vietnamese spiritual souls" (p. 25). Given the increasingly lower status of Confucianism and the more significant role of Buddhism in contemporary Vietnam (Chu & Vu, 2021), this is one possible explanation for the findings. My study suggests that future research investigate the influence of Buddhist philosophy in contemporary Vietnamese higher education practices, including faculty evaluation.

Overall, my study adds to the extant literature by suggesting that evaluation practices need to be understood in their religious context, spiritual values or philosophical backgrounds, not just in their cultural one. Confucian principles, which are associated with pragmatic and political purposes of learning (Chu & Vu, 2021; H. T. Ngo, 2019), are more related to the notion of formative purposes (Scriven, 1991) and improvement-orientation evaluation (Lonsdale, 1998). Evaluation in this view concerns both the improvement of individuals and organisations, but decision-making power belongs to the dominant group, and evaluation primarily serves political and managerial purposes. In contrast, Engaged Buddhism concerns self-transformation, reflexivity and contextuality (Chu & Vu, 2021; Thich Nhat Hanh, 1987), so it relates more to the transformational purposes (Patton, 1996) and learning orientation of

²³ Being *self-reliant* here means the ability to depend on oneself, on one's own ability, and to act on one own's values. Buddhists believe that "believe that while others can exert an influence on someone's life, the individual will in the end create his own kamma and be responsible for his own actions" as cited from <https://www.budsas.org/ebud/whatbudbeliev/191.htm>

evaluation (Dahler-Larsen, 2009) which promote practitioners' reflexivity, adaptability, and increased ownership to act effectively in each specific evaluation context. The Buddhist principles of context sensitivity might contribute to the theory of problem solving in Vietnam and other collectivist Confucian nations, as suggested earlier. For example, an alternative theory of problem solving could involve applying *skilful means* which "indicates a contextualised rather than a universal approach in response to concerns over contemporary issues" and "is applicable to contemporary organisational workplaces, which require a high degree of flexibility and reflexivity at many levels" (Chu & Vu, 2021, n. p.). My study suggests that skilful means (Chu & Vu, 2021) could be a way forward to problem solving in my study context as it matches with the Vietnamese participants' philosophical and cultural worldview while promoting active engagement, cultivating connectedness and embracing the diverse experience of those involved.

Improvability of the Current Theories of Action

This section explores how changes could increase the likelihood of desired outcomes being achieved and evaluation impacting learning and improvement. The improvement of a ToA is possible if the existing constraints can be altered or their significance can be reduced or removed (Robinson, 1993).

Can Faculty Evaluation be Improved?

My participants' theories of action concerning faculty evaluation are minimally improvable because the driving constraints related to the policies and culture are not easily alterable. However, there are some possibilities for change at the grassroots level.

It is challenging to revise the Vietnamese participants' theories of action about faculty evaluation for several reasons. First, the increasing focus on managerial purposes of faculty evaluation has been an unavoidable trend not only in Vietnam (K. A. Le et al., 2019; London, 2010; L. H. Phan & Doan, 2020) but also in many other countries (Appel, 2020; Macfarlane, 2021; Posselt et al., 2020). In the Vietnamese case, the neoliberal idea of market orientation has been integrated into the communist centralised governance, further intensifying faculty evaluations' managerial and political purposes. Thus, the voting evaluation is easily open to improvability because it belongs to the national emulation campaigns under the Vietnamese communist centralised governance that demands compliance rather than adapting practices (T. N. Pham & Goyette, 2019; Salmi & Pham, 2019). Unless the Vietnamese Communist Party's leaders and policymakers think otherwise, the HEIs' voting practice will remain under the government's control and follow the same patterns nationwide. Also, the Vietnamese participants' views of harmony orientation as suppressing personal wishes and avoiding conflicts prevented effective use of SET and voting evaluation for improvement. As discussed

earlier, these cultural norms are deeply embedded in the Vietnamese participants' ways of viewing and doing things, so they are not easily changeable.

However, even if several key policies and cultural approaches cannot be changed, the current theories of action regarding faculty evaluation are possible to be revised in several ways. First, many Vietnamese academics in this study have a learning attitude related to the Vietnamese culture of learning appreciation (N. T. Nguyen, 2019; P.-M. Nguyen, 2008). Many participants in my study tried to solve problems within the evaluation system to make the process more effective. For instance, the QA officers at HEI 2 developed their own SET forms and reports to address their institutional needs for quality feedback from students. The fact that some participants were also not passively compliant with the system predicts the possibility that the current system might be open to revision. Second, there has been a recent call for a more active form of accountability in Vietnam, which moves “beyond administrative compliance to holding stakeholders, both at the organisational and individual levels, accountable for the outcomes of their performance of duties and missions” (Hoang, 2017, p. 2). The need for transforming accountability could be a catalyst for Vietnamese HEIs' increased attention to professional accountability, especially faculty evaluation for learning and improvement.

How can Faculty Evaluation be Improved?

The possibilities for improving the original ToA are implicit in the ToA itself (Robinson & Lai, 2006). An improved ToA needs to alter the weighting of a constraint in ways that satisfy all the constraints, not just those preferred by the researcher or an intended group of participants (Robinson & Lai, 2006). Given the current theories of action of the participants, this study suggests improving faculty evaluation based on the revision and addition of existing constraints as follows.

Revision of Existing Constraints. One typical constraint that influenced the participants' disengagement from using evaluation for learning and improvement was the low quality and relevance of the evaluation focus, data, and processes. First, many faculty members expressed concerns over SET data due to the QA officers' improper collection process and students' lack of serious attitude to SET. Thus, one recommendation for improving the participants' theories of action is for QA officers to check the appropriateness of the evaluation process and the validity of the SET data. The evaluation data validity check involves making sure “what is reported” accurately reflects “what it is meant to represent” (Robinson & Timperley, 2000, p. 68). Many faculty members in this study commented that they would have engaged in using SET for learning and improvement if SET accurately reflected their teaching quality. Furthermore, the HEIs' hasty and bureaucratic evaluation processes with numerous evaluations and time-consuming paperwork limited faculty

members' active engagement. Thus, one more suggestion is to improve the system of managing faculty performance data. This system needs to integrate all faculty members' work into a single platform, which would help reduce administrative duties, including repetitive steps of reporting performance. Improving the quality of the evaluation processes would increase the likelihood that faculty members evaluate data for learning and improvement.

Addition to Existing Constraints. My study found that the participants' current approaches ruled in political and cultural constraints while ruling out individual constraints. Given the participants' current theories of action, individual constraints need to be added to improve the situation. Overall, my study suggests that the Vietnamese faculty evaluation policymakers can integrate individual needs for professional learning and improvement by reconceptualising SET and voting evaluation and treating faculty underperformance as a collective problem.

Voting Evaluation as a Learning Opportunity. Despite a normative view of evaluation as a learning opportunity, many participants in my study saw voting as an administrative process of just getting paperwork done to satisfy the managerial demands of the top-down management. However, voting evaluation needs to be defined as a social practice involving various stakeholders' needs (Saunders et al., 2011) rather than a top-down and one-off instance. A social-practice view is what agents do following their values, beliefs, and perceived meaning of evaluation either individually or at faculty, institutional, or national levels (Robinson, 1993; Saunders et al., 2011). Including individual faculty members' need for professional learning in voting evaluation means collectively viewing it as an opportunity for faculty members to learn from each other rather than a bureaucratic administrative process. For example, some administrators combined voting evaluation with mentoring less experienced faculty members. The faculty members in my study felt more satisfied with the voting evaluation when they gained professional knowledge from other members during their evaluation meetings.

SET as a Means for Inquiry Into Student Learning. SET needs to be seen as a means for inquiry into student learning rather than an evaluative means to judge faculty performance. First, some faculty members viewed student feedback as an effective inquiry into student learning. These participants conducted their own informal collection of student feedback to inform their improvement of teaching. My finding contrasts with previous studies (K. D. Nguyen, 2000; H. T. Pham, 2014, 2018; H. T. Pham & Nguyen, 2020), which attributed Vietnamese academics' lack of engagement in SET to the cultural norms of teacher-centredness and to their perception of SET as a "Western concept." This finding supports X. B. Tran (2010), who found that collecting student feedback to inform teaching and learning improvement is not a foreign practice but has been an ongoing practice of Vietnamese faculty members since the first Vietnamese HEIs were established. Many faculty members in my

study appreciated if SET were designed to reflect their teaching effectiveness and student learning, rather than their popularity quantitatively measured through questioningly valid SET processes.

Second, my study reveals a mismatch between the SET policymakers' conceptualisation of SET to identify faculty underperformance and the faculty members' normative view of student feedback as an inquiry into student learning. Because SET policy (Grand University, 2014) focuses on teaching staff and whether they are above or below the threshold, many participants in my study did not perceive SET as appropriate means for teaching and learning reflection. My findings support previous authors (N. D. Tran, 2015; N. D. Tran & Nguyen, 2015) who suggested that SET in Vietnam should focus on students and their learning rather than on judging faculty members. Outside Vietnam, various studies worldwide have shown a similar lack of favourable attitude amongst academics to teacher- or teaching-oriented SET, including academics in the USA (Medlen, 2019), New Zealand (Mutch & Tatebe, 2017), Norway (Borch, 2021), Canada (Vargas-Madriz, 2019), and Israel (Hammer et al., 2018). Also, the call for reconceptualising SET or learning-focused approaches is evident in Vietnam and elsewhere in the world (Borch, 2021; Darwin, 2016; Nicol, 2019; Roxå et al., 2021) (Nicol, 2019). A commonality among these studies is the possible but not yet widely enacted approach to SET that improves the teaching and learning process. My study supports previous authors (Mutch & Tatebe, 2017; P. V. Nguyen, 2020; Norsworthy & Sanders, 2021; Roxå et al., 2021; N. D. Tran & Nguyen, 2015; T. T. Tran, 2018) who have advocated that SET policy makers and designers incorporate the values and beliefs of faculty members and students when designing SET policies. Overall, my study suggests that focusing on how the results would support teaching and student learning is a subtle but potentially important shift that could result in SET being viewed as appropriate and used as a source of feedback even when the scores are above the acceptable threshold. Instead of overusing quantitative measures, SET can be designed in ways that are more valid by removing meaningless questions, including more qualitative questions and eliminating bias in students' responses (T. T. Tran, 2018). SET can also be implemented in ways that respect the collectivist Confucian norms of respecting learning and sharing among faculty members and students (P. V. Nguyen, 2020; N. D. Tran & Nguyen, 2015; T. T. Tran, 2018).

Treating Underperformance as a Collective Problem. The collectivist Confucian norm can be leveraged in ways that promote collaborative problem solving. My earlier constraint analysis in this chapter shows that some participants' professional, ethical, and education evaluation values were at odds with political and managerial demands. My finding confirms the heightened demands for faculty performance from market-oriented performance management and state-centralised government in Vietnam and global higher education (L. H. Phan & Doan, 2020). While it seems impossible for faculty to escape from these political and

managerial forces, there are possibly several ways for them to navigate the system. One navigation strategy involves collectively addressing the evaluation underperformance problem within a professional community as a whole rather than as a problem that an individual has to fix. This approach aligns well with the collectivist nature of Vietnam and the importance of harmony-oriented interpersonal relationships among Vietnamese academics (M.-P. Nguyen, 2008; Truong et al., 2017). For example, the combination of collaborative peer evaluation of teaching and SET through nonformal professional learning facilitated Vietnamese faculty members' teaching reflection and transformation (P. V. Nguyen, 2020).

As Coburn (2006) stated, professional, proximal communities and informal networks can be places where educators can engage in conversations that inspire members with new approaches to understand and resolve the problems. Interestingly, like my study suggestion, non-Asian scholars also advocate integrating a collective focus on performance management. For example, various international scholars in the USA, Canada, and New Zealand (Grant, 2017; Mountz et al., 2015; Mutch & Tatebe, 2017) share various strategies where academics can individually or collectively alleviate the performance requirements imposed upon them. These efforts also aim to strengthen individual and collective resilience to "resist being coerced into compliance" and to "put the heart back into the neoliberal university" (p. 231). Examples of these strategies include academics reflecting and sharing their own experience (Grant, 2019) or academics collaboratively engaging in conversation about a shared commitment to scholarship, teaching, and service informed by an ethics of care (Mountz et al., 2015). Another more practical strategy is co-teaching, in which two faculty members work reciprocally to plan courses, share resources, and give feedback on each other's teaching (Mutch & Tatebe, 2017).

Overall, my findings show that collectively focusing on faculty performance problems will likely lead to improvement. This involves the administrators and faculty members treating the quality problem as a collective one and working out ways to solve it collaboratively. For example, the problem of SET underperformance can be solved by peer observation of teaching or inquiry into student learning issues. This recommendation is consistent with T. T. Tran (2018), who argued that there are some opportunities for implementing agents to "fill in the gap" to turn a managerial demand like SET into "a powerful tool to serve teaching improvement" (p. 6) of faculty members.

Chapter 7: Conclusion

My thesis investigates faculty evaluation in contemporary higher education, where neoliberalism or market-oriented evaluation dominates various aspects of educational practices (see Chapter 1). My study proposes viewing summative, formative and transformational evaluation purposes in a continuum in which they are not mutually exclusive, and these purposes were associated with faculty evaluation for accountability, improvement, and learning, respectively (see Figure 1, Chapter 2). My study uses PBM and a qualitative case study to examine the SET and voting evaluation in some public HEIs in Vietnam, especially concerning how the implementing agents used evaluation for accountability, learning, and improvement. The PBM framework allowed me to understand the problem of the practices through exploring the participants' approaches, the constraints, and the consequences of these patterns (see Chapter 3). Key findings of my study indicate that faculty evaluation in the Vietnamese public HEI contexts was mainly used for managerial accountability purposes. While there were some efforts to use evaluation for improvement and learning, such efforts were inadequate to cause significant transformation of practice (see Chapters 4 and 5).

Overall, the dominance of accountability-oriented faculty evaluation approaches over improvement- and learning-oriented ones mostly caused single-loop learning rather than double-loop learning or an institutional change in the culture of evaluation for learning and improvement (Argyris & Schön, 1996). My constraint analysis suggests that the participants' approaches to faculty evaluation were not completely passive but were institutional and individual use of evaluation to address their managerial accountability demands (Røiseland et al., 2015). The participants' current approaches to problem solving, especially regarding underperformance issues, were strongly influenced by their collectivist Confucian harmony-oriented norm (M.-P. Nguyen, 2008; Truong et al., 2017), causing a limited impact on faculty members' learning and improvement. My study suggests further investigation into spiritual and philosophical constraints on Vietnamese academics' behaviours as Buddhism potentially influenced my Vietnamese participants' approaches to faculty evaluation. My study proposes improving the current theories of action about faculty evaluation by revising and adding some cultural and individual constraints to the existing constraints that influenced the participants' approaches (see Chapter 6). This chapter ends the thesis by presenting the key contributions of my study. It then outlines some limitations of the study, followed by suggestions for future research.

Contributions of the Study

The overall contribution of my study originates in the adoption of PBM (Robinson, 1993; Robinson & Lai, 2006) that investigated the problem of practices from the participants' theories of action in a non-Western higher education context like Vietnam. Under the PBM approach, my study contributes methodologically, theoretically, and practically to the existing literature in higher education faculty evaluation in several ways.

Methodological Contributions

My study has several methodological contributions to the existing literature. The first methodological contribution is about PBM and how it can be applied. Most PBM research examines educational issues in Western, or developed, countries such as New Zealand or Australia (Donald, 2013; Eastham, 2017; Finnerty, 2020; Hannah et al., 2019; Lalwani, 2019; Meyer & Slater-Brown, 2020; Nock, 2017; Slater-Brown, 2016). A few PBM studies in other contexts have explored the copying habits of Chinese students (Robinson & Lai, 1999), the curriculum design in Cambodian universities (Svay, 2017), and the Japanese leaders solving student incidents (Hannah et al., 2019). Also, except for studies by Svay (2017) and Robinson and Lai (1999), most previous PBM studies have been in general education (i.e., primary and secondary schools). My study demonstrates that PBM can be well applied to study education in Asian contexts like Vietnam and in higher education. This study creates an opportunity for PBM to be undertaken in a less developed Asian country like Vietnam. It adds to the existing PBM-related literature by suggesting how the methodology can be adapted to suit the Vietnamese higher education context.

Overall, my study adds to the literature by suggesting some strategies to address these challenges of conducting PBM research in Vietnam. For example, under the influence of Vietnamese communism and Confucianism, Vietnamese participants tend to suppress their personal voices and feelings, making it more difficult to carry out a true learning conversation, as defined by Robinson & Lai (2006). In a formal interview context, the Vietnamese participants might provide some diplomatic or safe answers, only focusing on their positive experiences of the faculty evaluation. Thus, researchers need to nourish their sensitivity by understanding and attending to their participants' needs. For example, I had an extended conversation with interviewees before each interview to get to know the participants personally and to unravel some of their participants' prior assumptions about the research topic. My study suggests one way to enable learning conversation interviews (Robinson & Lai, 2006) with the Vietnamese participants is to position oneself as an insider or occasionally invite the participants to react to viewpoints or scenarios different from theirs. By actively prompting the Vietnamese participants with critical questions, such as "what if" or "how do you react to ...?",

I could engage them in learning conversations about the practices they usually took for granted.

