

Jernej Širok, Martina Mravlja

PHASE MODALITY OF EXTERNAL EVALUATIONS IN HIGHER EDUCATION: EXPLORING THE PROPERTIES OF STUDY PROGRAMME EVALUATIONS IN SLOVENIA

ABSTRACT

To better understand the evaluation practices set by the Bologna process as well as their properties and implications, this paper critically explores the judgements of quality in external evaluations in higher education with regard to their orientation towards three phases: conditions, processes and end states. It furthermore explores how this modality is connected with how critical external evaluations are. It develops a theoretical framework for observing evaluation practices to support the statistical analysis of external evaluation reports for 485 study programmes in Slovenian higher education. The findings offer insight into how quality assurance impacts higher education in practice, considering that the subsequent measures taken by the higher education institutions correspond with the outcomes of the evaluations.

Keywords: *impact of quality assurance, properties of external evaluations, phase modality of external evaluations, criticality of external evaluations, higher education study programmes*

FAZNI NAČIN ZUNANJIH EVALVACIJ V VISOKEM ŠOLSTVU: RAZISKOVANJE LASTNOSTI EVALVACIJ ŠTUDIJSKIH PROGRAMOV V SLOVENIJI – POVZETEK

Kritična raziskava presoj kakovosti pri zunanjih evalvacijah v skladu z bolonjskim procesom pojasnjuje lastnosti in vpliv evalvacijskih praks tako, da obravnava način izrekanja sodb glede na tri faze: pogojno, procesno in fazo končnega stanja. Obravnava tudi, kako je ta način povezan s kritičnostjo presoj. Izbaha iz teoretskega okvira evalvacijskih praks, ki omogoča kontekstualizacijo rezultatov statistične analize evalvacijskih poročil strokovnjakov v postopkih podaljšanja akreditacije 485 študijskih programov v slovenskem visokošolskem prostoru. Rezultati raziskave prikazujejo, kako zagotavljanje kakovosti prek ukrepov za izboljšanje v praksi vpliva na visoko šolstvo, ter spodbujajo tako k skrbnemu opredeljevanju standardov kakovosti in usmeritev za presojo kot k boljšemu razumevanju lastnosti in vpliva evalvacijskih praks.

Ključne besede: *vpliv zagotavljanja kakovosti, lastnosti zunanjih evalvacij, fazna modalnost zunanjih evalvacij, kritičnost zunanjih evalvacij, visokošolski študijski programi*

Jernej Širok, PhD student, Faculty of Arts, University of Ljubljana, jernej.sirok@gov.si

Martina Mravlja, PhD student, Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, martina.mravljja@nakvis.si

INTRODUCTION

In the European system of quality assurance, external evaluations are based on establishing links between specifications, such as standards of quality, the actual state of affairs which is the perceived reality of what the evaluators scrutinised, and, in parts where specifications or the evaluators exceed the level of objectivity, also with the ideals of what is good. Research of this undertaking revolves around the conceptualisation of quality, organisation and operationalisation of external quality assurance, as well as around its outcomes, effects and implications (Biesta, 2010; Collini, 2012; Harvey & Green, 1993; Van Kemenade et al., 2008). However, prior research has arguably focused less on the functioning of evaluation practices at the ontological, epistemological and methodological level (Bornmann et al., 2006; Seyfried & Pohlenz, 2018; Tavares et al., 2016). To our knowledge, previous research has not yet focused on the phase modality of external evaluations, on corresponding offsets or on its relationship to the way standards of quality are defined. The purpose of the research is to observe the extrinsic properties of external evaluations in relation to sets of their intrinsic properties or modalities that are derived through interpretation from the content of proclaimed judgements.

The research revolves around the question how phase modality and its offsets function in practice, as well as how phase modality helps to understand and how it relates to the criticality of evaluations. It hypothesises that there is a link between the phase modality and criticality of external evaluations. The research question further touches on the possibility of evaluation practices to produce overly positive appearances of quality, or to conflate the techniques and processes in the name of quality assurance with what is or is not good in higher education. It is therefore also important whether evaluators exhibit reluctance to passing judgements on end states – i.e. on outcomes or results – where this should be done, and instead resort to judging conditions and processes and thus do not reveal the quality of what has really been observed or has happened.

THEORETICAL FRAMEWORK

The direct extrinsic properties of external evaluations are frequency and criticality. The first refers to the question whether evaluators passed a judgement based on a certain specification or not. Hence, it is binary. *The Standards and Guidelines for Quality Assurance in the European Higher Education Area* are not hierarchically structured (European Association for Quality Assurance in Higher Education [ENQA] et al., 2015). Following these guidelines, no Slovenian standards of quality have so far been prescribed as more important than others (Slovenian Quality Assurance Agency [SQAA], 2014). However, the empirical results show that in practice some specifications are more often the basis for judgement than others. The second property refers to the varying grades that judgements manifest. These grades may range from examples of excellence, strengths, compliance, opportunities for improvement, threats, to examples of inconsistencies or non-compliance depending on the specifics of the national external evaluation systems. If the grades in a

national quality assurance system, like in Slovenia, allow for it, criticality can for a given specification be understood as a relation. The latter can be defined as the ratio between the share of positive evaluations, proclaimed for instance as strengths, and the sum of shares of grades proclaimed as opportunities for improvement and inconsistencies, which represent evaluations that are critical or negative. While both extrinsic properties materialise through the proclamation of a grade, another, indirect one, can be derived from predispositions that judgements have in specifications or guidelines. Predisposition means that the way a specification or a guideline for evaluation is defined influences whether evaluators will at all evaluate a certain (aspect of the) state of affairs, and if so, how they will do it. Since all three properties come about as proclamations, predefined regulations or guidelines, they are extrinsic to or independent of the observer of evaluation practices, the evaluated state of affairs and the way a judgement is substantiated. Especially for the first two properties, no or little interpretation is required to identify them – they tell if and what they are. To demonstrate, evaluators may characterise a judgement as a strength even though it conveys no quality of the evaluated state of affairs or no relation to the ideal of quality harboured by the observer.

