

Waikato Journal of Education

ISSN 2382-0373

Website: <https://wje.org.nz>



Volume 26, Issue 2, 2021

The functions of NZQA moderation: Academic leaders' perceptions
Anna H. Williams, Michael B. Johnston & Robin Averill

Editors: Kerry Earl & David Taufui Mikato Fa'ava

To cite this article: Williams, A. H. Johnston, M. B., & Averill, R. (2021). The functions of NZQA moderation: Academic leaders' perceptions. *Waikato Journal of Education*, 26(2), 65–89. <https://doi.org/10.15663/wje.v26i2.738>

To link to this volume: <https://doi.org/10.15663/wje.v26i2>

Copyright of articles

Authors retain copyright of their publications.

Articles are subject to the Creative commons license: <https://creativecommons.org/licenses/by-nc-sa/3.0/legalcode>

Summary of the Creative Commons license.

Author and users are free to

Share—copy and redistribute the material in any medium or format

Adapt—remix, transform, and build upon the material

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms

Attribution—You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use

Non-Commercial—You may not use the material for commercial purposes

ShareAlike—If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Open Access Policy

This journal provides immediate open access to its content on the principle that making research freely available to the public supports a greater global exchange of knowledge.



The functions of NZQA moderation: Academic leaders' perceptions

Anna H. Williams, Michael B. Johnston & Robin Averill

Victoria University of Wellington
New Zealand

Abstract

In New Zealand, education organisations that internally assess against nationally-set standards must engage in national moderation—such as that conducted by the New Zealand Qualifications Authority (NZQA)—to quality assure those assessments. How policies are enacted by organisations is influenced by multiple factors and enactment may diverge from intent. One such factor is the people involved in policy work within organisations; however, little research has been done into their perspectives. In this mixed-methods study, we collected perception data from a sample of 215 academic leaders from three organisation types via an online survey. We explored their understandings of the functions of NZQA moderation and differences in views according to organisation type. We found that leaders perceive NZQA moderation to function in the areas of assessment quality, professional learning opportunities, organisational quality assurance and maintaining public confidence. Leaders from different organisation types mostly shared similar views in terms of the functions; however, those from private training enterprises considered NZQA moderation to place stronger emphasis on organisational quality assurance and to hold broader improvement functions than did those from other organisations. These findings are useful for informing the policy work of education organisations (and making better use of the opportunities afforded by moderation) and for NZQA, including to guide suitable framing of moderation communications to enhance alignment between policy intent and policy enactment. Such uses remain relevant in New Zealand's current climate of educational reform, irrespective of any changes made to NZQA moderation policy.

Keywords

Educational assessment; moderation; quality assurance; policy enactment; perceptions; context



Introduction

Effective implementation of moderation policy for high-stakes internal assessment is necessary to ensure the trustworthiness and credibility of the qualifications towards which those assessments contribute. This is particularly so for national moderation systems. A primary purpose of moderation is to ensure acceptable comparability of assessment judgement between organisations (Klenowski & Wyatt-Smith, 2014) and to support public confidence in assessments and in the qualifications they support (Controller and Auditor-General, 2012). The ways in which educational policies—including moderation policies—are put into practice does not always reflect policy intent (Ball et al., 2012); policy enactment is mediated by multiple factors, including the perspectives of those involved in the enactment.

In an investigation of the utilisation by New Zealand (NZ) education organisations of the national moderation conducted by the New Zealand Qualifications Authority (NZQA), we probed the beliefs of the people in schools and tertiary education organisations (TEOs) who were actively involved in the implementation of policy. While most secondary teachers and principals consider NZQA moderation feedback to be of *use* or *value* (Controller and Auditor-General, 2012; Hipkins, 2013), no studies to date have explored perceptions of those in schools or TEOs about the specific *functions* of NZQA moderation. We sought to address this research gap by exploring the following questions:

1. What do academic leaders in Institutes of Technology and Polytechnics (ITPs), registered Private Training Enterprises (PTEs), and state and state-integrated schools (schools) perceive the functions of NZQA moderation to be, and to what extent are the different dimensions of these perceptions associated?
2. To what extent are there differences between the three organisation types regarding academic leaders' perceptions of the functions of NZQA moderation, and regarding the associations between the different dimensions of those perceptions?