The second methodological contribution to research into faculty evaluation in Vietnam is my explicit connection between faculty evaluation policymaking and implementation. In the context of Vietnamese higher education, several studies (K. D. Nguyen, 2000; T. H. Nguyen, 2016; H. T. Pham & Nguyen, 2020; X. B. Tran, 2009) have pointed out some issues of the public faculty evaluation, such as lack of improvement focus or lack of strong material rewards. However, these studies did not explicitly link the faculty evaluation policymaking and implementing processes. Thus, it was unclear what the key faculty evaluation policies are and how they impact Vietnamese academics' approaches to the evaluation. My study complements previous research in faculty evaluation in Vietnamese higher education by analysing the national and institutional policy documents that influenced the implementing agents' approaches to faculty evaluation. I have also looked at the perspectives of participants in multiple roles, which makes a theoretical contribution by taking into account the complexity of the multiple roles involved in the practices. My study emphasises the need to examine policymaking and implementation as part of the methodology to study faculty evaluation.

Theoretical Contributions

My study contributes insights to the current theoretical debates about higher education faculty evaluation in various ways. First, my study findings in the Vietnamese context were juxtaposed with those from Western contexts considering the accountability, improvement, and learning orientations. Using these orientation lenses, I found that the issue of neoliberal influence or overemphasis on managerial focus is not limited to “the West” but also to “the rest” (Kostyrykina et al., 2018, p. 59), including an Asian country like Vietnam. My study looked beyond SET – a global neoliberal practice that combines with the voting evaluation – a unique practice specific to the Vietnamese communist ideology. Most studies on evaluation and continuous quality improvement of higher education were from developed countries like Australia, the United Kingdom, and the United States (Nasim et al., 2020), but not from non-English-speaking and less developed countries like Vietnam. Thus, my study provides insight into faculty evaluation that is not well understood and beyond the West, expanding the literature on faculty evaluation in higher education in several ways.

First, my study finds that both representatives of the neoliberal (i.e., SET) and communist ideologies (i.e., voting evaluation) were accountability oriented. My study confirms through the case of faculty evaluation in Vietnamese public HEIs that both globally market-based and communist centralised ideologies drive Vietnamese higher education faculty evaluation practices towards managerial rather than professional accountability (Røiseland et al., 2015). My findings confirm neoliberalism as a powerful constraint that drives higher

education practices not only in Western contexts but also in a communist Confucian country like Vietnam. By investigating faculty evaluation in the case of Vietnam, one of the centre-peripheral higher education systems in the world (Lo, 2011; Stack, 2021), my study adds to the plurality of the global knowledge system that "recognises and respects the fuller corpus of languages, theories, concepts and methods" (Marginson & Xu, 2021, p. 31). Notably, focusing on practitioners' theories of action rather than on prior educational theories, PBM provided me with the freedom to engage with reality from the practitioners' perspectives (Robinson, 1993; Robinson & Lai, 2006). This methodology enabled me to see the problems of faculty evaluation practices influenced by the Vietnamese participants' value system. My study is, therefore, another step towards "the globalisation of science and social science" that is "a move from cultural homogeneity centre on the old-world order to something like a unity-in-diversity approach" (Marginson & Xu, 2021, p. 30). My study finds that the faculty evaluation practices, like other higher education practices in Vietnam, are influenced by the country's dualism with combined market orientation (i.e., neoliberalism) and state-centralism (i.e., communism) (H. Tran, 2009). This challenges the Vietnamese policymakers' tendency to unquestioningly see the West or the Anglo-American knowledge economy paradigm as desirable (Kostyrykina et al., 2018). Thus, many have been eager to follow it, without questioning the applicability and relevance to the Vietnamese Confucian and communist-embedded context. My study recommends the addition of individual and cultural constraints in the policymaking and implementation of faculty evaluation in Vietnamese higher education. My study contributes to the literature by first discovering nuances in Confucian harmony orientation between Chinese and Vietnamese participants. The study also emphasises spiritual values or philosophical backgrounds as potential influences explaining the participants' approaches to faculty evaluation in Vietnamese higher education.

Second, my study extends the learning-oriented evaluation concept (Dahler-Larsen, 2009) by shifting its focus from programme evaluation to faculty evaluation. My study also contributes to the extant literature by expanding the evaluation purposes and by looking at these purposes on a continuum of summative, formative, and transformational purposes corresponding to three orientations: accountability, improvement, and learning (see Figure 1). Furthermore, my study is not limited to espoused purposes, but it draws attention to evaluation orientations which are purposes evident in practice. This specific attention paves the way for future studies focusing on evident-in-practice purposes (i.e., evaluation orientations). In line with the evaluation for learning suggested by Dahler-Larsen (2009), this study outlines the key conditions for successful faculty evaluation for learning approaches. One of the foremost conditions of this approach is to promote the ownership and capacity of administrators and faculty members, the end-users of the evaluation. My investigation supports Patton (2018), who advocated a collaborative and empowering evaluation approach. These alternative

approaches hopefully resolve prevalent problems of the current accountability and improvement-oriented evaluation approaches. The learning-oriented approach aims to engage the implementing agents in scientific inquiry; thus, it increases the voices of those who are often unheard in evaluation science (Patton, 2018).

Practical Contributions

Based on the participants' current theories of action, my study discusses the degree of potential improvability and suggests some conditions for improving approaches to faculty evaluation. Unlike most literature, my study focuses on describing the practice and seeking places for improvability as opposed to evaluating the practice. As discussed earlier, improving the existing faculty evaluation practices in Vietnamese higher education requires additional constraints and revising certain ones in the existing theories. Overall, my study provides some in-depth insights for higher education policymakers and practitioners to consider concerning their use of faculty evaluation for learning and improvement purposes.

First, my study provides policymakers with informed data on appropriate decisions to improve the quality of faculty evaluation. My study suggests three key strategies to improve faculty evaluation problems: improving the quality of evaluation data, reconceptualising SET and voting evaluation, and treating underperformance as a collective problem. My study suggests conditions for an improvement and learning-oriented evaluation approach include policymakers prioritising transformational purposes and the implementers committing to collaboration, dialogues, and growth among individuals and groups. My study findings will benefit higher education administrators, QA staff, leaders or policymakers who find themselves getting stuck in role conflicts between various demands. My study findings offer these evaluation-related stakeholders some alternative perspectives to understand and analyse their governance situations.

Second, my study highlights the practices from the perspectives of faculty members whose voice is essential but often missed during the higher education policymaking process (Kosrow, 2020; Messier, 2017; T. T. Tran, 2018). My study also provides holistic views about evaluation practitioners (especially QA officers, administrators, and faculty members). For example, teaching faculty members working in Vietnam or elsewhere can figure out what is happening to them concerning evaluation practices that matter to them. Faculty members will find their voices heard or their values articulated in my thesis. The necessary inclusion of faculty voices draws policymakers' attention to the expected impact of faculty evaluation on faculty professional learning and development.

Third, my study suggests a framework upon which higher education policymakers and practitioners can predict the likely success of future interventions on faculty evaluation. First, as external rewards were important but inadequate to motivate faculty members, it is unlikely

that increasing rewards will necessarily lead to increased use of evaluation for learning and improvement. What the participants found crucial was a fair, transparent, and convincing process that recognises their commitment to quality improvement of teaching, research, and service. Thus, a more practical strategy for improving the faculty evaluation system is to upgrade the evaluation criteria, processes, and data quality from faculty members' perspectives. Second, at the institutional level, the approaches taken by the QA officers and administrators were strongly constrained by the Vietnamese multilayered policymakers who expected a high degree of conformity from these participants. Thus, policymakers' emphasis on learning and improvement orientation will not be likely to work without the HEIs' autonomy and decision-making powers. In other words, from the HEI leaders' perspectives, the use of evaluation for learning and improvement would be unfeasible if it means ruling out their managerial and practical use of faculty evaluation.

Limitations of the Study

My study has two main limitations. A key limitation of my study is its focus on a case study limited to the Vietnamese public HEIs. Although its findings might not be generalisable, the focus on key stakeholders and the use of PBM to solve higher education problems of practice can be applied to other contexts. The applicability of research findings depends on the reader, not the researcher (Merriam, 2009), so the understanding from this specific case could be transferred to that of a similar situation. Therefore, the potential of this research finding cannot simply be ignored due to its limitations in formal generalisability, which is often overemphasised in general research (Merriam, 2009).

Due to the COVID-19 pandemic, I could not engage my participants in focus-group interviews nor follow-up "debate about the relative adequacy" (Robinson & Lai, 2006, p. 63) of my constructed theories of action. However, there are two features of my study that minimised this limitation. First, my unofficial data analysis started concurrently during my interviews with the participants. During the interviews, I occasionally disclosed my assumptions and preliminary understanding of the participants' theories of action to check if they were correctly understood. I also checked claims made by one group of participants by juxtaposing them with those of other groups. The early disclosure of my thoughts regarding the participants' theory of action and my triangulation of ideas across various groups allowed me to provide a robust explanation for their accounts of practice.

Suggestions for Future Research

My study suggests several issues that future research could focus on. First, future researchers could use PBM to explore the problem of faculty evaluation practices across a broader range of contexts. Second, I suggest that my proposed model of evaluation purposes

and orientations (see Figure 1) be adapted and applied in investigating faculty evaluation in other higher education contexts. My literature review generally reveals a shared problem of limited use of evaluation for learning and improvement in various contexts. Thus, my PBM research can be applied beyond Vietnam to other higher education contexts. Also, due to my inability to engage the previous participants in evaluating my constructed theories of action as I had planned, I would recommend future research to include the participants in a collaborative possesses of participants' feedback. Thus, some potential enablers for learning- and improvement-focused evaluation need further investigation. Although I did not explicitly probe the participants about their spiritual and philosophical backgrounds, my study suggests these are potentially important factors that influenced their faculty evaluation approaches. Thus, I recommend future research to examine how the Vietnamese participants' spiritual or philosophical worldview, such as Buddhism, influences Vietnamese participants' higher education practices. As Zen Master Thich Nhat Hanh (1967/2022) explained:

Most people do not seem to get very far in their understanding of Vietnamese Buddhism, and consequently, they cannot comprehend the Vietnamese problem... Objective conditions in Vietnamese society have compelled the Buddhist religion to engage itself in the life of the nation. (pp. 10–11)

Hence, I believe that future studies, which are based on Vietnamese religious contexts, individual spiritual values, or philosophical backgrounds, will contribute significantly to resolving the problems of higher education practices in Vietnam – a country with a unique value system.

References

- Agent, R. (2007). *Vietnam: The conflict in Indochina 1945–1975*. Pearson Longman.
- Algozzine, B., & Hancock, D. (2016). *Doing case study research: A practical guide for beginning researchers*. Teachers College Press.
- Alvesson, M., & Spicer, A. (2016). (Un)Conditional surrender? Why do professionals willingly comply with managerialism. *Journal of Organizational Change Management*, 29(1), 29–45. <https://doi.org/10.1108/JOCM-11-2015-0221>
- Anderson, L. (1994). *Espoused theories and theories-in-use: Bridging the gap: breaking through defensive routines with organisation development consultants*. [Master's thesis, University of Queensland].
http://www.aral.com.au/resources/LianeAnderson_thesis.pdf
- Anderson, L. (1997). *Argyris and Schon's theory on congruence and learning*.
<http://www.aral.com.au/resources/argyris.html>
- Argyris, Smith, M., & Putnam. (1985). *Action Science: Concepts, Methods and Skills for Research and Intervention*. Jossey-Bass Publishers.
- Argyris, C., & Schön, D. (1978). *Organizational learning: A theory of action perspective*. Addison-Wesley Publishing Company.
- Argyris, C., & Schön, D. A. (1974). *Theory in practice: Increasing professional effectiveness*. Jossey-Bass.
- Argyris, C., & Schön, D. A. (1996). *Organizational learning II: Theory, method, and practice* (Vol. 1). Addison-Wesley.
- Arreola, R. A. (2007). *Developing a comprehensive faculty evaluation system: A guide to designing, building, and operating large-scale faculty evaluation systems* (3rd ed.). Anker PubCo.
- ASEAN University Network. (2020). *ASEAN University Network quality assurance: Guide to AUN-QA assessment at programme level (Version 4.0)*. Author.

- Asher, N. "Engaging difference: Towards a pedagogy of interbeing." *Teaching Education* 14, no. 3 (2003): 235-247. <https://doi.org/10.1080/1047621032000135159>
- Atta-Owusu, K., & Fitjar, R. D. (2021). What motivates academics for external engagement? Exploring the effects of motivational drivers and organizational fairness. *Science and Public Policy*, 2021, scab075, <https://doi.org/10.1093/scipol/scab075>
- Babad, E., Sahar-Inbar, L., Hammer, R., Turgeman-Lupo, K., & Nessim, S. (2021). Student evaluations fast and slow: It's time to integrate teachers' nonverbal behavior in evaluations of teaching effectiveness. *Journal of Nonverbal Behavior*, 45(3), 321–338. <https://doi.org/10.1007/s10919-021-00364-4>
- Bach V. H., & Nguyen K. D. (2009). Đề nghị các chuẩn đánh giá giáo viên trong giai đoạn mới [Proposed standards for evaluating teachers in the new phase]. *Tạp chí Khoa học - Trường Đại học Sư phạm Tp. Hồ Chí Minh*, 17, 140–140. <https://vjol.info.vn/index.php/sphcm/article/view/14973>
- Barrow, M., & Grant, B. M. (2016). Changing mechanisms of governmentality? Academic development in New Zealand and student evaluations of teaching. *Higher Education*, 72(5), 589–601.
- Beerens, M. (2018). Evidence-based policy and higher education quality assurance: Progress, pitfalls and promise. *European Journal of Higher Education*, 8(3), 272–287.
- Bendermacher, G. W., Oude Egbrink, M. G., Wolfhagen, I. H., & Dolmans, D. H. (2017). Unravelling quality culture in higher education: A realist review. *Higher Education*, 73(1), 39–60. <https://doi.org/10.1007/s10734-015-9979-2>
- Berezvai, Z., Lukáts, G. D., & Molontay, R. (2021). Can professors buy better evaluation with lenient grading? The effect of grade inflation on student evaluation of teaching. *Assessment & Evaluation in Higher Education*, 46(5), 793–808. <https://doi.org/10.1080/02602938.2020.1821866>
- Berk, R. A. (2018). Beyond Student Ratings: Fourteen Other Sources of Evidence to Evaluate Teaching. In Ellis R. & Hogard E. (Eds.), *Handbook of Quality Assurance for University Teaching*. Routledge.

- Blanco, G., & Metcalfe, A. (2020). Visualizing quality: University online identities as organizational performativity in higher education. *The Review of Higher Education*, 43, 781–809. <https://doi.org/10.1353/rhe.2020.0007>
- Bogt, H. J., & Scapens, R. W. (2012). Performance management in universities: Effects of the transition to more quantitative measurement systems. *European Accounting Review*, 21(3), 451–497. <https://doi.org/10.1080/09638180.2012.668323>
- Borch, I. H. (2021). *Student evaluation practice: A qualitative study on how student evaluation of teaching, courses and programmes are carried out and used*. [Doctoral thesis, The Arctic University of Norway].
- Boyer, E. L. (1991). The scholarship of teaching from: Scholarship reconsidered: Priorities of the professoriate. *College Teaching*, 39(1), 11–13. <https://doi.org/10.1080/87567555.1991.10532213>
- Brannick, T., & Coghlan, D. (2007). In defense of being “native”: The case for insider academic research. *Organizational Research Methods*, 10(1), 59–74.
- Bui, B., Hoang, H., Phan, D. P., & Yapa, P. W. (2017). Governance and compliance in accounting education in Vietnam – Case of a public university. *Accounting Education*, 26(3), 265–290. <https://doi.org/10.1080/09639284.2017.1286603>
- Buller, J. L. (2012). *Best practices in faculty evaluation: A practical guide for academic leaders*. John Wiley & Sons. <https://www.wiley.com/en-au/Best+Practices+in+Faculty+Evaluation%3A+A+Practical+Guide+for+Academic+Leaders-p-9781118237885>
- Cadez, S., Dimovski, V., & Zaman Groff, M. (2017). Research, teaching and performance evaluation in academia: The salience of quality. *Studies in Higher Education*, 42(8), 1455–1473. <https://doi.org/10.1080/03075079.2015.1104659>
- Cardno, C. E., & Piggott-Irvine, E. (1997). *Effective performance appraisal: Integrating accountability and development in staff appraisal*. Longman.
- Carson, J. G., & Nelson, G. L. (1996). Chinese students’ perceptions of ESL peer response group interaction. *Journal of Second Language Writing*, 5(1), 1–19.