Modalities are intrinsic to the proclaimed judgement. They arise not from its external property, for instance, from a standardised grade, but from the way the judgement is articulated and substantiated. Since modalities depend on language and content, they are connected with ontological, epistemological and methodological properties of quality. Concretely, *conceptual modality* is a type of ontological modality derived from the possibility of applying different essentialist and functionalist concepts of quality (Harvey & Green, 1993). Another one, a type of *epistemic and methodological modality*, is that of treating quality as a matter of commensurable fact which results in a relative judgement, or that of passing an absolute judgement of value based on an ideal (Wittgenstein, 1965). While the former focuses on material existence and properties of the object under evaluation with regard to an objective specification, be it an indicator or a criterium, the latter focuses on qualities or recognition thereof and, however professional, unavoidably leans on quality related opinions, values, ideals or concepts.

This brings us to the third modality, which is the focal point of this research. *Phase modality* is a derivative of ontological modality that serves to distinguish whether judgements are passed on conditions or inputs (conditional phase of quality); on processes and procedures including the performative and transformative aspect of quality (process phase of quality); or on end states, end phenomena, results, outputs or outcomes (end phase of quality; hereinafter shortened to: end states) (Ben-Gal & Dror, 2016; Thareja, 2009).

Phase modality splits quality into three phases and invites the observer to approach quality on the axis from promise or possibility of quality to action or change towards enhanced, eventual or just possible quality, and finally to quality which has been achieved, obtained, demonstrated or recognised. Attention shifts towards how close judgements get to invoking the quality of end states, especially if the observed specifications too refer to end states and are defined with the predisposition of end states. Examples of such specifications

refer to learning outcomes, competences, employability of graduates, and research. To demonstrate using employability: the observer considers whether the evaluators commend or criticise a condition (for example, the structure of a university's employability survey), a process (for example, regularly including employers in curricular design), or an end state (for example, the actual employment rates of graduates).

The way phase modality has been laid out inadvertently triggers the question of offsets. The critical aspect of this question is twofold. It refers to the image of quality that evaluation practices paint when they resort to the quality of conditions or processes rather than to the quality of results where results should be addressed. But it also refers to quality as possibility, credit or promise, which through accreditation transforms into quality as an official guarantee granted by the overseeing institution, into public recognition of achieved or demonstrated quality.

An offset is a shift in modality resulting from a disconnect between the judgement, specification, actual state of affairs and, where applicable, the ideal. Therefore, phase modality is offset when the phase that is inscribed in the specification is shifted by the evaluator to another phase. The evaluation consequently turns towards a different phase of the concerning state of affairs. An offset in phase modality therefore has to be differentiated from any shift from end state evaluations towards evaluations of conditions and processes.

Anchoring Phase Modality

Research of quality and its assurance has readily touched on the notion of phase modality. It has done so through deducing their systemic building blocks or inducing from the effects of adopted policies rather than from the properties of evaluation practices or their modalities (Lagrosen et al., 2004; Seyfried & Pohlenz, 2018). In systems theory, Luhmann (2008) leans on the input-output model of how systems relatively autonomously from their environment transform inputs into outputs. Such transformations can be considered a matter of machine-like operations, which in the system of higher education touch on education and research, and which in the system of quality relate to a specifications-based programme of converting inputs into positive or negative outputs. Located in the cross-section of both systems is quality assurance. It appears as a control mechanism not for producing quality outputs but for controlling the transformations in order to avoid or reduce deviations from what is expected, and at the same time, as a mechanism for allocating specific values in the name of quality (Harvey & Green, 1993; Luhmann, 2008; Westerheijden, 2007).

Van Kemenade et al. (2008) remind us of past attempts at defining quality related value systems. Through the lens of quality management, the authors share an overview of quality assurance transitioning from controlling the process to continuous improvement, then to commitment, which in the educational sense relates to the transformative aspect of quality, and finally, to breakthrough that rests on flexflow and on opportunistic adaptation of processes. In all developmental stages, values of quality are process-laden and lean on acts of change.

In their seven models of educational quality, Asif and Raouf (2013) mention the resource-input model in which quality is linked with acquiring scarce resources and inputs, and the process model that attaches quality to the issue of how smoothly the internal processes of a higher education institution function. Harvey and Green (1993) pointed out that “quality is relative to ‘processes’ or ‘outcomes’” (p. 9). Their concept of quality as perfection or consistency is focused on consistency of processes and compliance with specifications rather than on essential quality of inputs and outputs. This concept together with that of quality as fitness for purpose to this day prevails in the *European Standards and Guidelines* (ENQA et al., 2015) as well as in the Slovenian *Criteria for Accreditation* (SQAA, 2014). These regulations integrate the concept of quality as fitness for purpose into specifications reducing it to prescribing, managing and processing stakeholder requirements, inclusion, and participation. Hence, the evaluation practices tend to blur the otherwise clear theoretical distinction between the two concepts.

Freitag (1995) claims that normativity in the technological and technocratic society directs towards procedures rather than synthetic values. Bourdieu and Passeron (1990) tie such procedures with measuring the efficiency and productivity of the education system to meet the requirements of economy. While discipline and surveillance are exercised through quality related processes, the latter thus also serve as means of reproducing economic behavioural patterns, values and norms (Biesta, 2010; Cannizzo, 2016; Charlton, 2002; Shore, 2008). Following Foucault’s concept of the technique of power, Cannizzo (2016) continues that agents entangled in processes are reified by performance evaluation and become visible through the documentation of their conduct. It becomes apparent that quality and its assurance pose a problem not only for the process phase but also for that of end states. Biesta (2010) observes how both reproduction and production of knowledge are systematically squeezed into objectively measurable quantities despite severe limitations of such conversion. Evidence-based practice assumes that the ends of professional or scientific action are given and that “the only relevant (professional and research) questions to ask are about the most effective and efficient way of achieving these ends” (Biesta, 2010, p. 35). The thread of this argument also winds around the problem of not measuring what we value, but instead valuing what we measure while neglecting that reification of education and research considerably limits our scope (Biesta, 2010). Biesta’s predicament that is applicable to quality assurance can be traced back to the historical rise of instrumentalised subjective reason (Horkheimer, 2004). Collini (2012), Findlow (2008), Harvey (2009), Rué et al. (2010) and Wittek and Kvernbekk (2011) provide arguments on how quality and its assurance have been unable to overcome the deficits of their positivist approach in converting education and research into quantities and then equating these with quality. With their unsolved ontological, epistemological and methodological problems, quality and its assurance thus foster a breeding ground for offsets in the modalities of external evaluations.

Alvesson (2013) has exposed the force of the image and the surface appearance that determine the behaviour of people and institutions alike. Higher education institutions,

their study programmes and representatives are driven to present themselves positively and pay attention to their appearance rather than to substance and the actual state of affairs. In doing so, they may resort to grandiosity, illusion tricks, exaggerated, pretentious and inflated claims, titles and labels while marginalising the issues of substance, veiling unfavourable appearances of the state of affairs, and possibly stopping short of misleading, of disguising inconvenient facts (Alvesson, 2013).

Grandiosity, more likely reduced to benevolence, seeks its domain in self-evaluation reports which serve as one of the main pieces of evidence in external evaluations. During site visits, evaluators check the information from such reports against the testimonies of interviewees who are also driven to resort to grandiosity and illusion tricks. In doing so, questionable information is occasionally escorted on its way to becoming evidence and the basis for passing judgements. In addition, eventual pseudo-structures may offer themselves either as false signs of quality or quality offset to conditions and processes. Creating organisational goals, appointing committees, participating in quality assurance projects, adopting quality related policies and trends, managerial practices, and continuous institutional reorganisation may be proposed to evaluators as evidence of achieved quality, even excellence. At the same time, specifications that govern external evaluations may steer evaluators towards paying attention to exactly such structures and practices over substance or essence. Although this negative approach to quality assurance is not to be adopted as a rule of thumb, Alvesson (2013) has nevertheless demonstrated that his findings cannot be neglected in the research of evaluation practices. Therefore, the question is, how do evaluators buy into this, and if or how do they perpetuate or even amplify this? By succumbing to appearance, quality and its assurance consequently also foster a breeding ground for offsets in the frequency and criticality of external evaluations.

The immanent quality of the depth of student knowledge, of a diploma thesis, and academic recognition of pedagogical or scientific achievement, such as a great lecture, monograph, patent or discovery, find themselves in the company of adopting managerial or administrative decisions, appointing task forces or focus groups, changing internal regulations, producing self-evaluations, surveying particular stakeholder groups, etc. Harvey (2009) points out that external quality assurance is also considered a “process designed to obscure what has really happened to higher education” (p. 10).

The field of quality in higher education has so far been governed by specifications and guidelines that rather than determining the value and substance of this field’s symbolic capital in its end state, instead focus, on the one hand, on efficiently processing its accumulation – on goal-oriented planning, measuring, documenting, reporting and overseeing –, and on the other hand, on the conditions for its accumulation – on rules themselves, on stakeholder inclusion and on minimum requirements that standardise higher education. These specifications and guidelines then influence the practices, the rituals of quality assurance, and colonise them with the vocabulary of bureaucrats and managers that Bourdieu and Wacquant (2001) termed *newspeak*. Therefore, the

immanent quality of end states is, more than that of conditions and especially processes, left to diverse disciplines, to dispersed external pressures on higher education and to the relativistic eye of a professional beholder. And in practice, this beholder wrestles with the politics and policing of his contractor, the quality assurance agency, with possibly enhanced presentations of what he or she evaluates, with disqualifying interests and tastes elevated by the necessity of stakeholder inclusion, as well as with the disparate academic and economic imperatives cultivating his or her habitus. This sets the richly layered context for interpreting phase modality, its offsets and relation to the criticality of evaluations.

METHOD AND SAMPLE

The research of phase modality proceeds from a system-wide analysis of quality in Slovenian tertiary education and of the properties of the external evaluation practices of SQAA (Širok, 2018). It examines the frequency, criticality, and phase modality of external evaluations according to 32 categorical variables that are derived from the specifications in the *Criteria for Accreditation* (SQAA, 2014). These variables, presented in Table 1, were selected out of 63 (of otherwise 123 in total) with a frequency greater than 20%, meaning that 20% of observed study programmes where evaluators proclaimed a strength, opportunity for improvement or an inconsistency (with regulations) for a corresponding specification were considered. This allowed us to test the sensitivity of evaluations to phase modality in variables with the greatest frequency across all areas that the *Criteria for Accreditation* cover, and therefore with the greatest potential to reflect the impact of quality assurance on higher education. The sample includes 485 study programmes which is 99% of all programme re-accreditations by the SQAA during 2014 and 2017, and 49% of all accredited study programmes in Slovenia according to the *Register of Higher Education Institutions and Study Programmes* in 2017 (Širok, 2018).

Observation of Phase Modality

Observing *frequency* and *criticality* required collecting the proclamations of the three qualitative categories according to individual variables. Interpretation was necessary only in assigning individual judgements to individual variables. Decisions had to be made whether a judgement is too broad or too narrow or whether it corresponds with the variable, while the proclamation of compliance or quality was clear (Širok, 2018).

In *establishing phase modality*, guidelines were introduced to limit the empirical gap. Judgements were interpreted as those of conditions if they referred to conditions, possibilities, motivations, interests, requirements, demands, guarantees and promises for quality or for an end state that has yet to be achieved or demonstrated. (1) Judgements of conditions referred to material, financial, organisational, managerial and intellectual conditions. They were expressed with verbs like: ensure, set-up, appoint, determine, check, consider, start, introduce, support, encourage, prepare, propose, promote, acquire, look or strive for, invite, regulate, include, define. (2) Judgements of processes had to point towards processes,

procedures, practices, action, motion and transformation with the possibility of eventually achieving the quality of an end state or enhancing something. They were to include signifiers such as: to participate, convene, coordinate, repeat, disseminate or inform, document, report, plan, systematise, formalise, monitor, improve, ensure, assure, organise, manage, lead, change, continue, pursue, strengthen, renovate, refresh, enhance, accommodate, adjust, function. (3) Judgements of end states were identified as such if they aimed at something final, terminal, accomplished, achieved, realised and completed – a result or an outcome. Such judgements leaned on signifiers like: a recognition, commendation, award, creation, publication, quotation, graduation and graduation related outcomes such as a diploma, habilitation, completion of a research project or a conference, patent, discovery, employment, promotion, tenure, acquisition.