- Cashin, W. E. (2003). Evaluating college and university teaching: Reflections of a practitioner. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (pp. 531–593). Springer.
- Cenamor, J. (2021). To teach or not to teach? Junior academics and the teaching-research relationship. *Higher Education Research & Development*, 1–19.
<https://doi.org/10.1080/07294360.2021.1933395>
- Chan, K. (2001). The difficulties and dilemma of constructing a model for teacher evaluation in higher education. *Higher Education Management*, 13(1).
- Chan, W. V., Tang, H. H., & Cheung, R. L. (2020). Freedom to excel: Performativity, accountability, and educational sovereignty in Hong Kong's academic capitalism. In Z. Hao & P. Zabielskis (Eds.), *Academic freedom under siege: Higher education in East Asia, the U.S. and Australia* (pp. 125–145). Springer International Publishing.
https://doi.org/10.1007/978-3-030-49119-2_6
- Chen, G.-M. (2011). *An Introduction to Key Concepts in Understanding the Chinese: Harmony as the Foundation of Chinese Communication*, 7(4), 13.
https://digitalcommons.uri.edu/cgi/viewcontent.cgi?article=1011&context=com_facpubs
- Cheong, G. (2017). Developing an instrument for self-evaluation of teaching and learning competencies: A review of faculty professional development and the changing higher education landscape in Singapore. *Education Research and Perspectives (Online)*, 44, 1–20. <https://www.proquest.com/scholarly-journals/developing-instrument-self-evaluation-teaching/docview/2012855372/se-2?accountid=8424>
- Cherry, B. D., Grasse, N., Kapla, D., & Hamel, B. (2017). Analysis of academic administrators' attitudes: Annual evaluations and factors that improve teaching. *Journal of Higher Education Policy and Management*, 39(3), 296–306.
- Chu, I., & Vu, M. C. (2021). The nature of the self, self-regulation and moral action: Implications from the Confucian relational self and Buddhist non-self. *Journal of Business Ethics*. <https://doi.org/10.1007/s10551-021-04826-z>

- Clayson, D. E. (2017). Student evaluation of teaching and matters of reliability. *Assessment & Evaluation in Higher Education*. <https://doi.org/10.1080/02602938.2017.1393495>
- Coburn, C. E. (2006). Framing the problem of reading instruction: Using frame analysis to uncover the microprocesses of policy implementation. *American Educational Research Journal*, 43(3), 343–349.
<https://psycnet.apa.org/doi/10.3102/00028312043003343>
- Cormier, G. (2018). The language variable in educational research: An exploration of researcher positionality, translation, and interpretation. *International Journal of Research & Method in Education*, 41(3), 328–341.
- Cranton, P. (2001). Interpretive and critical evaluation. *New Directions for Teaching and Learning*, 2001(88), 11–18. <https://doi.org/10.1002/tl.33>
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). SAGE Publications.
- Curby, T., McKnight, P., Alexander, L., & Erchov, S. (2020). Sources of variance in end-of-course student evaluations. *Assessment & Evaluation in Higher Education*, 45(1), 44–53. <https://doi.org/10.1080/02602938.2019.1607249>
- Dahler-Larsen, P. (2009). Learning-oriented educational evaluation in contemporary society. In K. E. Ryan & J. B. Cousins (Eds.), *The SAGE International handbook of educational evaluation* (pp. 307–322). Sage.
- Dao, V. K., & Hayden, M. (2019). Vietnam's Progress with Policies on University Governance, Management and Quality Assurance. In N. T. Nguyen & L. T. Tran (Eds.), *Reforming Vietnamese Higher Education: Global Forces and Local Demands* (pp. 37–50). Springer. https://doi.org/10.1007/978-981-13-8918-4_3
- Darojat, O., Nilson, M., & Kaufman, D. (2015). Perspectives on quality and quality assurance in learner support areas at three Southeast Asian open universities. *Distance Education*, 36(3), 383–399. <https://doi.org/10.1080/01587919.2015.1081734>
- Darwin, S. (2016). The emergence of student evaluation in higher education. In S. Darwin (Ed.), *Student evaluation in higher education: Reconceptualising the student voice*

- (pp. 1–11). Springer International Publishing. https://doi.org/10.1007/978-3-319-41893-3_1
- Darwin, S. (2021). From the local fringe to market centre: Analysing the transforming social function of student ratings in higher education. *Studies in Higher Education*, 46(9), 1978–1990. <https://doi.org/10.1080/03075079.2020.1712690>
- Dawson, M., DeFranco, A., & Draper, J. (2020). Are all student evaluation scores created equally? Identifying factors that hospitality administrators should consider when assessing student feedback. *Journal of Hospitality, Leisure, Sport & Tourism Education*, 26, 100246. <https://doi.org/10.1016/j.jhlste.2020.100246>
- Dewar, J., & Perkins, K. (2021). Facilitating Faculty Getting Started in SoTL: Reflections by Two Carnegie Scholars. *Journal of the Scholarship of Teaching and Learning*, 21(2), 107–120. <https://eric.ed.gov/?id=EJ1303356>
- Dick, R., & Dalmay, T. (1991). *Values in action: Applying the ideas of Argyris and Schon*. Interchange.
- Do, Q. T. (2019). Accreditation, ranking and classification in Vietnamese higher education: The localization of foreign-born QA models and methods. In N. T. Nguyen & L. T. Tran (Eds.), *Reforming Vietnamese higher education: Global forces and local demands* (pp. 133–148). Springer. https://doi.org/10.1007/978-981-13-8918-4_8
- Do, Q. T., Pham, H. T., & Nguyen, K. D. (2017). Quality assurance in the Vietnamese higher education: A top-down approach and compliance-driven QA. In M. Shah & Q. T. Do (Eds.), *The rise of quality assurance in Asian higher education* (pp. 191–207). Elsevier. <https://doi.org/10.1016/B978-0-08-100553-8.00005-7>
- Donald, R. M. (2013). *The impact of practitioners' theories of action on curriculum change* [Master's thesis, University of Auckland]. <https://researchspace.auckland.ac.nz/handle/2292/20343>
- Duong, T. T. X. (2016). Nâng cao năng lực của đội ngũ giảng viên đại học đáp ứng yêu cầu đào tạo nguồn nhân lực chất lượng cao [Improving the capacity of university lecturers to meet the requirements of training high-quality human resources]. *Tạp Chí*

Lý Luận Chính Trị, 11. <http://lyluanchinhtri.vn/home/index.php/dao-tao-boi-duong/item/2054-nang-cao-nang-luc-cua-doi-ngu-giang-vien-dai-hoc-dap-ung-yeu-cau-dao-tao-nguon-nhan-luc-chat-luong-cao.html>

Eastham, D. A. (2017). *Partnerships for learning: Teachers' efforts to engage whānau* [Master's thesis, University of Auckland].

<https://researchspace.auckland.ac.nz/handle/2292/36377>

Ehlers, U. D. (2009). Understanding quality culture. *Quality Assurance in Education*, 17(4), 343–363. <https://doi.org/10.1108/09684880910992322>

Eiszler, C. F. (2002). College students' evaluations of teaching and grade inflation. *Research in Higher Education*, 43(4), 483–501.

Felix, J. J. (2020). Commentary: Modernity and reflexivity in Vietnamese higher education—Situating the role of the ideological, capacity building, learning organisation, and policy reform. In P. Le Ha & D. Ba Ngoc (Eds.), *Higher Education in market-oriented socialist Vietnam: New players, discourses, and practices* (pp. 107–122). Springer International Publishing. https://doi.org/10.1007/978-3-030-46912-2_6

Field, L. (2015). Appraising academic appraisal in the new public management university. *Journal of Higher Education Policy and Management*, 37(2), 172–189. <https://doi.org/10.1080/1360080X.2014.991534>

Finnerty, A. (2020). *Collaboration in educational networks: A theory of action case study of collaboration within and across schools* [Master's thesis, University of Auckland]. <https://researchspace.auckland.ac.nz/handle/2292/54059>

Giovannetti, M. M. (2015). *Faculty and administrator perceptions of a faculty evaluation process and its ability to advance faculty enrichment* [Unpublished doctoral thesis]. Simon Fraser University.

Gonzales, L., & Núñez, A.-M. (2014). The ranking regime and the production of knowledge: Implications for academia. *Education Policy Analysis Archives*. <https://doi.org/10.14507/epaa.v22n31.2014>

- Grand University. (2010). *University guidelines on collecting undergraduate student evaluation*.
- Grand University. (2014). *University quality evaluation from feedback of stakeholders*.
- Grand University. (2015). *University decision on emulation and commendation*.
- Grant, B. (2017). On delivering the consumer-citizen: New pedagogies and their affective economies. In C. Shore & S. Wright (Eds.), *Death of the Public University*.
- Grant, B. (2019). The future is now: A thousand tiny universities. *Philosophy and Theory in Higher Education*, 1(3), 9–28. <https://doi.org/10.3726/ptihe.2019.03.02>
- Gu, J., & Levin, J. S. (2021). Tournament in academia: A comparative analysis of faculty evaluation systems in research universities in China and the USA. *Higher Education*, 81(5), 897–915. <https://doi.org/10.1007/s10734-020-00585-4>
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Sage Publications.
- Haig, B. (2013). Reflections on problem-based methodology. *A Companion to Research in Education*, 353–356. https://doi.org/10.1007/978-94-007-6809-3_47
- Hajdin, G., & Pažur, K. (2012). Differentiating between student evaluation of teacher and teaching effectiveness. *Journal of Information and Organizational Sciences*, 36(2), 123–134.
- Hallinger, P. (2010). Using faculty evaluation to improve teaching quality: A longitudinal case study of higher education in Southeast Asia. *Educational Assessment, Evaluation and Accountability*, 22(4), 253–274.
- Hammer, R., Peer, E., & Babad, E. (2018). Faculty attitudes about student evaluations and their relations to self-image as teacher. *Social Psychology of Education*, 21(3), 517–537. <https://doi.org/10.1007/s11218-018-9426-1>
- Hannah, D., Sinnema, C., & Robinson, V. (2019). Theory of action accounts of problem-solving: How a Japanese school communicates student incidents to parents. *Management in Education*, 33(2), 62–69.
- Hayden, M., & Chinh, L.-N. (2020). A review of the reform agenda for higher education in Vietnam. *Higher Education in Market-Oriented Socialist Vietnam*, 21–39.

- Hayden, M., & Le-Nguyen, D. C. (2020). A review of the reform agenda for higher education in Vietnam. In P. Le Ha & D. Ba Ngoc (Eds.), *Higher education in market-oriented socialist Vietnam: New players, discourses, and practices* (pp. 21–39). Springer International Publishing. https://doi.org/10.1007/978-3-030-46912-2_2
- Hazelkorn, E., Coates, H., & McCormick, A. C. (2018). Quality, performance and accountability: Emergent challenges in the global era. In E. Hazelkorn, H. Coates, & A. C. McCormick (Eds.), *Research handbook on quality, performance and accountability in higher education* (pp. 3–14). Edward Elgar Publishing.
- Henderson, C., Beach, A., & Finkelstein, N. (2011). Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature. *Journal of Research in Science Teaching*, 48(8), 952–984. <https://doi.org/10.1002/tea.20439>
- Herr, K., & Anderson, G. L. (2014). *The action research dissertation: A guide for students and faculty*. Sage Publications.
- Hightower, W. H., Jr. (2010). Perceptions of Virginia community college system faculty and administrators on the purposes for and composition of a comprehensive evaluation system for teaching faculty members [Doctoral dissertation, Old Dominion University]. *ProQuest Dissertations and Theses* (763007497). <https://www.proquest.com/dissertations-theses/perceptions-virginia-community-college-system/docview/763007497/se-2?accountid=8424>
- Hoang P. (2003). *Từ Điển Tiếng Việt (Vietnamese Dictionary)*. Da Nang Publishing House. <https://nhatbook.com/2017/09/06/tu-dien-tieng-viet/>
- Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software of the mind* (Vol. 3). McGraw-Hill.
- Homutová, L. (2018). *Vietnamese political power, authority and legitimacy in the context of emulation movements* [Doctoral dissertation, Charles University]. <https://dspace.cuni.cz/bitstream/handle/20.500.11956/103320/140066431.pdf?sequence=1>

- Hornstein, H. A. (2017). Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Education*, 4(1), 1304016. <https://doi.org/10.1080/2331186X.2017.1304016>
- Hoskins, J. A., & Ninh, T.-H. (2017). Globalizing Vietnamese religions. *Journal of Vietnamese Studies*, 12(2), 1–19. <https://doi.org/10.1525/jvs.2017.12.2.1>
- Huang, Y.-T. (2020). Responding to the neoliberal and managerial changes: A generational perspective of Chinese academics. *Compare: A Journal of Comparative and International Education*, 1–19. <https://doi.org/10.1080/03057925.2020.1716305>
- Hung, L. N. (2013). *A social cognitive and cross-cultural investigation of performance appraisal in Australian and Vietnamese universities*. [Doctoral thesis, University of Wollongong]. <https://ro.uow.edu.au/cgi/viewcontent.cgi?referer=https://scholar.google.com/&httpsredir=1&article=4944&context=theses>
- Hunt, L. H. (2008). *Grade inflation: Academic standards in higher education*. Suny Press.
- Hutaibat, K., Alhatabat, Z., Alberti-Alhtaybat, L., & Al-Htaybat, K. (2021). Performance habitus: Performance management and measurement in UK higher education. *Measuring Business Excellence*, 25(2), 171–188. <https://doi.org/10.1108/MBE-08-2019-0084>
- Huynh, T. T. (2016). *Vietnamese academics' research capacity in tertiary contexts*. <https://1library.net/document/yrmn538q-vietnamese-academics-research-capacity-in-tertiary-contexts.html>
- Jeffs, C., Nelson, N., Grant, K. A., Nowell, L., Paris, B., & Viceer, N. (2021). Feedback for teaching development: Moving from a fixed to growth mindset. *Professional Development in Education*, 0(0), 1–14. <https://doi.org/10.1080/19415257.2021.1876149>
- Jha, P. K. (2005). *Assessment and evaluation in higher education*. Vista International Publishing House.
- Jones, D. R., Visser, M., Stokes, P., Örtenblad, A., Deem, R., Rodgers, P., & Tarba, S. Y. (2020). The performative university: 'Targets,' 'terror' and 'taking back freedom' in

academia. *Management Learning*, 51(4), 363–377.

<https://doi.org/10.1177/1350507620927554>

Jungblut, J., Stensaker, B., & Vukasovic, M. (2018). What kind of quality assurance leads to improved performance and accountability? The views of European students. In E. Hazelkorn, H. Coates, & A. C. McCormick (Eds.), *Research handbook on quality, performance and accountability in higher education* (pp. 321–331). Edward Elgar Publishing.

Kallio, K.-M., & Kallio, T. J. (2014). Management-by-results and performance measurement in universities—Implications for work motivation. *Studies in Higher Education*, 39(4), 574–589.

Kenny, J. (2017). Academic work and performativity. *Higher Education*, 74(5), 897–913.

<https://doi.org/10.1007/s10734-016-0084-y>

Kenny, J. (2018). Re-empowering academics in a corporate culture: An exploration of workload and performativity in a university. *Higher Education*, 75(2), 365–380.

<https://doi.org/10.1007/s10734-017-0143-z>

Khtere, A. (2020). Performance Appraisal of Faculty Members Based on Internal Quality Assurance System: A Delphi Study. *International Journal of Higher Education*, 9(6), 87–95. <https://eric.ed.gov/?id=EJ1277879>

Ko, E. Y.-L. (2001). *Performance appraisal in higher education: The impact of new managerialism on the motivation of academic staff* [Unpublished doctoral thesis]. University of Bristol.

Koçe, T., Jansone-Ratinika, N., & Koka, R. (2017). Formative assessment on an individual and an institutional level on the way to quality culture in higher education. *Problems of Education in the 21st Century*, 75(1), 54–71.

<https://www.proquest.com/docview/2343796812/abstract/622D04BB80DE49A9PQ/1>

Kosrow, L. (2020). *“We Don’t Work For Free”: Tenured Faculty Engagement In Assessment Initiatives On A Unionized Community College Campus*. [Doctoral thesis, National Louis University].