A proclamation of compliance or quality has the possibility to produce all three phase modalities for each specification and each state of affairs regardless of the underlying specification's relation to an outcome or result. For instance, if a judgement on premises and equipment, which are only conditions for eventual accomplishment in study, teaching or research, refers to a promise of their acquisition, to the process of their renovation, or to them already having served as a quality basis for accomplishing an educational goal, a judgement for either of the variables may be that of a condition, process or end state. Similarly, if a judgement on the functioning of the internal quality assurance system refers to introducing a new quality manual, formalising stakeholder participation in ongoing processes, or to the impact of quality assurance related improvements, the first is that of a condition, the second is that of a process and the third is that of an end state. In summary, a crude approach to interpreting phase modality could be to ask whether evaluators judged the target implied by the specification, an underlying process or a condition for what is implied.

Following the same guidance, the predisposition of phase modality was assigned to each variable by observing the way respective specifications are defined in the valid *Criteria for Accreditation* (SQAA, 2014). Even though premises and equipment are only conditions for eventual accomplishment, they are specified as end states, meaning that the criterium for the re-accreditation of a study programme requires the higher education institution to already have availed appropriate premises and equipment. The specification regarding the scientific, research, professional or artistic work of students, however, is neither defined as a possibility that a higher education institution must provide to students nor as an end state meaning that student achievements such as publications are not expected. Instead, it is defined as a process, as a requirement of ongoing student participation in research (SQAA, 2014). The more evaluators are sensitive to phase modality and careful in applying the specifications, the more external evaluations are likely to be influenced by definitions of specifications. Therefore, the effects of the predisposition of phase modality on evaluation practices were observed as well.

Structure of Collected Data

The acquired database structures the results according to categorical variables as frequencies and total counts of strengths (S), opportunities for improvement (OI), inconsistencies (with regulations) (I) and at the same time of conditions (C), processes (P) and end states (ES). The frequency of judgements is labelled either with the category mentioned (M) or its complementary category not mentioned (NM). All (M) are either (S), (OI) or (I), and at the same time either (C), (P) or (ES), while all (NM) are neither. For comparison, the results for these categories are reduced to two ratios – criticality ratio (CR) and phase modality ratio (PMR). Both are weighted by the frequency of judgements (M). The former is given by the following formula:

$$CR = (S - (OI + I)) * M * 100,$$

and the latter by:

$$PMR = (ES - (C + P)) * M * 100.$$

Variables are assigned the predisposition of phase modality (PPM) ranging from 0 (condition) to 1 (process) and 2 (end state).

Averages and standard deviations are given for all 32 variables and the entire sample of study programmes. Averages are calculated for top and bottom quartiles of observed modalities, and in case of predisposition of phase modality for all three groups of variables. The association between the variables is further explored with Paerson's Chi Square Test to observe the differences between criticality and phase modality, as well as phase modality and its predisposition. In the supporting contingency tables that are stated in case of $p < 0.05$, the observed counts for phase modality are structured into ES and C+P, and the observed counts for criticality are structured into S and OI+I.

RESULTS

The first array of results (Table 1) gives the shares for phase modality, criticality, frequency, as well as both ratios, CR and PMR, for individual variables at the level of the entire sample of evaluated study programmes. It shows that some specifications are more frequently used as a basis for evaluation (see changes in NM) and how criticality and phase modality of evaluations vary. On average, the selected variables had a 50% chance of being connected with a commendation or a recommendation for the evaluated study programme. Although more than half of the evaluations in relative terms referred to end states, more than a third were evaluations of processes or conditions, which is considerable.

Table 1

Variables according to phase modality ratio (PMR) for all 485 study programmes

Variables	PPM	S	OI	I	NM	CR	C	P	ES	PMR
1 Premises for teaching, research or artistic creation	2	59.59%	20.41%	0.00%	20.00%	31.34	4.12%	1.86%	74.02%	54.43
2 Equipment for teaching, research or artistic creation	2	52.78%	24.74%	0.00%	22.47%	21.74	10.52%	5.77%	61.24%	34.85
3 General support to extra-curricular activities of students	2	50.72%	25.77%	0.00%	23.51%	19.08	16.49%	9.07%	50.93%	19.40
4 Library resources	2	31.13%	16.91%	0.00%	51.96%	6.83	3.71%	3.30%	41.03%	16.34
5 Self-evaluation at the higher education institution (HEI) - completion of quality loop	0	8.25%	34.85%	0.21%	56.70%	-11.61	0.82%	2.89%	39.59%	15.53
6 Informing about the findings of self-evaluation	2	7.84%	37.73%	3.71%	50.72%	-16.56	4.95%	5.57%	38.76%	13.92
7 Pedagogical workload of teachers	2	5.57%	35.46%	0.62%	58.35%	-12.71	7.01%	0.82%	33.81%	10.82
8 Participation of important stakeholder groups in internal quality assurance	2	16.91%	41.03%	0.21%	41.86%	-14.15	4.33%	17.32%	36.49%	8.63
9 Delivery of organised contact hours, their extent and distribution	2	14.64%	23.71%	1.24%	60.41%	-4.08	7.22%	2.68%	29.69%	7.84
10 Correspondence of study programme's content with the allocated credits or student workload	2	2.89%	38.56%	2.68%	55.88%	-16.92	7.22%	10.10%	26.80%	4.18
11 Financial resources for teaching, research or artistic creation	2	28.66%	20.82%	0.21%	50.31%	3.79	6.19%	14.85%	28.66%	3.79
12 Functioning of the internal quality assurance system	1	20.62%	10.31%	0.41%	68.66%	3.10	5.98%	4.33%	21.03%	3.36
13 Completion of study programme's structure	2	3.71%	22.27%	0.21%	73.81%	-4.91	5.57%	1.86%	18.76%	2.97
14 Scientific, research, professional or artistic work (SRPAW) of teachers	0	38.76%	14.02%	0.41%	46.80%	12.94	12.37%	11.55%	29.28%	2.85
15 Correspondence of study programme content with programme objectives, anticipated knowledge, competences and learning outcomes	2	16.08%	29.69%	0.82%	53.40%	-6.73	16.91%	3.51%	26.19%	2.69

Variables	PPM	S	OI	I	NM	CR	C	P	ES	PMR
16 Complexity of study (of contents, anticipated knowledge, enrolment criteria, criteria for progression and completion)	2	9.90%	21.44%	0.21%	68.45%	-3.71	8.66%	3.92%	18.97%	2.02
17 Organisation of the HEI	0	19.18%	16.08%	0.00%	64.74%	1.09	7.84%	7.84%	19.59%	1.38
18 Validity of habilitations and their correspondence with the cycle of the study programme	2	11.55%	7.01%	1.86%	79.59%	0.55	4.74%	2.89%	12.78%	1.05
19 Transfer of SRPAW outcomes into teaching	2	19.18%	3.51%	0.00%	77.32%	3.55	0.41%	8.66%	13.61%	1.03
20 Management of the HEI	0	16.29%	5.77%	0.00%	77.94%	2.32	5.77%	3.51%	12.78%	0.77
21 Graduate competences	0	29.90%	12.58%	0.21%	57.32%	7.30	9.28%	11.75%	21.65%	0.26
22 Cooperation with the business sector through pedagogical work	1	29.69%	8.45%	0.00%	61.86%	8.10	4.74%	14.43%	18.97%	-0.08
23 Student participation in HEI management	2	44.12%	28.04%	0.62%	27.22%	11.26	13.81%	23.51%	35.46%	-1.35
24 Scientific or professional cooperation with the business sector	1	37.94%	11.13%	0.00%	50.93%	13.15	10.52%	16.29%	22.27%	-2.23
25 Quality culture	1	17.94%	7.84%	3.51%	70.72%	1.93	12.58%	7.01%	9.69%	-2.90
26 Changes or upgrades to study contents (curriculum, syllabi)	2	35.46%	7.42%	0.00%	57.11%	12.03	7.01%	19.38%	16.49%	-4.24
27 Including students into SRPAW at the HEI	1	43.51%	18.97%	0.41%	37.11%	15.17	17.11%	22.68%	23.09%	-10.50
28 Activity of central organisational units in the field of graduate employability	1	18.76%	60.41%	0.00%	20.82%	-32.98	15.46%	31.55%	32.16%	-11.75
29 Employability or employment of graduates	1	26.80%	34.43%	0.82%	37.94%	-5.25	16.08%	26.60%	19.38%	-14.46
30 Quality of teaching	2	37.32%	19.59%	0.00%	43.09%	10.09	31.13%	10.93%	14.85%	-15.49
31 Student mobility	0	19.38%	62.68%	0.00%	17.94%	-35.53	43.09%	8.87%	30.10%	-17.94
32 SRPAW at the level of HEI or its founder	2	54.43%	21.44%	0.21%	23.92%	24.94	21.44%	31.34%	23.30%	-22.43
Average		25.92%	23.22%	0.58%	50.28%	1.41	10.72%	10.83%	28.17%	3.27
Standard deviation		15.80%	14.48%	0.99%	18.57%	15.10	8.73%	8.75%	14.18%	14.81

Note. Abbreviations: predisposition of phase modality (PPM), strengths (S), opportunities for improvement (OI), inconsistencies (with regulations) (I), not mentioned (NM), criticality ratio (CR), conditions (C), processes (P), end states (ES), phase modality ratio (PMR).

Individual shares for categories of frequency, criticality, and phase modality are similarly affected by how broadly the specifications are defined. Unlike the specifications of content and delivery of study programmes, those of material conditions and student support are less fragmented and consequently exhibit greater frequency. It is also due to this that variables of conditions and support for study, teaching and research (variables 1, 2 and 3) receive greater absolute shares of positive evaluations with a stronger reference to the target implied by the specification. But other variables with highly positive PMR are also essentially more closely connected with inputs or processes (variables 4, 5 and 6) rather than outputs.

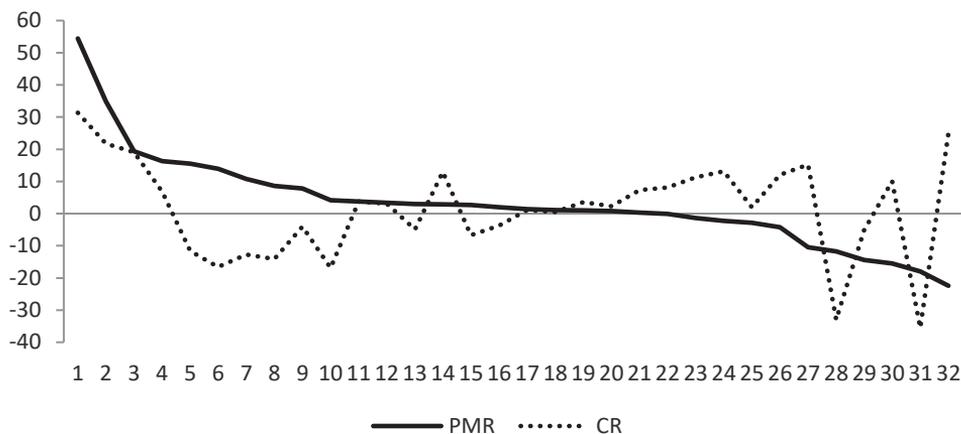
Strong drops in evaluations of end states can be observed in variables more closely connected with outputs. Despite its PPM, scientific, research, professional or artistic work at the institutional level (variable 32) shows a strong offset. Here, evaluators are preoccupied with conditions and especially processes leading to research outcomes. To exemplify, they evaluate support for research, funding, pending research projects and project applications, as well as research related strategic objectives. Rather than assessing the quality or impact of completed research, they only encourage research or emphasize raising awareness of its importance. Quality of teaching (variable 30) with the predisposition of an end state and essentially referring to an end state demonstrates a strong offset towards conditions. Here, evaluations focus on the funding of compulsory teacher training, introducing trending policies in teaching, teaching methods, modes of assessment and supporting technologies, as well as on incentives for efficiency or excellence in teaching, rather than on direct quality of teachers and their work. Evaluations are again strongly process-laden in case of employability or employment of graduates (variable 29, PPM = 1). While mostly critical, they pay attention to monitoring and surveying the employability of graduates as well as to the underlying methodology or to informing about employability (SQAA, n.d.).