- Kostrykina, S., Lee, K., & Hope, J. (2018). The west, the rest and the knowledge economy: A game worth playing? *Perspectives: Policy and Practice in Higher Education*, 22(2), 58–67. <https://doi.org/10.1080/13603108.2017.1372317>
- Lalwani, S. (2019). *Deciding to implement the International Baccalaureate (IB) primary years programme: A case study* [Master's thesis, Australian Catholic University]. <https://doi.org/10.26199/6m4w-1693>
- Lammers, W. J., Gillaspay Jr, J. A., & Hancock, F. (2017). Predicting academic success with early, middle, and late semester assessment of student–instructor rapport. *Teaching of Psychology*, 44(2), 145–149.
- Law on Emulation and Commendation 2003.
- Law on Emulation and Commendation Amendment 2013.
- Le, C. L., & Do, Đ. T. (2017). Sự thay đổi hoạt động giảng dạy môn học của giảng viên dưới góc nhìn phản hồi từ người học [Changes in teaching viewed from student feedback]. *VNU Journal of Science: Education Research*, 33(2), 23–32.
- Le, K. A., Hayden, M., & Nhan, T. T. (2019). Development of the higher education sector in Vietnam. In C. H. Nguyen & M. Shah (Eds.), *Quality assurance in Vietnamese higher education* (pp. 1–23). Springer.
- Le, T. M., & Yu, N. (2019). Ideological and Philosophical Underpinnings of Attitudes Toward Sexual Minorities in Vietnamese Society. *Sexuality & Culture*, 23(2), 444–457. <https://doi.org/10.1007/s12119-018-9577-4>
- Lee, H., & Lee, K. (2013). Publish (in international indexed journals) or perish: Neoliberal ideology in a Korean university. *Language Policy*, 12(3), 215–230. <https://doi.org/10.1007/s10993-012-9267-2>
- Lei, W., Li, J., Li, Y., Castaño, G., Yang, M., & Zou, B. (2021). The boundary conditions under which teaching–research conflict leads to university teachers' job burnout. *Studies in Higher Education*, 46(2), 406–422. <https://doi.org/10.1080/03075079.2020.1811218>

- Li, Y. (2011). Translating Interviews, Translating Lives: Ethical Considerations in Cross-Language Narrative Inquiry. *TESL Canada Journal*, 28(5), 16–30.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.
- Lo, W. Y. (2011). Soft power, university rankings and knowledge production: Distinctions between hegemony and self-determination in higher education. *Comparative Education*, 47(2), 209–222. <https://doi.org/10.1080/03050068.2011.554092>
- Lohman, L. (2021). Evaluation of university teaching as sound performance appraisal. *Studies in Educational Evaluation*, 70, 101008. <https://doi.org/10.1016/j.stueduc.2021.101008>
- London, J. D. (2010). Globalization and the governance of education in Viet Nam. *Asia Pacific Journal of Education*, 30(4), 361–379. <https://doi.org/10.1080/02188791.2010.520202>
- Lonsdale, A. (1998). Performance appraisal, performance management and quality in higher education: Contradictions, issues and guiding principles for the future. *Australian Journal of Education*, 42(3), 303–320. <https://doi.org/10.1177/000494419804200307>
- Lu, Y. (2014). *An exploration of merit pay, teacher and student satisfaction, and teacher performance evaluation from an instructional perspective* [Doctoral thesis, University of Missouri-Saint Louis]. <http://www.proquest.com/docview/1545690962/abstract/14287DA75E14354PQ/1>
- Lyde, A. R., Grieshaber, D. C., & Byrns, G. (2016). Faculty teaching performance: Perceptions of a multi-source method for evaluation. *Journal of the Scholarship of Teaching and Learning*, 16(3), 82–94. <https://doi.org/10.14434/josotl.v16i3.18145>
- Lyotard, J.-F. (1984). *The postmodern condition* (Vol. 1). Manchester University Press. <https://www.marxists.org/reference/subject/philosophy/works/fr/lyotard.htm> (Original work published 1979)
- Macfarlane, B. (2021). The neoliberal academic: Illustrating shifting academic norms in an age of hyper-performativity. *Educational Philosophy and Theory*, 53(5), 459–468. <https://doi.org/10.1080/00131857.2019.1684262>

- Madhan, M., Gunasekaran, S., & Arunachalam, S. (2018). Evaluation of research in India—Are we doing it right. *Indian J Med Ethics*, 3(3), 221–229.
- Mampane, S. T. (2020). Understanding and application of the institutional performance management system within higher education institution departments. In *BCES Conference Books – Volume 18 of 2020*. Bulgarian Comparative Education Society.
- Marginson, S. (2010). The global knowledge economy and the culture of comparison in higher education. In A. Kaur, M. Sirat, & W. G. Tierney (Eds.), *Quality assurance and university rankings in higher education in the Asia Pacific: Challenges for universities and nations* (pp. 23–55). National Higher Education Research Institute.
- Marginson, S. (2011). Higher education in East Asia and Singapore: Rise of the Confucian model. *Higher Education*, 61(5), 587–611.
- Marginson, S., & Xu, X. (2021). *Moving beyond centre-periphery science: Towards an ecology of knowledge*. The Centre for Global Higher Education.
<https://www.researchcghe.org/publications/working-paper/moving-beyond-centre-periphery-science-towards-an-ecology-of-knowledge/>
- Mark, M. M., Greene, J. C., & Shaw, I. F. (2006). Introduction: The evaluation of policies, programs, and practices. In I. Shaw, J. C. Greene, & M. M. Mark (Eds.), *The SAGE handbook of evaluation* (pp. 1–30). <https://methods.sagepub.com/book/the-sage-handbook-of-evaluation>
- McNess, E., Arthur, L., & Crossley, M. (2015). Ethnographic dazzle and the construction of the “other”: Shifting boundaries between the insider and the outsider. In *Revisiting insider-outsider research in comparative and international education*. Symposium Books.
- Medlen, M. J. (2019). *Faculty Perceptions of Course Evaluation Processes* [Doctoral dissertation, California Lutheran University].
- Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. Jossey-Bass.

- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation*. Jossey-Bass.
- Mertens, D. M. (2010). Social transformation and evaluation. *Evaluation Journal of Australasia*, 10(2), 3–10. <https://doi.org/10.1177/1035719X1001000202>
- Messier, J. (2017). Shared governance and academic freedom: Yes, this is union work. *Thought & Action*, 33(2), 63–76.
- Meyer, F., & Slater-Brown, K. (2020). Educational change doesn't come easy: Lead teachers' work as change agents. *Mathematics Education Research Journal*. <https://doi.org/10.1007/s13394-020-00333-y>
- Mills, M., & Hyle, A. E. (1999). Faculty evaluation: A prickly pair. *Higher Education*, 38(3), 351–371. <https://doi.org/10.1023/A:1003735227936>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.
- Ministry of Education and Training. (2010). *Công văn 2754/BGDĐT-NGCBQLGD hướng dẫn lấy ý kiến phản hồi từ người học về hoạt động giảng dạy của giảng viên* [Correspondence 2754/BGDĐT-NGCBQLGD on collecting student evaluation of teaching]. <https://thuvienphapluat.vn/cong-van/Giao-duc/Cong-van-2754-BGDDT-NGCBQLGD-huong-dan-lay-y-kien-phan-hoi-tu-nguoi-hoc-107682.aspx>
- Ministry of Education and Training. (2013). *Công văn 7324/BGDĐT-NGCBQLGD hướng dẫn tổ chức lấy ý kiến phản hồi từ người học về hoạt động giảng dạy của giảng viên* [Correspondence 7324/BGDĐT-NGCBQLGD on organising the collection of student evaluation of teaching]. <https://luatvietnam.vn/giao-duc/cong-van-7324-bgddt-ngcbqlgd-bo-giao-duc-va-dao-tao-81940-d6.html>
- Ministry of Education and Training. (2019). *Khảo thí và kiểm định chất lượng giáo dục* [Assessment and accreditation of higher education Institutions]. <https://moet.gov.vn/giaoducquocdan/khao-thi-va-kiem-dinh-chat-luong-giao-duc/Pages/Default.aspx>

- Ministry of Education and Training. (2020). *Vietnamese higher education statistics during 2018–2019*. <https://moet.gov.vn/thong-ke/Pages/thong-ko-giao-duc-dai-hoc.aspx?ItemID=6636>
- Ministry of Education and Training. (2021, August 24). *Giáo dục đại học năm học mới: Tăng cường thích nghi, tiếp tục lộ trình đổi mới* [Higher education in the new school year: Strengthening adaptation, continuing the innovation roadmap]. Ministry of Education and Training. <https://moet.gov.vn/tintuc/Pages/tin-tong-hop.aspx?ItemID=7489>
- Ministry of Home Affairs. (2020). *Đại hội Thi đua yêu nước ngành Giáo dục: Đổi mới sáng tạo trong dạy và học* [The National Emulation Congress in Education: Innovating in teaching and learning]. Nguyễn Thanh Bình. <https://moha.gov.vn/baucu/van-ban-huong-dan/dai-hoi-thi-dua-yeu-nuoc-nganh-giao-duc-doi-moi-sang-tao-trong-day-va-hoc-44996.html>
- Morris, L., Stanton, P., & Young, S. (2007). Performance management in higher education—Development versus control. *New Zealand Journal of Employment Relations (Online)*, 32(2), 17–31. <https://www.proquest.com/openview/421297336e3a9037debe2b4ed0c8d98d/1?pq-origsite=gscholar&cbl=39964>
- Mountz, A., Bonds, A., Mansfield, B., Loyd, J., Hyndman, J., Walton-Roberts, M., Basu, R., Whitson, R., Hawkins, R., & Hamilton, T. (2015). For slow scholarship: A feminist politics of resistance through collective action in the neoliberal university. *ACME: An International Journal for Critical Geographies*, 14(4), 1235–1259.
- Muliaina, T. (2018). In search of meaningful assessment in the university curriculum: The case for culturally relevant pedagogy. *Australian Geographer*, 49(4), 517–535. <https://doi.org/10.1080/00049182.2018.1440689>
- Musselin, C. (2013). Redefinition of the relationships between academics and their university. *Higher Education*, 65(1), 25–37.
- Mutch, C., & Tatebe, J. (2017). From collusion to collective compassion: Putting heart back into the neoliberal university. *Pastoral Care in Education*, 35(3), 221–234. <https://doi.org/10.1080/02643944.2017.1363814>

- Nasim, K., Sikander, A., & Tian, X. (2020). Twenty years of research on total quality management in Higher Education: A systematic literature review. *Higher Education Quarterly*, 74(1), 75–97. <https://doi.org/10.1111/hequ.12227>
- Ngo, H. T. (2020). ‘Standing between the flows’: Interactions among neoliberalism, socialism and Confucianism in Vietnamese higher education. In P. Le Ha & D. Ba Ngoc (Eds.), *Higher education in market-oriented socialist Vietnam: New players, discourses, and practices* (pp. 41–61). Springer International Publishing. https://doi.org/10.1007/978-3-030-46912-2_3
- Ngo, T. L. (2020). Critiquing the promotion of American biased “liberal arts education” in post- “Đổi Mới” Vietnam. In P. Le Ha & D. Ba Ngoc (Eds.), *Higher education in market-oriented socialist Vietnam: New players, discourses, and practices* (pp. 125–144). Springer International Publishing. https://doi.org/10.1007/978-3-030-46912-2_7
- Nguyen, C. H., & Shah, M. (Eds.). (2019). *Quality assurance in Vietnamese higher education: Policy and practice in the 21st century*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-26859-6>
- Nguyen, C. P. (2020). *Education in the reform era: How policy and politics transformed Vietnamese students, school leaders, and communities*. [Doctoral dissertation, The Pennsylvania State University]. https://etda.libraries.psu.edu/files/final_submissions/20977
- Nguyen, H. T. (2020). A review of university research development in Vietnam from 1986 to 2019. In P. Le Ha & D. Ba Ngoc (Eds.), *Higher education in market-oriented socialist Vietnam: New players, discourses, and practices* (pp. 63–86). Springer International Publishing. https://doi.org/10.1007/978-3-030-46912-2_4
- Nguyen, K. D. (2000). *The potential of appraisal policies to improve the quality of teaching in Ho Chi Minh City College of Education, Vietnam* [Master’s thesis, The University of Melbourne]. <http://minerva-access.unimelb.edu.au/handle/11343/37419>

- Nguyen, K. D. (2003). *International practices in quality assurance for higher education teaching and learning: Prospects and possibilities for Vietnam* [Doctoral thesis, The University of Melbourne]. <http://minerva-access.unimelb.edu.au/handle/11343/37419>
- Nguyen, K. D., & Mcinnis, C. (2002). The possibility of using student evaluations in Vietnamese higher education. *Quality in Higher Education*, 8(2), 151–158.
- Nguyen, K. T., Griffin, P., & Nguyen, C. (2006). *Generating criteria for assessing lecturers in Vietnam's universities: A conceptual paper*. Australian Association for Research in Education (AARE) International Education Research Conference, Adelaide, SA.
- Nguyen, L. U. (2014). Creating core conditions for quality teaching evaluation in Vietnamese higher education. In A. Kwan, E. Wong, T. Kwong, P. Lau, & A. Goody (Eds.), *Research and development in higher education: Higher education in a globalized world* (Vol. 37, pp. 245–254). Higher Education Research and Development Society of Australasia.
- Nguyen, N. T. (2019). Cultural modalities of Vietnamese higher education. In N. T. Nguyen & L. T. Tran (Eds.), *Reforming Vietnamese higher education: Global forces and local demands* (pp. 17–33). Springer. https://doi.org/10.1007/978-981-13-8918-4_2
- Nguyen, P.-M. (2008). *Culture and cooperation: Cooperative learning in Asian Confucian heritage cultures. The case of Viet Nam*. [Doctoral thesis, Utrecht University]. <https://dspace.library.uu.nl/bitstream/handle/1874/27818/nguyen.pdf?sequence=2>
- Nguyen, P. V. (2020). TT-SET augmented with POT: A potential framework for academics' pedagogical reasoning. *Reflective Practice*, 21(1), 14–27.
- Nguyen, T. H. (2016). Building human resources management capacity for university research: The case at four leading Vietnamese universities. *Higher Education*, 71(2), 231–251.
- Nguyen, T. Q. (2015). Conducting semi-structured interviews with the Vietnamese. *Qualitative Research Journal*, 15(1), 35–46. <https://doi.org/10.1108/QRJ-04-2014-0012>

- Nguyen V. K. (2015). *Chỉ số thực hiện nhiệm vụ của giảng viên: So sánh tự đánh giá và đánh giá của phía cán bộ phụ trách [Performance indicators of faculty evaluation: Comparison of self-evaluation and administrator evaluation (Case study of Vietnam National University, Hanoi)]* [Master's Thesis, Vietnam National University Hanoi]. http://repository.vnu.edu.vn/handle/VNU_123/8717
- Nickles, T. (1981). What is a problem that we may solve it? *Synthese*, 47(1), 85–118. <http://www.jstor.org/stable/20115620>
- Niessen, T. J., Abma, T. A., Widdershoven, G. A., & Van der Vleuten, C. P. (2009). Learning-in-(inter) action: A dialogical turn to evaluation and learning. *The SAGE International Handbook of Educational Evaluation*. Thousand Oaks, CA: Sage, 377–395.
- Nock, M. (2017). *Teacher perceptions and practice for increasing student literacy engagement in low decile 1:1 laptop classrooms* [Master's thesis, University of Auckland]. <https://researchspace.auckland.ac.nz/handle/2292/36927>
- Norsworthy, D. B., & Sanders, D. M. (2021). *Working with “the SET process” to improve teachers’ teaching and learners’ learning*. Ako Aotearoa. <https://ako.ac.nz/knowledge-centre/working-with-the-set-process/>
- O’Leary, M., & Savage, S. (2020). Breathing new life into the observation of teaching and learning in higher education: Moving from the performative to the informative. *Professional Development in Education*, 46(1), 145–159. <https://doi.org/10.1080/19415257.2019.1633386>
- Oon, P.-T., Spencer, B., & Kam, C. C. (2017). Psychometric quality of a student evaluation of teaching survey in higher education. *Assessment & Evaluation in Higher Education*, 42(5), 788–800. <https://doi.org/10.1080/02602938.2016.1193119>
- Ory, J. C. (2000). Teaching evaluation: Past, present, and future. *New Directions for Teaching and Learning*, 83, 13–18.
- Palermo, J. (2013). Linking student evaluations to institutional goals: A change story. *Assessment & Evaluation in Higher Education*, 38(2), 211–223.

- Patton, M. Q. (1996). A world larger than formative and summative. *Evaluation Practice*, 17(2), 131–144.
- Patton, M. Q. (2018). Evaluation science. *The American Journal of Evaluation*, 39(2), 183–200. <https://doi.org/10.1177/1098214018763121>
- Patuawa, J., Robinson, V., Sinnema, C., & Zhu, T. (2021). Addressing inequity and underachievement: Middle leaders' effectiveness in problem solving. *Leading and Managing*, 27(1), 51–78.
- Patuawa, J., Robinson, V., Sinnema, C., & Zhu, T. (2021). Addressing inequity and underachievement: Middle leaders' effectiveness in problem solving. *Leading and Managing*, 27(1), 51–78.
- Pham H. T., & Nguyen A. D. (2019). *Trách nhiệm giải trình của cơ quan hành chính Nhà nước ở Việt Nam hiện nay [Accountability of state administrative agencies in Vietnam today]*. http://repository.vnu.edu.vn/handle/VNU_123/99343
- Pham, H. T. (2014). *Quality culture in Vietnamese universities: A multiple case study of quality assurance and quality culture of business English undergraduate programmes at three universities in Vietnam* [Doctoral thesis, Victoria University of Wellington].
- Pham, H. T. (2018). Assuring quality in higher education in a Confucian collectivist culture: The Vietnamese experience. *Journal of Contemporary Educational Research*, 2(4), 10–19.
- Pham, H. T. (2019). Limited legitimacy among academics of centrally driven approaches to internal quality assurance in Vietnam. *Journal of Higher Education Policy and Management*, 41(2), 172–185.
- Pham, H. T., & Nguyen, C. H. (2020). Academic staff quality and the role of quality assurance mechanisms: The Vietnamese case. *Quality in Higher Education*, 26(3), 262–283.
- Pham, H. T., & Starkey, L. (2016). Perceptions of higher education quality at three universities in Vietnam. *Quality Assurance in Education*, 24(3), 369–393. <https://doi.org/10.1108/QAE-07-2014-0037>

- Pham, T. L. (2012). The renovation of higher education governance in Vietnam and its impact on the teaching quality at universities. *Tertiary Education & Management*, 18(4), 289–308. <https://doi.org/10.1080/13583883.2012.675350>
- Pham, T. N., & Goyette, K. (2019). Higher Education Governance in Vietnam: Statism Versus Institutional Autonomy. In D. S. L. Jarvis & K. H. Mok (Eds.), *Transformations in Higher Education Governance in Asia: Policy, Politics and Progress* (pp. 197–212). Springer. https://doi.org/10.1007/978-981-13-9294-8_10
- Pham, T. T., & Renshaw, P. (2013). How to enable Asian teachers to empower students to adopt student-centred learning. *Australian Journal of Teacher Education*, 38(11), 65–65. <https://doi.org/10.14221/ajte.2013v38n11.4>
- Pham, T. V. (2016). *Công tác quản lý nhà nước về Thi đua, Khen thưởng tại các bệnh viện công lập trên địa bàn Hà Nội* [The state governance of Emulation and Commendation at public hospitals in Hanoi]. National Academy of Public Administration. <https://www1.napa.vn/saudaihoc/pham-thi-viet-anh-quan-ly-nha-nuoc-ve-thi-dua-khen-thuong-tai-benh-vien-tuyen-trung-uong-tren-dia-ban-thanh-pho-ha-noi.htm>
- Phan, A. N. (2021). Under the impacts of globalisation: The rising power of English as a foreign language (EFL) and the corresponding response of EFL policy in Vietnam. *SN Social Sciences*, 1(1), 31. <https://doi.org/10.1007/s43545-020-00047-9>
- Phan, L. H., & Dang, V. H. (2020). Engaging (with) new insights: Where to start to move scholarship and the current debate forward. In L. H. Phan & B. N. Doan (Eds.), *Higher education in market-oriented socialist Vietnam* (pp. 363–378). Springer. https://doi.org/10.1007/978-3-030-46912-2_20
- Phan, L. H., & Doan, B. N. (2020). Introduction and foregrounding the work: ‘New’ players, ‘new’ discourses, ‘new’ practices, and “new flavours.” In L. H. Phan & B. N. Doan (Eds.), *Higher education in market-oriented socialist Vietnam* (pp. 3–19). Springer International Publishing. https://doi.org/10.1007/978-3-030-46912-2_1

- Phan, N. (2014). Faculty perceptions of the performance appraisal process [Doctoral thesis, California State University, Fullerton]. *ProQuest Dissertations and Theses*.
<http://www.proquest.com/docview/1616727216/abstract/C253F4026E4C400CPQ/1>
- Posselt, J., Hernandez, T. E., Villarreal, C. D., Rodgers, A. J., & Irwin, L. N. (2020). Evaluation and decision making in higher education: Toward equitable repertoires of faculty practice. In L. W. Perna (Ed.), *Higher education: Handbook of theory and research* (Vol. 35, pp. 1–63). Springer International Publishing.
https://doi.org/10.1007/978-3-030-11743-6_8-1
- Postiglione, G., & Jung, J. (2017). *The changing academic profession in Hong Kong: Challenges and future* (pp. 3–14). https://doi.org/10.1007/978-3-319-56791-4_1
- Prasetya, I. (2020). *An evaluation of the national examination policy in the Indonesian education system* [Doctoral Thesis, University of Auckland].
- Ramirez, F. O. (2013). World society and the university as formal organization. *The European Educational Space: New Fabrications*, 1(1).
- Ray, B., Babb, J., & Wooten, C. A. (2018). Rethinking SETs: Retuning student evaluations of teaching for student agency. *Composition Studies*, 46(1), 34–56, 192–194.
<https://eric.ed.gov/?id=EJ1181078>
- Resolution 14/2005/NQ-CP on Fundamental and Comprehensive Reform of Higher Education in Vietnam in the Period 2006–2020. (2005).
<https://thuvienphapluat.vn/van-ban/Giao-duc/Nghi-quyet-14-2005-NQ-CP-doi-moi-co-ban-va-toan-dien-giao-duc-dai-hoc-Viet-Nam-giai-doan-2006-2020-5013.aspx>
- Robinson, V. (1992). Why doesn't educational research solve educational problems? *Educational Philosophy and Theory*, 24(2), 8–28.
- Robinson, V. (1993). *Problem-based methodology: Research for the improvement of practice*. Pergamon Press.
- Robinson, V. (2011). *Student-centered leadership* (Vol. 15). John Wiley & Sons.
- Robinson, V. (2018). *Reduce change to increase improvement*. Corwin.

- Robinson, V. M., & Lai, M. K. (1999). The explanation of practice: Why Chinese students copy assignments. *International Journal of Qualitative Studies in Education*, 12(2), 193–210.
- Robinson, V., & Lai, M. K. (2006). *Practitioner research for educators: A guide to improving classrooms and schools*. Corwin Press. <https://us.corwin.com/en-us/nam/practitioner-research-for-educators/book225422>
- Robinson, V., & Timperley, H. (2000). The link between accountability and improvement: The case of reporting to parents. *Peabody Journal of Education*, 75(4), 66–89.
- Røiseland, A., Pierre, J., & Gustavsen, A. (2015). Accountability by professionalism or managerialism? Exploring attitudes among Swedish and Norwegian local government leaders. *International Journal of Public Administration*, 38(10), 689–700. <https://doi.org/10.1080/01900692.2014.955196>
- Roman, O., Valentin, M., & Nina, S. (2018). Neoliberalism in higher education as a challenge for future civilization. *Философия и Космология*, 20.
- Roxå, T., Ahmad, A., Barrington, J., Van Maaren, J., & Cassidy, R. (2021). Reconceptualizing student ratings of teaching to support quality discourse on student learning: A systems perspective. *Higher Education*. <https://doi.org/10.1007/s10734-020-00615-1>
- Rutherford, D. (1992). Appraisal in action: A case study of innovation and leadership. *Studies in Higher Education*, 17(2), 201–210. <https://doi.org/10.1080/03075079212331382667>
- Ryan, K. E., & Cousins, J. B. (Eds.). (2009). *The SAGE international handbook of educational evaluation*. SAGE Publications.
- Salmi, J., & Pham, I. T. (2019). Academic governance and leadership in Vietnam: Trends and challenges. *Journal of International and Comparative Education (JICE)*, 8(2), 103–118. <https://doi.org/10.14425/jice.2019.8.2.103>

- Sampson, J. P., Driscoll, M. P., Foulk, D. F., & Carroll, P. S. (2010). Successful faculty performance in teaching, research and original creative work, and service. *Florida State University*.
- Sannadan, J. G., & Lang-ay, P. (2021). Neoliberal restructuring in Philippine education: Towards performativity. *International Journal of English Literature and Social Sciences*, 6, 355–366. <https://doi.org/10.22161/ijels.61.45>
- Saunders, M., Trowler, P., & Bamber, V. (2011). *Reconceptualising evaluation in higher education: The practice turn*. Open University Press; Society for Research into Higher Education, McGraw Hill.
- Schreiner, L. A., & Tobolowsky, B. F. (2018). The role of faculty in sophomore success. *New Directions for Higher Education*, 2018(183), 59–70. <https://doi.org/10.1002/he.20293>
- Scriven, M. (1991). Beyond formative and summative evaluation. *Faculty Members College Record*, 92(6), 18–64.
- Seldin, P. (1984). *Changing practices in faculty evaluation*. Jossey-Bass.
- Seldin, P. (2006). *Evaluating faculty performance: A practical guide to assessing teaching, research, and service*. Anker PubCompany.
- Serdyukova, N., Tatum, B. C., & Serdyukova, P. (2010). Student evaluations of courses and teachers. *Journal of Research in Innovative Teaching*, 3(1), 173–182.
- Shah, M., Cheng, M., & Fitzgerald, R. (2017). Closing the loop on student feedback: The case of Australian and Scottish universities. *Higher Education*, 74(1), 115–129. <https://doi.org/10.1007/s10734-016-0032-x>
- Shaked, H. (2021). Instructional leadership in higher education: The case of Israel. *Higher Education Quarterly*, 75(2), 212–226. <https://doi.org/10.1111/hequ.12274>
- Shao, L. P., Anderson, L. P., & Newsome, M. (2007). Evaluating teaching effectiveness: Where we are and where we should be. *Assessment & Evaluation in Higher Education*, 32(3), 355–371.
- Shavelson, R. J., Zlatkin-Troitschanskaia, O., & Mariño, J. P. (2018). Performance indicators of learning in higher education institutions: An overview of the field. In E. Hazelkorn,

- H. Coates, & A. C. McCormick (Eds.), *Research handbook on quality, performance and accountability in higher education* (pp. 249–263).
<https://doi.org/10.4337/9781785369759.00029>
- Shelley, S. (1999). Diversity of appraisal and performance-related pay practices in higher education. *Personnel Review*, 28(5/6), 439–454.
<https://doi.org/10.1108/00483489910286765>
- Shi, J., Luo, Y., Wen, W., & Guo, F. (2018). Governing quality in a transforming higher education system: The case of China. In E. Hazelkorn, H. Coates, & A. C. McCormick (Eds.), *Research handbook on quality, performance and accountability in higher education* (pp. 371–381). <https://doi.org/10.4337/9781785369759.00039>
- Shields, R., & Watermeyer, R. (2020). Competing institutional logics in universities in the United Kingdom: Schism in the church of reason. *Studies in Higher Education*, 45(1), 3–17.
- Shklarov, S. (2007). Double vision uncertainty: The bilingual researcher and the ethics of cross-language research. *Qualitative Health Research*, 17(4), 529–538.
- Shreeve, R. L. (2020). Globalisation or westernisation? The influence of global university rankings in the context of the Republic of China (Taiwan). *Compare: A Journal of Comparative and International Education*, 50(6), 922–927.
- Sikes, P., & Potts, A. (2008). *Researching education from the inside: Investigations from within*. Routledge.
- Sinnema, C. E. (2005). *Teacher appraisal: Missed opportunities for learning* [Doctoral thesis, University of Auckland]. <https://researchspace.auckland.ac.nz/handle/2292/32757>
- Sinnema, C., Hannah, D., Finnerty, A., & Daly, A. (2021). A theory of action account of an across-school collaboration policy in practice. *Journal of Educational Change*.
<https://doi.org/10.1007/s10833-020-09408-w>
- Slater-Brown, K. J. (2016). *Ability grouping in mathematics: We know it is not an equitable practice, so why are we doing it?* [Doctoral thesis, University of Auckland].
<https://researchspace.auckland.ac.nz/handle/2292/71>

- Smith, R. A. (2001). Formative evaluation and the scholarship of teaching and learning. *New Directions for Teaching and Learning*, 2001(88), 51–62.
- Song M. (2021, November 24). *Bệnh mùa bỏ phiếu* [Voting season diseases]. *Journal of Communist Propaganda*. <https://tuyengiao.vn/cung-suy-ngam/benh-mua-bo-phieu-125916>
- Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, 72(3), 387–431.
- Stack, M. (2021). *Global university rankings and the politics of knowledge*. <https://doi.org/10.14288/1.0398205>
- Stake, R. E. (1995). *The art of case study research*. Sage Publications.
- Stein, S. J., Goodchild, A., Moskal, A., Terry, S., & McDonald, J. (2021). Student perceptions of student evaluations: Enabling student voice and meaningful engagement. *Assessment & Evaluation in Higher Education*, 46(6), 837–851. <https://doi.org/10.1080/02602938.2020.1824266>
- Stephens, C. S. (1999). Faculty evaluation and continuous improvement in higher education: A literature review. *Allied Academies International Conference. Academy of Educational Leadership. Proceedings*, 4(1), 81–90. <https://www.proquest.com/scholarly-journals/faculty-evaluation-continuous-improvement-higher/docview/192407418/se-2?accountid=8424>
- Sułkowski, Ł., Przytuła, S., Borg, C., & Kulikowski, K. (2020). Performance appraisal in universities—Assessing the tension in public service motivation (PSM). *Education Sciences*, 10(7), 174. <https://doi.org/10.3390/educsci10070174>
- Sutrisno, A., Nguyen, N. T., & Tangen, D. (2014). Incorporating translation in qualitative studies: Two case studies in education. *International Journal of Qualitative Studies in Education*, 27(10), 1337–1353. <https://doi.org/10.1080/09518398.2013.837211>

- Svay, T. (2017). *Understanding the theories of action underpinning the foundation year curriculum design at Cambodian tertiary institutions*. [Unpublished master's thesis, University of Auckland].
- Taitt, M. (2011). *Components of merit compensation systems for faculty at four-year universities*. [Master's thesis, Southern Illinois University Carbondale].
https://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1085&context=gs_rp
- Teelken, C. (2018). Teaching assessment and perceived quality of teaching: A longitudinal study among academics in three European countries. *European Journal of Higher Education*, 8(4), 382–399. <https://doi.org/10.1080/21568235.2018.1490661>
- Temple, B., & Young, A. (2004). Qualitative research and translation dilemmas. *Qualitative Research*, 4(2), 161–178.
- Thanh, P. (2016). Ghi sâu lời Bác dạy: “Thi đua là yêu nước, yêu nước thì phải thi đua” [Inscribing in our minds Uncle Ho’s teaching: ‘Emulation means patriotism, those who are patriotic must emulate’]. *Central Committee of Emulation and Commendation*.
<http://www.thiduakhenthuongvn.org.vn/dien-dan/ghi-sau-loi-bac-day-thi-dua-la-yeu-nuoc-yeu-nuoc-thi-phai-thi-dua>
- Theall, M., & Centra, J. A. (2001). Assessing the scholarship of teaching: Valid decisions from valid evidence. *New Directions for Teaching and Learning*, 2001(86), 31–43.
- Thich Nhat Hanh. (1987). *Old path white clouds: Walking in the footsteps of the Buddha*. Parallax Press. <https://www.penguin.co.nz/books/old-path-white-clouds-9780938077268>
- Thich Nhat Hanh. (1991). *Peace Is every step*. Bantam.
<https://www.penguinrandomhouse.com/books/74752/peace-is-every-step-by-thich-nhat-hanh/>
- Thich Nhat Hanh. (2014, March 16). *On restlessness | Dharma talk*. Plum Village.
<https://www.youtube.com/watch?v=Qq6B4OX6uUE>

- Thich Nhat Hanh. (2022). *Vietnam: Lotus in a sea of fire*. SCM PRESS. (Original work published 1967)
- Thiel, J. J. (2019). *The production of “student voice” and its effects on academia* [Doctoral thesis, Manchester Metropolitan University]. <https://e-space.mmu.ac.uk/622766/>
- Tight, M. (2018). Quality assurance and management in higher education. *Vortrag Gehalten Auf Der 3rd International Conference on Quality Engineering and Management, Barcelona*.
- Tight, M. (2019). The neoliberal turn in higher education. *Higher Education Quarterly*, 73(3), 273–284. <https://doi.org/10.1111/hequ.12197>
- Top Universities. (2021). *QS World University Rankings – Methodology*. Top Universities. <https://www.topuniversities.com/qs-world-university-rankings/methodology>
- Tran, H. (2009). *The transformation of higher education in Vietnam after DoiMoi: A story of “dualism”* [Doctoral thesis, St. John’s University (New York)]. <http://search.proquest.com/docview/751537087/abstract/57BF1D6CEBC14A01PQ/1>
- Tran, L. T., & Marginson, S. (2014). Education for flexibility, practicality and mobility. In L. T. Tran, S. Marginson, H. Do, T. Le, N. Nguyen, T. Vu, & T. Pham (Eds.), *Higher education in Vietnam: Flexibility, mobility and practicality in the global knowledge economy* (pp. 3–25). Springer.
- Tran, L. T., & Marginson, S. (2018a). Internationalisation of Vietnamese higher education: An overview. In L. T. Tran & S. Marginson (Eds.), *Internationalisation in Vietnamese higher education Internationalisation in Vietnamese Higher Education* (pp. 1–16). Springer.
- Tran, L. T., & Marginson, S. (2018b). Internationalisation of Vietnamese higher education: Possibilities, challenges and implications. In L. T. Tran & S. Marginson (Eds.), *Internationalisation in Vietnamese higher education* (pp. 253–257). Springer.
- Tran, L. T., Nghia, T. L., Nguyen, M. N., & Ngo, M. (2020). ‘Let go of out-of-date values holding us back’: Foreign influences on teaching-learning, research and community