Other individual variables, such as student mobility (variable 31, PPM = 0) or activity of central organisational units in the field of graduate employability (variable 28, PPM = 1) behave differently than those previously presented. Both may essentially be deemed as conditions or processes that contribute to eventual outcomes in education, and yet they have a strongly negative PMR. Despite such exceptions, individual results indicate a pattern that evaluators resort to offsets or shifts from end states in variables that are essentially more closely connected with end states (to some extent variables 14 and 21, but especially variables 27, 29, 30 and 32), whereas greater shares of end state evaluations can be found in variables at the top of Table 1 that essentially refer to conditions and processes – to prerequisites for relevant outcomes. Apart from this pattern and the influence of PPM which will be discussed below, no other distinctive property of individual specifications could be identified that influences the phase modality of evaluations.

Leaving immanent properties of variables aside, there seems to be no obvious relation between phase modality and criticality on the level of individual variables. This can be observed in the visualisation of the relation between PMR and CR for variables 1 through 32 in Figure 1.

Figure 1

PMR against CR for 32 variables and the entire sample of study programmes



Note. Abbreviations: criticality ratio (CR), phase modality ratio (PMR).

Based on the amount of scatter in Figure 1, it seems that phase modality and criticality behave independently and differently. Looking past the individual variables, the changes in phase modality and its relation to criticality were examined for variables grouped according to results for CR and PMR. When comparing the averages for categories of phase modality in the quartile of variables least critically evaluated with those in the quartile of most critically evaluated variables, individual phases including PMR differ little. Similarly, the comparison of averages for strengths, opportunities for improvement or inconsistencies in the quartile of variables with greatest shares of end state evaluations to that of variables with least shares of end state evaluations produces hardly any difference. Table 2 suggests that great changes in the phase modality of evaluations result in smaller changes in their criticality, which remains close to average values for all 32 variables and vice versa. However slightly, the more critical the evaluations are, the more they deviate from end states.

Table 2

Averages for variables grouped according to CR and PMR

Quartiles	S	OI	I	NM	CR	C	P	ES	PMR
CR - Upper quartile (aver. top 8)	46.65%	17.99%	0.13%	35.23%	18.80	12.45%	14.74%	37.58%	9.02
CR - Lower quartile (aver. bottom 8)	11.96%	42.55%	1.03%	44.46%	-18.40	12.47%	10.08%	32.99%	3.26
PMR - Upper quartile (aver. top 8)	29.10%	29.61%	0.59%	40.70%	3.00	6.49%	5.82%	46.98%	21.74
PMR - Lower quartile (aver. bottom 8)	31.70%	29.10%	0.62%	38.58%	-1.20	20.49%	19.79%	21.13%	-12.46

Note. Abbreviations: strengths (S), opportunities for improvement (OI), inconsistencies (with regulations) (I), not mentioned (NM), criticality ratio (CR), conditions (C), processes (P), end states (ES), phase modality ratio (PMR).

The Chi Squared Test (2x2) shows that there indeed is significant association between the criticality and phase modality of evaluations $\chi^2(1) = 6845, p < 0.05$. The observed counts of strengths and the sum of counts of opportunities for improvement and inconsistencies on the one hand, and the observed counts of end states and the sum of counts of conditions and processes on the other, produce the following contingency table:

Table 3
Contingency table – phase modality vs. criticality

Observed	ES	C+P	Total	Expected	ES	C+P
S	2477	1546	4023		2279.2	1743.8
OI+I	1895	1799	3694		2092.8	1601.2
Total	4372	3345	7717			

Note. Abbreviations: strengths (S), opportunities for improvement (OI), inconsistencies (with regulations) (I), conditions (C), processes (P), end states (ES), phase modality ratio (PMR).

When compared to expected values, an increase in the criticality of evaluations produces a statistically significant decrease in end state evaluations. The size of this change in phase modality, which can be derived from the ratio between the observed and expected counts of end state evaluations, amounts to 9%. It is similar to the size of the excess of observed counts of end states on the positive end of evaluations. Despite the scatter in Figure 1 and the small compared differences in Table 2, this general association cannot be neglected. Evaluators to some extent tend not to pass critical judgements on end states and positive judgements on conditions and processes.

Proceeding from the above contingency table (Table 3) while returning to the results for individual variables, overly positive appearances of quality cannot be confirmed. Although a combination of positive evaluations and strong offsets in phase modality is a condition for eventual false signs of quality or succumbing to appearance, few variables, such as the above presented variables 30 and 32, have highly negative PMRs and highly positive CRs. Instead, the results for several variables show that evaluators also prefer to attach positive judgements to specifications that regardless of the phase modality of evaluations essentially refer to conditions or processes for education and research rather than end states. Such are the results for variables of material conditions (variables 1, 2, 3 and 4) or stakeholder inclusion which is also aimed at the process of serving society (variables 22, 23 and 24). Evaluators tend to focus on end states when outputs are not constitutive elements of specifications. In the critical spectrum of evaluations with higher positive PMRs, specifications essentially unrelated to end states crop up again. Such are the results for variables 5, 6 and 8, which refer to internal quality assurance processes and stakeholder participation therein.

Next is the question of the influence of the predisposition of phase modality (PPM) on the phase modality of evaluations. Averages were calculated for groups of variables with

varying PPM for the entire sample of evaluated study programmes (Table 4). 6 variables are predisposed as conditions (PPM = 0), 7 as processes (PPM = 1) and 19 as end states (PPM = 2).

Table 4
Averages for variables grouped according to PPM

Groups of variables	S	OI	I	NM	CR	C	P	ES	PMR
PPM - All 0 (aver. 6 variables)	21.96%	24.33%	0.14%	53.57%	-3.91	13.20%	7.73%	25.50%	0.48
PPM - All 1 (aver. 7 variables)	27.89%	21.65%	0.74%	49.72%	0.46	11.78%	17.56%	20.94%	-5.51
PPM - All 2 (aver. 19 variables)	26.45%	23.45%	0.66%	49.44%	3.44	9.55%	9.33%	31.68%	7.39

Note. Abbreviations: predisposition of phase modality (PPM), strengths (S), opportunities for improvement (OI), inconsistencies (with regulations) (I), not mentioned (NM), criticality ratio (CR), conditions (C), processes (P), end states (ES), phase modality ratio (PMR).