- engagement in Vietnamese universities. *Cambridge Journal of Education*, 50(3), 281–301. <https://doi.org/10.1080/0305764X.2019.1693504>
- Tran, N. D. (2015). Reconceptualisation of approaches to teaching evaluation in higher education. *Issues in Educational Research*, 25(1), 50–61. <https://doi.org/10.3316/informit.373427507401317>
- Tran, N. D., & Nguyen, T. T. (2015). Shifting the focus to student learning in teaching evaluation: A new direction for Vietnamese higher education institutions? *Interchange*, 46(2), 113–119.
- Tran, N. T., & Vu, P. A. (2019). Fifteen years of accreditation in Vietnam. In C. H. Nguyen & M. Shah (Eds.), *Quality assurance in Vietnamese Higher Education* (pp. 97–120). Springer.
- Tran, T. T. (2018). Sinh viên đánh giá giảng viên: Liệu có trái với thuần phong mỹ tục Việt? [Students 'evaluation of teachers: Is it against the Vietnamese cultural norm?]. *VNU Journal of Foreign Studies*, 34(6), 1-14.
- Tran, X. B. (2010). *Đánh giá giảng viên đại học theo hướng chuẩn hóa trong giai đoạn hiện nay [Faculty evaluation towards standardisation in the current period]* [Doctoral thesis, Đại học Giáo dục – Đại học Quốc gia Hà Nội]. https://repository.vnu.edu.vn/flowpaper/simple_document.php?subfolder=74/67/94/&doc=74679400953330133693556038268383400415&bitsid=34792b79-c35e-48fe-945e-c8402e68b3f0&uid=
- Tran, X. B. (2011). Đánh giá giảng viên ở các trường đại học-vấn đề bức thiết trong giai đoạn hiện nay [Faculty evaluation: An urgent issue in the current period]. *Tạp Chí Khoa Học và Công Nghệ - Đại Học Đà Nẵng*. https://scholar.google.com.vn/citations?view_op=view_citation&hl=vi&user=VDTO8ScAAAAJ&citation_for_view=VDTO8ScAAAAJ:hC7cP41nSMkC
- Travers, M. (2019). *Qualitative interviewing methods: Uncovering social meaning through interviews*.
- Truong, D. T. (2013). *Confucian values and school leadership in Vietnam*. [Doctoral thesis, Victoria University of Wellington].

<http://researcharchive.vuw.ac.nz/xmlui/bitstream/handle/10063/2774/thesis.pdf?sequence=2>

Truong, T. D., Hallinger, P., & Sanga, K. (2017). Confucian values and school leadership in Vietnam: Exploring the influence of culture on principal decision making. *Educational Management Administration & Leadership*, 45(1), 77–100.

<https://doi.org/10.1177/1741143215607877>

Turner, S. (2010). Challenges and dilemmas: Fieldwork with upland minorities in socialist Vietnam, Laos and southwest China. *Asia Pacific Viewpoint*, 51(2), 121–134.

Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, 54, 22–42.

<https://doi.org/10.1016/j.stueduc.2016.08.007>

Vargas-Madriz, S., Luis Francisco ;. Nocente, Norma ;. Best-Bertwistle, Rebecca ;. Forgie. (2019). "Somebody has to teach the 'broccoli' course": Administrators Navigating Student Evaluations of Teaching (SET). *Canadian Journal of Higher Education / Revue Canadienne d'enseignement Supérieur*, 49(1), 85–103.

<https://doi.org/10.7202/1060825ar>

Vietnam Law on Public Employees (No. 58/2010/QH12) 2010.

http://www.ilo.org/dyn/natlex/natlex4.detail?p_lang=en&p_isn=91606&p_country=VN
[M&p_count=548](#)

Vu, H. V. (2019). Comparative Buddhism in India, China, Vietnam and the spirit of localization in Vietnamese Buddhism. *International Journal of Recent Scientific Research*, 10(6), 33114–33120. <https://doi.org/DOI: 10.24327/IJRSR>

Vu, M. C., & Tran, T. (2021). Trust issues and Engaged Buddhism: The triggers for skillful managerial approaches. *Journal of Business Ethics*, 169(1), 77–102.

<https://doi.org/10.1007/s10551-019-04273-x>

Vuong, Q.-H., Bui, Q.-K., La, V.-P., Vuong, T.-T., Nguyen, V.-H. T., Ho, M.-T., Nguyen, H.-K. T., & Ho, M.-T. (2018). Cultural additivity: Behavioural insights from the interaction of

- Confucianism, Buddhism and Taoism in folktales. *Palgrave Communications*, 4(1), 1–15. <https://doi.org/10.1057/s41599-018-0189-2>
- Wei, W. (2015). Using summative and formative assessments to evaluate EFL teachers' teaching performance. *Assessment & Evaluation in Higher Education*, 40(4), 611–623. <https://doi.org/10.1080/02602938.2014.939609>
- Williams, H. A., & Preziosi, R. C. (2004). An analysis of the proposed performance based pay system for academic staff at the University of Technology, Jamaica. *Journal of College Teaching & Learning*, 1(8), Article 8. <https://doi.org/10.19030/tlc.v1i8.1974>
- The World and Vietnamese Report. (2021, June 9). Hai đại học nào của Việt Nam lọt top 1000 thế giới? [Which two Vietnamese universities are in the top 1000 in the world?]. *Báo Thế giới và Việt Nam*. <https://baoquocte.vn/hai-dai-hoc-nao-cua-viet-nam-lot-top-1000-the-gioi-147835.html>
- World Bank. (2021). *The World Bank in Vietnam: Overview*. World Bank. <https://www.worldbank.org/en/country/vietnam/overview>
- World Population Review. (2022). *Communist countries 2022*. <https://worldpopulationreview.com/country-rankings/communist-countries>
- World Population Review. (2022). *Vietnam Population 2022*. <https://worldpopulationreview.com/countries/vietnam-population>
- Xét chiến sĩ thi đua: Hiệu trưởng sao lại tranh phần giáo viên? [Considering emulation title: Why does the principal compete with teachers?]. (2017, April 11). *Tien Phong Newspaper*. <https://tienphong.vn/post-945559.tpo>
- Yao, Y., & Grady, M. L. (2005). How do faculty make formative use of student evaluation feedback? A multiple case study. *Journal of Personnel Evaluation in Education*, 18(2), 107. <https://doi.org/10.1007/s11092-006-9000-9>
- Yin, R. K. (2018). *Case study research and applications*. Sage.
- Yokuş, G., & Yelken, T. Y. (2019). Analyzing academic members' expectations from a performance evaluation system and their perceptions of obstacles to such an

evaluation system: education faculties sample. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi*, 42.

Appendices

Appendix A. Interview Protocol

A. INTRODUCTION

- **Greeting**
 - My name is .. I'm a PhD student at ...
- **Explain the purpose and procedure of the interview**
 - This research aims to understand better the faculty evaluation practices that are implemented in your institution and the impact it has on (teaching) faculty members' learning and improvement. The research findings are expected to contribute to improving faculty evaluation practices that have a positive impact on faculty members.
 - The purpose of our conversation today is to learn about your thoughts, feelings, and experiences with the faculty evaluation practices at your institution.
 - Your reports will be written in a manner that no personal attribution is made to a particular person.
 - The interview will take between 60 and 90 minutes.
- **Reviewing the participant information sheet**
- **Clarifying ethics**
 - Your participation in this interview is absolutely voluntary. Have you any questions you would like to ask? Are you willing to participate? Are you willing to be recorded?
 - Please sign the consent form.
 - Do you have any other questions before we begin?
 - Thank you.
- **Begin recording**

B. THE INTERVIEW PHASES

Phase 1: Familiarisation

- How long have you been in this role at the institution?
- What are your main responsibilities as a quality assurance officer, administrator (i.e., faculty dean or head of department), and faculty members?
- What evaluation practices have you experienced? How long have you experienced them?

Phase 2: Eliciting theories of action

Constraint Topics	Action Topics	Consequence Topics
<p>The documents and actual conditions that drive faculty evaluation practices will be probed, including the goals, beliefs, attitudes, and values that drive faculty evaluation actions, e.g., own beliefs about:</p> <ul style="list-style-type: none"> - faculty evaluation purpose, criteria, and processes - roles, responsibility & authority 	<ul style="list-style-type: none"> • Responses to evaluation requirements • Level of engagement • Uses of evaluation results 	<p>Impact on:</p> <ul style="list-style-type: none"> • Managerial decisions (e.g., accreditation, grading, reward) • Faculty members' professional learning and teaching improvement • Others (e.g., overall satisfaction level)

Phase 3: Checking theories of action

- **Checking accuracy**

- So you have identified several common evaluation practices... Have I understood you correctly?
- I would like to summarise what you've said so far to check whether I've understood you. You talked about ... Do you want to modify any of that?

- **Checking coherence**

- Checking possible coherence issues in the interviewees' account of practice, e.g., by asking explanations for differences between the stated policies and purposes and/or actual implementation: It is stated in the guidelines about ... that ... But you mentioned that ...
- It seems that there is a difference between the stated policies and actual implementation. Am I correct?

C. INTERVIEW TECHNIQUES

- **Probes**

- I am interested to hear about I am not clear about ... I am curious to know ...
- I'm not sure if I understood you correctly. Could you please clarify?
- You said ... Could you give me an example of what you mean by that? (Repeated Probe Questions)
- You said ... What was exactly the words that he used?
- You said ... Can you elaborate on what you mean by ...?

- Theory of action elements to probe for:
- **Critical incident method**
 - Can you tell me a little about your most recent faculty evaluation when you were involved as an evaluator/ one who was evaluated?
 - Can you think of an example of something that happened recently that typified these factors that you have been describing?
- **Critiquing thinking**
 - How did you reach that conclusion?
 - What leads you to that conclusion?
 - There may be other possible explanations for that ...
 - Are there other possible explanations for that?
 - What else happened? Do you think of other practices?
 - What leads you to that conclusion?
- **Notes**
 - The setting
 - Non-verbal language
 - Others

Appendix B. Sample of SET Report*Grand University – Member HEI 2***STUDENT EVALUATION OF COURSES REPORT**

(KẾT QUẢ LẤY Ý KIẾN CỦA SINH VIÊN VỀ HỌC PHẦN)

Course title: xxx

Lecturer: xxx

Semester: I

School year: 2017-2018

Number of student responses: 23

Mode of the survey: online (using Google Doc)

Rating scale (highlight a rating that corresponds to your opinion)

①	②	③	④	⑤
Completely Disagree	Generally Disagree	Not sure	Generally Agree	Completely Agree

PART I. TEACHING ACTIVITIES

	Evaluation content	Rating scores	Standard Deviation
1	The faculty member provides adequate information about the course content, schedule and assessment methods.	4.57	0.788
2	The faculty member gives students instructions on learning methods during the course.	4.43	0.788
3	The faculty member creates opportunities for students to participate actively in learning activities.	4.61	0.722
4	The faculty member helps students develop some basic soft skills (e.g., communication, presentation, group work/ working independently, problem-solving skills).	4.61	0.783
5	The faculty member helps the student develop thinking skills (critical thinking, creativity, thinking logically).	4.39	0.783
6	The faculty member is concerned about student learning matters.	4.35	0.732
7	The faculty member delivers the lessons with clarity and comprehensibility.	4.26	0.764
8	The faculty member conducts the adequate amount and content of the course as scheduled in the course outline.	4.52	0.730
9	The faculty member uses effectively teaching equipment and facilities	4.57	0.788
10	The faculty member comes to class on time.	4.57	0.788

Average rating scores (statements from 1-10): 4.48

PART II. CONDITIONS FOR QUALITY ASSURANCE AND ASSESSMENT ACTIVITIES

Evaluation content		Rating	Standard
2.1. COURSE OBJECTIVES, PROGRAM AND CONTENT		scores	Deviation
1	The course content meets the objectives in terms of knowledge, skills and attitude.	4.04	0.878
2	The course content is balanced and logical.	4.13	0.920
3	The course content is updated.	4.39	0.839
4	The course content has contributed to the equipment of knowledge and skills for students' (future) jobs.	4.13	0.014
5	The learning materials are updated and adequate.	4.35	0.885
2.2. (LEARNING) ASSESSMENT ACTIVITIES			
1	The assessment methods match with the course content and teaching methods.	4.48	0.790
2	The assessment methods can evaluate students' level of knowledge and skill attainment.	4.35	0.885
3	The assessment methods are objective and accurate.	4.48	0.790
4	The assessment results are announced within the specified timeframe.	4.30	0.876
5	Feedback from assessment results helps students improve learning outcomes	4.30	0.822
INFRASTRUCTURE CONDITIONS			
1	Lecture halls meet teaching and learning requirements (furniture, lights)	4.30	0.947
2	The equipment in the lecture hall (audio-visual equipment, utility applications, internet) satisfies the teaching and learning requirements.	4.26	0.964
3	Facilities meet students' self-study requirements.	4.26	0.864
4	Laboratory equipment (for practice sessions) meet learning requirements.	3.91	1.164

Other comments:

(blank/ no comments)

Hanoi, January 2018

Appendix C. Sample of Translation Excerpts

	Original quotes in Vietnamese	Translated version
1	<p>Sử dụng kết quả Các ban chức năng, Viện Đảm bảo Chất lượng Giáo dục sử dụng kết quả lấy ý kiến phản hồi từ các bên liên quan làm căn cứ đề xuất, tư vấn với Giám đốc ĐH Grand các giải pháp quản lý, chỉ đạo các đơn vị trong những hoạt động đào tạo cụ thể, đồng thời giám sát hoạt động cải tiến chất lượng của các đơn vị nhằm nâng cao chất lượng hoạt động đào tạo chung trong toàn ĐH Grand.</p> <p>Thủ trưởng các đơn vị đào tạo và các đơn vị liên quan khác sử dụng kết quả phản hồi từ các bên liên quan làm căn cứ để điều chỉnh các chính sách, xây dựng các giải pháp nhằm nâng cao chất lượng hoạt động đào tạo và hỗ trợ đào tạo tại đơn vị;</p> <p>Kết quả đánh giá chất lượng thông qua phản hồi từ các bên liên quan được cập nhật hàng năm và được tích hợp vào cơ sở dữ liệu đảm bảo chất lượng chương trình đào tạo của đơn vị và của ĐH Grand. (ĐH Grand, 2014, p. 5)</p>	<p>The Institute for Education Quality Assurance staff are expected to use the results of key stakeholders' surveys as a basis to:</p> <ul style="list-style-type: none"> (i) consult the Grand University's President in the management and advancement of member higher education institutions (HEIs)' specific training programs, (ii) monitor member institutions' quality improvement towards the advancement of university-wide training activities <p>Rectors of member institutions are expected to use the results of key stakeholders' surveys as a basis to:</p> <ul style="list-style-type: none"> (i) adjust policy planning of institutional training support and quality improvement <p>Results of key stakeholders' surveys are annually updated and integrated into the quality assurance database of the Grand University and its member institutions. (Grand University, 2014, p. 5)</p>
2	<p>... thế còn đưa ra cái công tác đánh giá giảng viên thì, thứ nhất là ở Khoa chỉ chính sách chủ trương phải là do Ban chủ nhiệm Khoa đưa ra, đánh giá giảng viên ở góc độ nào, đánh giá để làm gì, dừng ở đến đâu, và xử lý kết quả như thế nào, vân vân, thì tất cả những cái đấy phải do Ban chủ nhiệm Khoa đưa xuống. Ban Chủ nhiệm Khoa sẽ đưa xuống tổ chức cán bộ, tổ chức cán bộ làm công văn hướng dẫn, ví dụ trong năm học này tôi sẽ đánh giá giảng viên trên những khía cạnh này, sau đó sẽ trả xuống các đơn vị liên quan. Từ những cái chủ trương chung đấy thì bên chị sẽ tự lên một cái kế hoạch là bây giờ bên chị được phân công là đánh giá giảng viên mảng này, để đạt được việc đấy thì bên chị sẽ phải làm những việc gì trong năm để ra được cái kết quả là đánh giá giảng viên mặt nào trong năm đấy. Chung</p>	<p>... at our institution [faculty evaluation] policy intentions must be set out by the Rector Board. [Concerning] the aspects, scopes and uses of evaluation, as well as ways to deal with SET results, etc, all must be sent down by the Rector Board. The Board would send [their policy intentions] to the Office of Personnel and Administrative Affairs staff who then prepare the guideline documents ... to forward to related faculties and departments. From these guidelines, regarding the faculty evaluation have been assigned to us, we would make our own plan about what we need to do, what aspects of faculty evaluation we need to focus on during the year to achieve the intended faculty evaluation outcomes. It's our general procedures ... our specific plans tend to be different every year." (I1Q1).</p>

	chung như thế nhé, cụ thể thì chị sẽ lấy năm nay ra làm ví dụ, không có mỗi năm lại làm một kiểu khác nhau (I1Q1)	
3	Nếu mà chúng ta nói là phải theo cái tiêu chí đánh giá thì chúng tôi vẫn gọi là hai cái nó song song với nhau 1 là cái tri thức chuyên môn của người được tuyển dụng đấy nhưng cái đấy nó chỉ được 1 nửa thôi còn nửa kia thậm chí quan trọng hơn nó tự là cách đạo đức, phẩm chất chính trị tư tưởng bởi vì cái khoa này nó là như thế. (I2A4)	We normally find that the evaluation criteria can be divided into two parallel facets: one facet being the professional knowledge of the recruited faculty members while the other, which is even more important facet, is faculty moral quality, or the quality of political ideology because this is how thing goes here at our faculty. (I2A4)
4	Theo kinh nghiệm của mình thì ở Việt Nam mà không có cái ép từ trên xuống thì không ai làm. Nhưng mà chỉ có điều là cái pressure nó chỉ a little bit, nó không phải là a lot (I4A2)	In my experience in Vietnam, without top-down pressure, no one does [the jobs]. One thing to note is that the pressure is only a little bit, not a lot. (I4A2)
5	“Chị...chị cũng chỉ để tham khảo thôi, chứ còn cái mà bản thân chị đánh giá cho bản thân chị ý, bản thân cá nhân thì chị dựa trên cái mức độ xem là sinh viên, thứ nhất là cái điểm kiểm tra cuối kỳ của sinh viên, tức là mình đánh giá ở cái mức độ bài của sinh viên họ có làm, có đáp ứng được đúng cái yêu cầu của môn học hay không, vì mỗi cái môn học có một cái barem riêng.” (I2T1)	“I... I just used it for reference, but what I personally evaluate for myself, I personally base on the degree of student [learning]. At the end of the term, I evaluate the extent to which my students meet the requirements of the subject, because each subject has its own [assessment] rubrics. (I2T1)