A comparison of averages between the three groups of variables reveals that variables with PPM = 2 have considerably greater shares of end state evaluations. Phase modality in those with PPM = 1 is predominantly shifted towards evaluations of processes and in those with PPM = 0 mostly towards conditions. The average PMR for all 32 variables (see Table 1) is also sizeably smaller than the average PMR for variables with PPM = 2 and larger than the average PMR for variables with PPM = 1. However, the averages for conditions, processes and end states in variables with PPM = 2 hardly differ from the averages of these categories in all 32 variables. The results of the Chi Squared Test (3x2) nevertheless confirm that there is significant association between PPM and the phase modality of evaluations $\chi^2(2) = 51242, p < 0.05$.

Table 5
Contingency table – phase modality vs. PPM

Observed	ES	C+P	Total	Expected	ES	C+P
PPM - All 0 (6 variables)	742	609	1351		765.4	585.6
PPM - All 1 (7 variables)	711	996	1707		967.1	739.9
PPM - All 2 (19 variables)	2919	1740	4659		2639.5	2019.5
Total	4372	3345	7717			

Note. Abbreviations: predisposition of phase modality (PPM), conditions (C), processes (P), end states (ES).

Table 5 reaffirms that with regard to expected values, variables with PPM = 2 are more likely to produce evaluations of end states. It is therefore important to consider phase modality when defining specifications. The less the definition of standards of quality targets end states, the more likely evaluations will focus on conditions and especially processes.

DISCUSSION

This analysis of external evaluations reveals considerable offsets in phase modality which are evident both in individual variables and in averages for all variables at the level of the entire sample of the observed study programmes. Offsets towards processes and conditions can be tied with the process character of quality assurance as it is also evident in its prevailing definitions. Quality is readily “in danger of being defined in terms of the existence of suitable mechanisms and procedures, but in and of themselves they tell us nothing about the quality of the results” (Wittek & Kvernbekk, 2011, p. 674). Let us remember that preoccupation with the operationalisation of quality and the processes of its assurance has been observed on numerous occasions (Charlton, 2002; Findlow, 2008; Harvey, 2009; Lorenz, 2012; Shore, 2008). With regard to phase modality, processes are central both in the prevailing concepts of quality as well as in quality assurance mechanisms, while conditions and end states tend to get limited to processes and reduced to matter that is processed. Processes may not only assimilate conditions and end states but may also be offset to other phases. For instance, the continuous process of quality management at a higher education institution may through proclamation of compliance with the respective specification become a symbol of achieved quality.

Evaluations that are based on several outcome related specifications tend to aim at conditions and processes while unexpectedly greater shares of end state evaluations are likely to arise from specifications that have essentially to do with conditions and processes. Thus, conditions and processes are likely to manifest themselves as ends of quality assurance, as a sign of good or bad quality. Although to a lesser extent, offsets are also evident in variables with the predisposition of end states. On the one hand, this points to the tendency of evaluators to relate to quality through processes and techniques rather than to identify what in the observed education and research is or is not good or what in terms of quality has happened to it. On the other hand, however, it is apparent that the kind of phase modality that is inscribed in the standards of quality has a significant chance to surface in evaluations.

Before expanding on the results of how critical the evaluations are, one should notice that in total averages the shares of strengths are well balanced with the sums of shares of opportunities for improvement and inconsistencies. This is because SQAA's evaluators were expected to produce critically balanced and sufficiently motivating assessments.

Considering also the criticality of evaluations, positive evaluations are in general not characterised by strong offsets in phase modality or shifts from end states. Instead of smuggling praise through evaluations of conditions and processes, evaluators actually tend to reserve the latter for criticism. When external quality assurance intervenes, it prefers to intervene in processes and conditions. External evaluations therefore do not so much catalyse grandiosity, illusion tricks or pseudo-structures that might have resulted from the higher institution's presentation of the actual state of affairs as much as they divert

attention away from end states. The quality that evaluators proclaim to some extent ends up being not an inflated but a skewed image of what may be considered good education and research or what has happened to both. In response to such external evaluations, higher education institutions are then more likely to reply with action plans saturated with administrative and managerial measures as a technique of gradual and constant improvement, rather than indirectly assuring desired or required end states. This then is the character of the impact of external evaluations on higher education.

Arguably, evaluators might avoid criticism of end states because it requires greater professional exposure, exactness and confronting disparate economic and academic questions of relevance, value and achievement. And since specifications as well as quality assurance processes are framed by process-laden values, evaluation of end states according to ideals is likely to give way to administrative and managerial issues of transparency, stakeholder inclusion, efficiency and effectiveness.

The tendency to avoid critically evaluating end states might lastly be traced back to eventual backlash from those who are being evaluated. In the case of critical evaluations, institutions can expect sanctions or are at least faced with having to act by adopting corrective measures. But remedying insufficiencies in end states may prove far more resource or time consuming than adjusting organisational and administrative processes or improving less immediate conditions. Lack of good research, teachers, students and eventually employable or accomplished and well-educated graduates is something a higher education institution may not be able to overcome, regardless of how well it tunes its internal quality assurance system.

CONCLUSION

This research offers specific insight into how quality assurance impacts higher education in practice. Across the spectrum of quality and its standards, evaluators do not consistently focus their evaluations on end states where they should do so. While there is statistically significant association between phase modality and criticality, the increase in the criticality of evaluations does not produce a strong decrease in the evaluations of end states. Nevertheless, the more evaluators focus on conditions and processes, the more their evaluations are critical. Significant association was also found between phase modality and its predisposition. It is therefore important to consider phase modality when defining specifications that govern quality in higher education. The presented theoretical framework and research could be either further developed to other modalities of external evaluations or they could cover other evaluation practices, for instance, external evaluations of higher education institutions, in order to better understand the impact of quality assurance in practice and to raise critical awareness of this impact when drafting the standards of quality.