Appendix D. Individual ToA Example

VOTING EVALUATION THEORIES OF ACTION_ FIRST ITERATION

Participant: Administrator I2A1

Constraints	Actions	Consequences
<p>EVALUATION POLICIES</p> <ul style="list-style-type: none"> • Basic criteria for the annual evaluation included faculty members’ teaching hours, research hours and research outputs. The annual evaluation is conducted twice a year – one in June when a school year ends and another in December which is related to the end-of-year financial bonus. Generally, the evaluation is based on the number of teaching hours, your research hours, and your research outputs or publications – three determinant criteria for considering a teacher’s task accomplishment • There are procedures that require certain steps: <ol style="list-style-type: none"> Individuals and organisations are required to write a report of achievements and rate their own performances Institutions/ Faculties/ Departments are required to organise meetings, collect comments on individual achievement reports, and select people with emulation titles by vote of confidence. (Individuals and units with the rate of confidence vote higher than 2/3 the total number of voters will meet the requirements for considering higher levels of titles). (Grand University, 2015, article 31) • The purpose of the annual evaluation is to rank individual faculty members with the following titles: <ol style="list-style-type: none"> Grade 1: Advanced Labourer, Advanced Fighter Grade 2: Emulation Fighter of the Grassroots Level (institutional level) Grade 3, 4 & 5: Emulation Fighter of the University, Ministerial Level & National Level (Grand University, 2015, article 10) • The number of “Advanced Labourer” titles is not restricted. However, the number of the “Grassroots-Level Emulation Fighter” is limited – being no more than 15% of the number of 	<p>MID-YEAR (VOTING) EVALUATION</p> <p>Before evaluation meeting</p> <ul style="list-style-type: none"> • Administrator I2A1 and his faculty members completed their own self-evaluation forms <p>During evaluation meeting</p> <ul style="list-style-type: none"> • Each faculty member shared their own self-evaluation report. Then they evaluate each other by a vote of confidence by ticking on paper ballots. <p>Each teacher reports his or her own self-evaluation level, and then we grade each other by a [secretive] vote of confidence</p> <ul style="list-style-type: none"> • Administrator I2A1 chose not to be critical about faculty evaluation results <p>Well, teaching staff generally receive quite a low salary, and they’ve worked hard during the year, so they deserve to get a sum of financial bonus. That’s why I don’t want to cause unnecessary tension related to evaluation results ... <i>I don’t want to give a low grade for any members because they may suffer a loss in financial bonus from our institution. I don’t like that to happen, and I want to support and encourage all faculty members</i></p> <ul style="list-style-type: none"> • Administrator I2A1 often stated general comments about faculty performance at the meeting • Administrator I2A1 often produced a “quite nice” report of faculty performance <p>I only reminded our staff with things like “well, there’re some tasks we need to make an effort to</p>	<p>INTENDED CONSEQUENCES</p> <ul style="list-style-type: none"> • The existing evaluation system helped monitor faculty members’ quantity of work <p>The influence of the current evaluation system is limited to control faculty members in terms of how many teaching hours, research hours and journal articles one has completed</p> <ul style="list-style-type: none"> • The existing evaluation system helped to maintain the stability of the institution <p>Well... how this evaluation is meaningful... well, in terms of organisational management, the stability of an organisation is reinforced by ... the culture or the interaction with its staff. If, for example, we could be fair and give more opportunity to staff during the evaluation, then the related tension and rumour would diminish, thus creating stability to the organisation</p> <p>UNINTENDED CONSEQUENCES</p> <ul style="list-style-type: none"> • The evaluation results did not reflect the quality of faculty members’ work and research impact <p>The quality of faculty members’ work is not recognised during the evaluation. We know this faculty</p>

<p>individuals with “Advanced Labourer” titles; AND with no more than 1/3 individuals holding the leadership roles within the Department (Grand University, 2015, article 10)</p> <p>NORMS</p> <ul style="list-style-type: none"> • Administrator I2A1 perceived an affection-over-rule norm which means that the Vietnamese people do not often critically or negatively evaluate others. <p>... I don't want to give a low grade to any members because they may suffer a loss in financial bonus from our institution. I don't like that to happen, and I want to support and encourage all faculty members (I2A1)</p> <p>Normally everyone is rated as having completed all tasks ... As you know, our Vietnamese culture of practices being ruled by “affection” so rarely do people (evaluate others critically) ... Sometimes teachers may have poor performance, but when it comes to the evaluation, other people are afraid of those teachers suffering from a salary cut, so they just let it go .. that's the way it is here (I2A1)</p> <p>In some cases, even the violation of the regulation, after consideration, was then proposed as “light misconduct” and “let bygones be bygones” sort of thing. That's why no one would be graded in the low level and so all would receive grade A ☺ that's the evaluation situation here (I2A1)</p> <p>VALUES AND BELIEFS</p> <ul style="list-style-type: none"> • Administrator I2A1 value teachers' roles and individual contribution <p>Previously special titles used to be awarded to leaders. But since I became the dean, I tried to promote the important roles of the teaching staff who deserve (to be appreciated by others)</p> <p>Frankly speaking, the biggest workload and contribution belong to the faculty leader, but I refused to be nominated for special titles as I followed the private sectors' approach to honour staff than leaders</p> <ul style="list-style-type: none"> • Administrator I2A1 expected to motivate faculty members to become the best experts in their field. 	<p>complete such as this or that ...,” while the evaluation results on the paper look quite nice</p> <ul style="list-style-type: none"> • Administrator I2A1 asked the faculty to nominate teachers with outstanding achievements with special titles • Administrator I2A1 typically limited number but sometimes allows more to be nominated. <p>Well, before voting, I ask each teacher group to share their opinions about who deserves which title, such as Advanced Labourer or Emulation Fighter, and we have a limited number of chosen teachers for such titles</p> <p>I normally tried to negotiate teachers to nominate a limited number of candidates corresponding to the chosen number so that a high percentage of votes is yielded for each chosen candidate. For example, if we can only choose two teachers while three are nominated, then each of the nominated will not get 100% votes – the results may go like 70%, 80% and 30%, for example. I tried to avoid such circumstances</p> <p>For the past two years, I find that it's better to allow a little bit more teachers to be nominated so that the voting is democratic ... Previously some teaching staff were given more priority in nomination but when it comes to the voting, they did not get enough votes by other members ...</p> <ul style="list-style-type: none"> • Administrator I2A1 turned down his nomination to give more opportunities for faculty members. <p>I know other faculties who were very intense in proposing certain members to be rewarded, but I want to ensure fairness in our faculty. I myself refused to be rewarded, giving the opportunities to</p>	<p>member has a publication, but we don't know its impact. Some teaching staff pay attention to their teaching quality, whereas others just go to the classroom to get enough teaching hours</p> <ul style="list-style-type: none"> • “Everyone wins” evaluation hindered individual effort and motivation for improvement <p>I think the collectivist nature of the current evaluation results – everyone rated at grade A as having completed all tasks. Such “everyone wins” evaluation situation hinders motivation for better performance</p> <ul style="list-style-type: none"> • Evaluation results linked to unfair incentives may cause low motivation and dissatisfaction to Administrator I2A1 and other teachers. <p>The current salary and financial bonus is scale-based in which the scale depends on your years of experience rather than your actual position and performance</p> <p>Some senior officers who only specialise in one certain work get higher salary than me. My workload, responsibility and working hours as a faculty dean is much higher than them, but I get a lower salary because of my lower seniority. Such salary and financial scheme are to show that the institution's evaluation and reward</p>
---	--	---

<p>My point is to motivate every faculty member to become the leading expert in the field they are working by learning continuously, participating in projects, engaging in knowledge exchange for better professional development</p> <ul style="list-style-type: none"> • Administrator I2A1 wanted to adopt a more practical and motivational approach to faculty evaluation ... <ul style="list-style-type: none"> ○ <i>By advocating the bottom-up construction of Key Performance Index (KPI), a performance-based evaluation, though he is aware of its obstacles</i> <p>My expectation to construct a KPI system for evaluation originates from my experience working in the private sector. I want to create a working environment that booster faculty members' motivation to work but it's difficult to do so in the public sector... Obviously, it's like a bottom-up approach in which teaching staff at our faculty could discuss the evaluation content ...</p> <p>However, KPI evaluation may pose the problem of linking more to the quantity than the quality. Now we need a KPI system that can help evaluate both the quantity and quality of faculty work – such a difficult task</p> <ul style="list-style-type: none"> ○ <i>By proposing a more specific job description and corresponding benefits linked to the evaluation system.</i> <p>The decision-making still belongs to the Institutional Management Board and the HR department who suggest different categories for staff of different levels such as associate professor, doctor, or master. Though I haven't figured out what the add-on evaluation would be like, but generally a clear set of job description is a must</p> <p>I think that using KPI and position-based salary scheme would motivate faculty members to work better</p> <ul style="list-style-type: none"> ○ <i>By being aware of potential obstacles linked to the proposed KPI system</i> ○ <i>By asking for more autonomy</i> <p>I think evaluation would be more effective if I had more power, autonomy, and resources for evaluating faculty members</p>	<p>the teaching staff to motivate them. They help me in my daily work, so they deserve to be honoured</p> <p>After evaluation meeting</p> <ul style="list-style-type: none"> • Administrator I2A1 reconsidered the voting result and then forwarded it to an upper level of administration. <p>I consider the proposed evaluation results from the teacher groups then send it to the institutional meeting for another round of voting</p> <ul style="list-style-type: none"> • The upper administrators brought the nomination list to an institutional meeting and all staff evaluated members in the list by vote of confidence. <p>The voting is conducted at the faculty level before taking place at the institutional staff meeting</p> <ul style="list-style-type: none"> • Administrator I2A1 rewarded staff financially for national and international publication <p>Before monthly meetings, I ask my assistant to summarise the number of faculty members with publication to grant them with small bonus (i.e., 13 USD for a national publication and 33 USD for an international publication) from the faculty-led budget to encourage faculty members to have more publication</p> <ul style="list-style-type: none"> • Administrator I2A1 tried to find a more systematic way to monitor faculty performance and evaluation <p>Besides other requirements from the institution, I make my own statistics of faculty performance. I like statistics, so I make a list containing 20 points related to faculty members' teaching hours, research hours, publication types, and conference attendances to monitor their performance. I base on the list to monitor what our teaching staff are doing ...</p>	<p>is not accurate, causing low motivation among us</p> <p>To be fair, I am not satisfied at all with the current evaluation practices</p> <ul style="list-style-type: none"> • Administrator I2A1 had difficulty in balancing between urging teaching staff to work and meeting individual teachers' demands. <p>Most staff in our faculty work with their internal enthusiasm for the job, being supportive of executing my new ideas and initiatives. However, many staff don't like that, and they are, how to say, ... negligent .. just to get the work done</p> <p>What is the normal tendency of our teaching staff? Most teachers want their reward to be stable while their workload and requirements reduced... It means that they keep focusing on their individual demands (without focusing on institutional requirements that I must bear) ... I get frustrated with it and find it unfair ... "You asked for your benefits to be stable, but you also asked institutional requirements to reduce ... It's unreasonable, no way!" That's how it is.</p>
--	--	--

<p>The Dean thinks his responsibility with student evaluation is to monitor and solve problems related to evaluation results from students. If the score is lower than 3 or 2 (below average), I will work with teachers to solve it. ..</p> <p>However, there's not many problems related to evaluation results from students. .. Except for some students whose opinions are straightforward, the rest of the average scores were all good.</p> <p>Students are a bit intuitive when evaluating teachers ... so the scores given are quite high, without variations. Most scores are the same.. So mostly I have nothing to analyse.</p> <p>The Dean thinks student evaluation of teachers are not so practical due to its commonly high results.</p> <p>However, the majority of student evaluation scores are over 4, so I do not have much work to do... I don't have anything to warn our teachers... The fact that evaluation scores range from 4.2 to 4.5 does not make any difference.</p> <p>I think the quality of formal student evaluation results is not high because it does not differentiate different levels of teaching. In my personal opinion, the student feedback is a bit formalistic (<i>characterised by formal rules than practical matters</i>)</p> <p>It's a norm at the faculty that students give negative feedback about teaching through informal channels However, students (normally) give negative feedback about teaching through informal channels ... They talk to each other, then through another teacher. They have groups to share their evaluation of teachers, which is not present in the paper forms.</p> <p>The Dean values feedback from informal channels. Student feedback from informal channels is sometimes more practical than those taken from formal sources.</p>	<p>SET</p> <p>10. Monitor student evaluation results The QA director sends me (the results of all members) so that I can forward to them. They also send me a separate pile of all each teacher's scores and feedback.</p> <p>11. Talk to the leaders and teachers when seeing negative comments from students With some specific written comments from students like "this teacher is like this or like that," I will work with the teacher group leader to talk to the teacher to improve a little bit – in a way that brings a bit more educational value to students. Examples of such circumstances were the teachers being late to class or talking to students in an improper manner.</p> <p>I also talk to that specific teacher, but generally, we have received few specific comments from students.</p> <p>Informal student feedback</p> <p>12. Try to discuss particular student feedback with the faculty members in the open rather than a judgmental manner We have never had any negative feedback from students. There's sometimes feedback (from informal sources). When it happens, I try to discuss with the teachers' feedback with faculty members in the open rather than judgmental manner. For example, I ask faculty members questions like, "We've got such feedback from students, what's your opinion about it? What change do you think you would make for your next classes?"</p>	<p>Student rating scores being one source of reference only We only use formal student rating scores as one source of reference.</p> <p>Perceived low level of impact The impact of formal evaluation from students is not much ... because the evaluation is not so practical.</p>
---	---	---

Appendix E. Group ToA Example

ADMINISTRATORS' SET THEORIES OF ACTION_FIRST ITERATION

CONSTRAINTS	ACTIONS	CONSEQUENCES
<p>EVALUATION POLICIES</p> <p>Institutional SET policy expectations:</p> <ul style="list-style-type: none"> - purpose of using SET results for quality improvement - administrators to be notified of faculty members' SET underperformance (Grand University, 2010, 2014) <p>VALUES</p> <p>Administrators found SET results are useful for them to</p> <ul style="list-style-type: none"> - monitor and discover any problems with faculty members' teaching (I2A2, I4A1, I4A2, I4A5, I6A1) - compare teaching performance across individuals and faculties (I6A1, I4A1, I4A3, I4A5) - warn underperformed faculty members (I1A2) 	<p>GROUP ACTIONS 1. CHECKING SET RESULTS</p> <p>1.1. Administrators looked at SET scores to see</p> <ul style="list-style-type: none"> - how students generally evaluated teachers. (I4A5) - if any faculty members' scores fell below standard. (I2A2, I4A5) - if students have negative comments or specific requirements. (I4A1, I4A2, I6A1) <p>1.2. Administrators compared SET scores across individual members and across different faculties. (I6A1, I4A1, I4A3, I4A5)</p> <p>1.3. Administrators mainly used SET results to monitor teachers. (I1A1, I2A3)</p>	<p>INTENDED CONSEQUENCES</p> <p>Faculty members, especially young ones, became more mindful about teaching (I4A1 & I2A2)</p> <p>UNINTENDED CONSEQUENCES</p> <p>Many faculty members did not value SET (I4A4)</p>
<p>NORMS</p> <p>Administrators thought they should not interfere with faculty members teaching because of the Vietnamese culture of harmony: an obligation to show respect to colleagues, especially senior ones (I1A1, I2A1, I2A4, I3A1, I4A1)</p> <p>BELIEFS</p> <p>Administrators believed that students' low SET ratings and negative comments were caused by faculty members' ineffective communication with students. (I1A3, I2A3, I2A3).</p>	<p>GROUP ACTION 2. SOLVING SET PROBLEMS</p> <p>2.1. Administrators rarely had private talks with faculty members who received low SET scores. (I1A2, I1A3, I2A1, I2A2, I2A3, I3A1, I4A4, I4A5)</p> <p>2.2. Administrators stated general reminders in faculty meetings that some teachers need to pay attention to teaching. (I2A4, I4A5)</p> <p>2.3. Administrators avoided mentioning or criticizing faculty members' low SET scores. (I1A1, I2A1)</p>	<p>Administrators were uncertain that SET had an impact on individual faculty members' teaching improvement (I2A4, I6A1)</p>

<p>Administrators did not think that they could use SET for teaching improvement due to a lack of academic autonomy. (I1A1, I2A4)</p> <p>PERCEPTIONS</p> <p>Administrators thought SET results are only used to discover if something unusual happens or if teachers had any serious problems. (I1A1, I2A3, I4A1, I2A3, I4A2)</p> <p>Administrators thought making some changes related to teacher arrangement would temporarily prevent further complaints from students. (I1A1, I4A4)</p>	<p>2.4. Administrators made some changes to avoid students' complaints (e.g., changing class for underperformed faculty members). I1A1, I4A4</p> <p>2.5. Administrators helped faculty members and students to find common ground (I4A3, I2A3, I4A4)</p> <p>2.6. An administrator combined peer evaluation of teaching, class meeting and giving feedback to faculty members. (I1A3)</p>	
<p>NORMS</p> <p>Administrators found SET results were generally positive: very few teachers had SET scores below standards (I2A2, I2A4, I4A5, I2A1, I4A2, I5A1, I2A1).</p> <p>PERCEPTIONS</p> <p>Administrators were concerned about the values of SET for teaching improvement (I1A1, I2A4, I4A1, I4A2)</p> <p>Administrators saw SET as an administrative task. (I5A1, I2A3).</p> <p>SET POLICIES</p> <p>Lack of policy indication guidelines on using SET for individual and institutional improvement</p>	<p>GROUP ACTIONS 3. INACTIONS</p> <p>3.1. Administrators did not have private talk with teachers (I2A4, I4A5).</p> <p>3.2. Administrators did not use SET scores for any specific purposes (I5A1, I4A2)</p> <p>3.3. Administrators did not use SET to</p> <ul style="list-style-type: none"> - motivate faculty members (I2A3) - provide professional support to faculty members (I1A1) 	

Appendix F. Participant Information Sheet



School of Learning, Development & Professional Practice
Epsom Campus

4 Epsom Avenue
Auckland, New Zealand
Telephone: +64 9 623 8899

The University of Auckland
Private Bag 92019
Auckland, New Zealand

PARTICIPANT INFORMATION SHEET

Individual and Focus Group Interviews
For the Teaching Faculty

Project Title: Higher education faculty evaluation practices in Vietnam: A theory-of-action multiple case study

Supervisor: Associate Professor Claire Sinnema

Student Researcher: Lan Anh Thi Nguyen

I am writing to invite you to participate in our research project about faculty evaluation practices in higher education in Vietnam. Please see the following table of questions and answers that outlines key information about the study:

<p>What is the research about?</p>	<p>The purpose of this research is to investigate faculty evaluation in higher education in Vietnam. Specifically, the research aims to address the following questions:</p> <ol style="list-style-type: none"> 1. <i>What approach is currently taken to faculty evaluation?</i> 2. <i>What explains the approach currently taken to faculty evaluation?</i> 3. <i>What is the impact of the approach currently taken to faculty evaluation?</i>
<p>What am I being asked to do?</p>	<p>Your participation in the research is entirely voluntary. The Rector of your institution has given an assurance that your participation or non-participation will not affect your employment conditions and relationship with the institution in any way. A maximum of 7 participants, including 2 midlevel administrators, 1 quality control officer, and 4 teaching faculty from your institution, will be selected to take part in the research. If you choose to participate, you will be requested to take part in</p> <ul style="list-style-type: none"> - Up to two individual interviews with the researcher about your experience of the current faculty evaluation approach. <ul style="list-style-type: none"> ▪ <i>The first individual interviews will last about 60 minutes.</i> ▪ <i>If further clarification is needed, you will be asked to attend a 30-minute follow-up interview in no more than six weeks from the first interview.</i> ▪ <i>In the individual follow-up interview, you will be asked to bring key documents describing faculty evaluation purposes and processes you have referred to in the first individual interview.</i> ▪ <i>The interviews will take place at a location that is convenient to you.</i> - One focus group interview where you discuss the researcher's summary of faculty evaluation at your institution.

	<ul style="list-style-type: none"> ▪ <i>The focus group interview with up to 6 participants in your institution who have participated in the individual interviews.</i> ▪ <i>The focus group interview will last between 60 and 90 minutes.</i> ▪ <i>The focus group interview will be conducted within 10 weeks from the first interview. The focus group interview will take place at your institution.</i> <p>All the interviews will be audio-recorded and transcribed by the researcher.</p>
Can I change my mind about participating?	<p>You have the right to withdraw at any time without explanation. You can withdraw your data at any time up until after the latest of the following dates: one month following the data having been gathered, or one month after you have received your transcripts.</p> <p>During the individual interviews, you can request the voice recording device to be turned off at any time. During the focus group interview, you may refuse to answer any questions and are free to leave the group interview without explanation. However, due to the nature of the group situation, the recording device cannot be turned off during the discussion and, if you withdraw from the research, the information you contribute up to that point cannot be withdrawn.</p>
Can I review and edit my interview recordings or transcripts?	<p>You will be given an opportunity to review and edit transcripts of your own individual interviews within two weeks after the receipt of the transcripts. You can edit the transcripts of your individual interviews within two weeks after receipt of the transcripts. However, you cannot review or edit the transcript of the focus group interview due to the complex nature of a group discussion.</p>
Will it be confidential?	<p>Your identity will be kept confidential. Participants will not be identified – pseudonyms will be used for all Departments/ Faculties and participants. Identifying details will be removed or changed to limit the possibility of participants being identified through written material.</p> <p>You will be requested to give assurance that you keep focus group discussion confidential and will not identify other participants. However, it cannot be guaranteed that other participants will comply with this request.</p>
Is there any compensation to be made to participants?	<p>Participants in this research will receive a book voucher (approximately 10 USD) for each round of the interviews.</p>
Why am I being asked to take part?	<p>All 12 higher education institutions in the national university have been invited. Your institution is among four institutions whose Rectors gave permission to participate and were selected based on the presence of a quality assurance centre, the training programs (e.g. the liberal arts and professional training programs), and the duration of establishment. If your Faculty/ Department has more than 40 faculty teaching members, you are in one of 40 members randomly selected from the list of faculty. If your faculty/ department has fewer than 40 faculty members, all have been invited to take part in this research. I am seeking teaching faculty to participate who:</p> <ul style="list-style-type: none"> (i) hold a full-time position; (ii) have been in the institution for at least three years; and (iii) have recent experience (in the last year) with the faculty evaluation program

	In case there are more participants willing to participate than needed, I will randomly select four teaching faculty to ensure the final sample has two tenure and two non-tenure teaching faculty.
How will the findings be used?	Data from this study will be used for the purposes of writing a Doctoral thesis and for academic publications including journal articles, book chapters and conference presentations. The data will only be used for the purpose of this research.
How is the data stored and deleted?	<p>While the researcher is in Vietnam, during data collection, electronic data will be stored in the researcher's password-protected personal computer and backed up with the storage in a secure University of Auckland server. The paper files will be stored in locked cabinets in a secured location such as in the researcher's office.</p> <p>In New Zealand, electronic data will be stored on a University of Auckland password-protected computer and backed up in a secure University of Auckland server. Participants' signed Consent Forms will be kept separate from the data and stored in the supervisor's office. Other types of paper data will be stored in a locked cabinet on university premises for six years.</p> <p>After six years, all electronic data will be permanently deleted from the research team's computer and any paper will be shredded and placed in secure disposal bins at the University of Auckland.</p>
Can I receive a summary of the findings	If you wish to receive a summary of the findings, please indicate that in the Consent Form. A summary of the findings will be emailed to you at the conclusion of the study.
What should I do next?	If you are willing to participate, please read and complete the attached consent form and email it to Lan Anh Thi Nguyen – nthi176@aucklanduni.ac.nz .

Thank you for reading the information and considering your participation in the research. If you have any queries about the research, please do not hesitate to contact:

<p>Researcher Lan Anh Thi Nguyen nthi176@aucklanduni.ac.nz Phones: +84 3 5595993 (xtn: 13, Vietnam) +64 22 477 1647 (New Zealand) Address: 99 Nguy Nhu Kon Tum, Thanh Xuan, Dong Da, Hanoi, Vietnam</p>	<p>Supervisor Associate Professor Claire Sinnema c.sinnema@auckland.ac.nz Phone: +64 9 6238899 (xtn: 46426)</p>	<p>Head of School <i>School of Learning, Development and Professional Practice</i> Dr Richard Hamilton, rj.hamilton@auckland.ac.nz, Phone +64 9 923 5619</p>
<p>Or for any queries regarding ethical concerns, you may contact the Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Research Office, Private Bag 92019, Auckland 1142. Telephone: +64 9 373-7599 (ext. 83711) Email: ro-ethics@auckland.ac.nz</p>		

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 17/08/2017 FOR THREE YEARS, REFERENCE NUMBER 021735

Appendix G. Consent Form



School of Learning, Development & Professional Practice
Epsom Campus

4 Epsom Avenue
Auckland, New Zealand
Telephone: +64 9 623 8899

The University of Auckland
Private Bag 92019
Auckland, New Zealand

CONSENT FORM

For the Teaching Faculty (Individual and Focus Group Interviews)

THIS FORM WILL BE HELD FOR A PERIOD OF 6 YEARS

Project Title: Higher education faculty evaluation practices in Vietnam: A theory-of-action multiple case study

Supervisor: Associate Professor Claire Sinnema

Student Researcher: Lan Anh Thi Nguyen

I have read the Participant Information Sheet and have understood the nature of the research. I have had the opportunity to ask questions and have them answered to my satisfaction.

I understand that

- Participation in this research is entirely voluntary. The Rector of our higher education institution has given an assurance that my participation or non-participation will not affect my employment conditions and relationship with the institution in any way.
- I will be requested to take part in up to two individual interviews and one focus group interview.
- I am free to withdraw my participation at any time without giving a reason and to withdraw any data traceable to me up to one month following the data having been gathered or one month after I have received my transcripts.
- During the individual interviews, I can request that the recording device be switched off at any time during the interview. During the focus group interview, I can refuse to answer any questions in the group discussion and am free to leave the discussion at any time without giving a reason. The recording device cannot be turned off during the group discussion, and that if I choose to withdraw from the research, information that I have contributed up to that point cannot be withdrawn.
- In the follow-up individual interview, I will be asked to bring key documents describing faculty evaluation purposes and processes I have referred to in the first individual interview.
- I can review and edit the transcripts of my own individual interviews within two weeks after receipt of the transcripts.
- I cannot review or edit the transcript of the focus group interview due to the complex nature of a group discussion.
- I give assurance that I keep focus group discussions confidential and will not identify other participants. However, it cannot be guaranteed that other participants will comply with this request.

- The findings will be reported in a doctoral thesis and for academic publications, including journal articles, book chapters and conference presentations.
- Any reports or publications arising from this research will not identify my institution's name or my identity.
- Electronic data will be stored on a password-protected computer backed up with the storage in a secure University of Auckland server. Participants' signed Consent Forms will be kept separate from the data and stored in the supervisor's office. Other types of paper data will be stored in a locked cabinet on university premises.
- Data will be kept for six years, after which they will be destroyed by shredding of paper data and permanent deletion of electronic files.

I agree to take part in this research. In particular, I agree

- To participate in up to two individual interviews with the researcher taking up to 90 minutes.
- To participate in a focus group interview between 60 and 90 minutes.
- To be audio-recorded.
- To give assurance that I will not identify other participants in the focus group interview.

Please circle your option: I wish/ do not wish to receive a summary of the findings, which can be emailed to me at this email address:

.....

Please put in front of each sentence a tick (✓) if it is true for you and a cross (✗) if it does not apply to you:

- I hold a full-time position;
- I have been in the institution for at least three years; and
- I have recent experience (in the last year) with the faculty evaluation program

Please circle the option that is true to you: I am a tenure/ non-tenure teaching faculty member.

Name: Click here to enter text.	Institution: Click here to enter text.
Signed: Click here to enter text.	Date: Click here to enter a date.

By typing your full name, you are signing this consent form electronically. You agree that your electronic signature is the equivalent of your written signature on this consent form. Once you have completed this form, please go to the File menu, press "Save and Send" and then "send as attachment" to email to nthi176@aucklanduni.ac.nz.

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON 17/08/2017 FOR THREE YEARS, REFERENCE NUMBER 021735

Appendix H. Sample of SET Form

Grand University – HEI 1

STUDENT EVALUATION OF TRAINING COURSES

(PHIẾU LẤY Ý KIẾN CỦA SINH VIÊN VỀ HỌC PHẦN)

Course title:

Code:

Major:

Faculty members' full name:

School year: 2017-2018 Semester:

Date of survey:

Dear students,

To improve the training quality, please provide your feedback about the past course. The table below includes various statements about the course. Please read each statement carefully and indicate your degree of agreement/ disagreement with each of them.

Rating scale

①	②	③	④	⑤
Completely Disagree	Generally Disagree	Neutral	Generally Agree	Completely Agree

Please highlight a rating that corresponds to your opinion.

PART I. TEACHING ACTIVITIES

Evaluation content		Rating scale				
1	The faculty member provides adequate information about the course content, schedule and assessment methods.	①	②	③	④	⑤
2	The faculty member gives students instructions on learning methods during the course.	①	②	③	④	⑤
3	The faculty member creates opportunities for students to participate actively in learning activities.	①	②	③	④	⑤
4	The faculty member helps students develop some basic soft skills (e.g., communication, presentation, group work/ working independently, problem-solving skills).	①	②	③	④	⑤
5	The faculty member helps the student develop thinking skills (critical thinking, creativity, thinking logically).	①	②	③	④	⑤
6	The faculty member is concerned about student learning matters.	①	②	③	④	⑤
7	The faculty member delivers the lessons with clarity and comprehensibility.	①	②	③	④	⑤
8	The faculty member conducts the adequate amount and content of the course as scheduled in the course outline.	①	②	③	④	⑤
9	The faculty member uses effectively teaching equipment and facilities	①	②	③	④	⑤
10	The faculty member comes to class on time.	①	②	③	④	⑤

PART II. CONDITIONS FOR QUALITY ASSURANCE AND ASSESSMENT ACTIVITIES

Evaluation content		Rating scale				
2.1. COURSE OBJECTIVES, PROGRAM AND CONTENT						
1	The course content meets the objectives in terms of knowledge, skills and attitude.	①	②	③	④	⑤
2	The course content is balanced and logical.	①	②	③	④	⑤
3	The course content is updated.	①	②	③	④	⑤
4	The course content has contributed to the equipment of knowledge and skills for students' (future) jobs.	①	②	③	④	⑤
5	The learning materials are updated and adequate.	①	②	③	④	⑤
2.2. (LEARNING) ASSESSMENT ACTIVITIES						
1	The assessment methods match with the course content and teaching methods.	①	②	③	④	⑤
2	The assessment methods can evaluate students' level of knowledge and skill attainment.	①	②	③	④	⑤
3	The assessment methods are objective and accurate.	①	②	③	④	⑤
4	The assessment results are announced within the specified timeframe.	①	②	③	④	⑤
5	Feedback from assessment results helps students improve learning outcomes	①	②	③	④	⑤

Other comments for the improvement of the course teaching

.....

.....

.....

Thank you!