REFERENCES

- Alvesson, M. (2013). *The Triumph of Emptiness: Consumption, higher education and work organization*. Oxford University Press. <https://doi.org/10.1093/oso/9780199660940.001.0001>
- Asif, M., & Raouf, A. (2013). Setting the course for quality assurance in higher education. *Quality and Quantity*, 47(4), 2009–2024. <https://doi.org/10.1007/s11135-011-9639-2>
- Ben-Gal, I. E., & Dror, S. (2016). Taguchi method for off-line quality control. *Wiley StatsRef: Statistics Reference Online*. <https://doi.org/10.1002/9781118445112.stat03122.pub2>
- Biesta, J. J. G. (2010). *Good Education in an Age of Measurement: Ethics, politics, democracy*. Paradigm Publishers.
- Bornmann, L., Mittag, S., & Daniel, H.-D. (2006). Quality assurance in higher education – meta-evaluation of multi-stage evaluation procedures in Germany. *Higher Education*, 52(4), 687–709. <https://doi.org/10.1007/s10734-004-8306-0>
- Bourdieu, P., & Passeron, J.-C. (1990). *Reproduction in education, society and culture* (2nd ed.). Sage.
- Bourdieu, P., & Wacquant, L. (2001). NewLiberalSpeak: Notes on the new planetary vulgate. *Radical Philosophy*, 105, 2–5.
- Cannizzo, F. (2016). The transformation of academic ideals: An Australian analysis. *Higher Education Research & Development*, 35(5), 881–894. <https://doi.org/10.1080/07294360.2016.1138454>
- Charlton, B. G. (2002). Audit, accountability, quality, and all that: The growth of managerial technologies in UK universities. In S. Prickett & P. Erskine-Hill (Eds.), *Education! Education! Education! Managerial ethics and the law of unintended consequences* (pp. 13–32). Imprint Academic.
- Collini, S. (2012). *What are Universities for?* Penguin Books.
- European Association for Quality Assurance in Higher Education, European Students' Union, European University Association, & European Association of Institutions in Higher Education. (2015). *Standards and guidelines for quality assurance in the European higher education area*. https://www.enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf
- Findlow, S. (2008). Accountability and innovation in higher education: A disabling tension. *Studies in Higher Education*, 33(3), 313–329. <https://doi.org/10.1080/030757070802049285>
- Freitag, M. (1995). *Le naufrage de l'Université et autres essais d'épistémologie politique*. MAUSS.
- Harvey, L. (2009). *Deconstructing Quality Culture*. <https://www.qualityresearchinternational.com/Harvey%20papers/Deconstructing%20quality%20culture%20EAIR%20Vilnius%202009.pdf>
- Harvey, L., & Green, D. (1993). Defining quality. *Assessment & Evaluation in Higher Education*, 18(1), 9–34.
- Horkheimer, M. (2004). *Eclipse of Reason*. Continuum.
- Lagrosen, S., Seyyed-Hashemi, R., & Leitner, M. (2004). Examination of the dimensions of quality in higher education. *Quality Assurance in Education*, 12(2), 61–69.
- Lorenz, C. (2012). If you're so smart, why are you under surveillance? Universities, neoliberalism, and new public management. *Critical Inquiry*, 38(3), 599–629. <https://doi.org/10.1086/664553>
- Luhmann, N. (2008). *Einführung in die Systemtheorie* (4th ed.). Carl-Auer-Systeme Verlag.
- Rué, J., Amador, M., Gené, J., Rambla, F. X., Pividori, C., Pividori, I., Torres-hostench, O., Bosco, A., Armengol, J., & Font, A. (2010). Towards an understanding of quality in higher education: the ELQ/AQA08 model as an evaluation tool. *Quality in Higher Education*, 16(3), 285–295. <https://doi.org/10.1080/13538322.2010.506723>
- Seyfried, M., & Pohlenz, P. (2018). Assessing quality assurance in higher education: Quality managers' perceptions of effectiveness. *European Journal of Higher Education*, 8(3), 258–271. <https://doi.org/10.1080/21568235.2018.1474777>

- Shore, C. (2008). Audit culture and illiberal governance: Universities and the politics of accountability. *Anthropological Theory*, 8(3), 278–298. <https://doi.org/10.1177/1463499608093815>
- Slovenian Quality Assurance Agency. (n.d.). *Information on higher education institutions*. <https://www.nakvis.si/analize-in-publikacije/porocila-strokovnjakov-in-odlocbe/>
- Slovenian Quality Assurance Agency. (2014). *Criteria for the accreditation and external evaluation of higher education institutions and study programmes*. <https://web.archive.org/web/20160308095211/http://test.nakvis.si/en-GB/Content/GetFile/561>
- Širok, J. (2018). *Quality in the Slovenian higher education and higher vocational education area in the period from 2014 to 2017*. SQAA. <https://www.nakvis.si/wp-content/uploads/2018/12/Quality-in-Slovenian-HE-2014-2017-Final.pdf>
- Tavares, O., Sin, C., & Amaral, A. (2016). Internal quality assurance systems in Portugal: What their strengths and weaknesses reveal. *Assessment and Evaluation in Higher Education*, 41(7), 1049–1064. <https://doi.org/10.1080/02602938.2015.1064515>
- Thareja, P. (2009). ‘The essence of quality’: Delineating a poetical treatment. *Quality World*, 6(12). <https://ssrn.com/abstract=1522424>
- Van Kemenade, E., Pupius, M., & Hardjono T. W. (2008). More value to defining quality. *Quality in Higher Education*, 14(2), 175–185. <https://doi.org/10.1080/13538320802278461>
- Westerheijden, D. F. (2007). States and Europe and quality of higher education. In D. F. Westerheijden, B. Stensaker, & M. J. Rosa (Eds.), *Quality Assurance in Higher Education: Trends in Regulation, Translation and Transformation* (pp. 73–95). Springer.
- Wittek, L., & Kvernbekk, T. (2011). On the problems of asking for a definition of quality in education. *Scandinavian Journal of Education Research*, 55(6), 671–684. <https://doi.org/10.1080/00313831.2011.594618>
- Wittgenstein, L. (1965). A lecture on ethics. *The Philosophical Review*, 74(1), 3–12.