This paper reports on a subset of findings from a larger (doctoral) study that explored NZ academic leaders' perceptions (Williams, 2019). Although this study predates the current NZ education system reforms (Reforms of Tomorrow's Schools, and Vocational Education, and review of National Certificates of Educational Achievement—NCEA; more on these later), the findings presented in this paper remain relevant, irrespective of whether these reforms result in substantive changes to NZQA moderation policy or approach.

Background

Internal assessment

Educational assessments are used to evaluate and assist students' progress towards intended learning outcomes; assessment, teaching and learning are thus interrelated (Black & Wiliam, 2018). Internal assessments are those in which judgements of students' learning are made *within* education organisations, including *during* courses. In some cases, responsibility for the development of assessment instruments is also devolved to individual organisations. In contrast, external assessments are usually run at the conclusion of a course, and responsibility for assessment instruments and judgements sit with a super-ordinate authority (see Crooks, 2011).

This article is concerned with internal assessments used for summative purposes, including credentialing. Such assessments contribute to the award of nationally-registered qualifications (e.g., the NCEA) and must be quality assured to ensure their rigour and trustworthiness and, by extension, the credibility of the qualifications to which they contribute (Crisp, 2017; Crooks, 2011; Hipkins et al., 2016; Newton & Shaw, 2014).

Moderation of internal assessment is usually conducted by a subject matter expert ('expert moderation'), by a group of assessors ('social moderation'), or by statistical comparison of internal assessment results with historical results or with external assessments ('statistical moderation') (Hipkins et al., 2016; Klenowski & Wyatt-Smith, 2014). Moderation can be conducted within or between organisations ('internal moderation'), or by quality assurance bodies such as NZQA ('national moderation'; see Crisp, 2017 and Hipkins et al., 2016).

Functions of moderation

Internationally, the functions of moderation are widely recognised as serving accountability and improving assessment practices (Grant, 2012; Klenowski & Wyatt-Smith, 2014).

Accountability, quality control and public confidence are closely linked. Quality control—checking the integrity and trustworthiness of internal assessments—ensures that organisations are accountable for utilising sound assessment practices and that assessment results and qualifications they support are credible (Broadbent, 2018; Crisp, 2017; Hipkins et al., 2016). Through these functions, moderation enables the public to retain confidence in education and qualifications systems (Controller and Auditor-General, 2012; Misko, 2015).

In terms of improving assessment practices, moderation helps to promulgate common understandings of assessment requirements through communities of teachers and assessors (Crisp, 2017) and can also provide learning opportunities for organisations and individuals (Bloxham et al., 2016; Broadbent, 2018; Crimmins et al., 2016). Such learning opportunities may relate not only to the assessment processes themselves but also to the wider teaching and curriculum processes that the assessment supports (Bloxham et al., 2016; Grant, 2012; Klenowski & Wyatt-Smith, 2014; Misko, 2015).

Moderation in NZ

In the NZ context, moderation of internal assessment is usually conducted using expert and social approaches. Education organisations that internally assess against nationally-set standards must engage in internal moderation and in national moderation with the relevant quality assurance body with responsibility for those standards—for example, NZQA (NZQA, 2020a). The present work is concerned with NZQA moderation, comprising expert moderation of internal assessments.

NZQA moderation

NZQA conducts three national moderation systems: a 'national systems check', a 'school check' and a 'TEO check'. The 'national systems check' involves a representative, random sample of assessments drawn from schools across the country, from which a publicly-reported national (moderator-teacher) agreement rate for the secondary sector is generated. The 'school check' and 'TEO check' systems involve samples for assessment standards purposively selected by NZQA from individual organisations (NZQA, 2019, 2020b).

All three systems involve moderators evaluating samples of assessed student work, and the associated assessor judgements, against relevant assessment standard(s). In some cases, the assessment instruments are also evaluated. In the 'school check' and 'TEO check' systems, written feedback is usually provided to the education organisation (NZQA, 2019, 2020b).

Policy enactment

Policy implementation is complex and does not always reflect the policy intent (Ball et al., 2012). Mediating factors in the enactment process include the people involved in the policy work within each organisation, contextual factors and the nature of the policy (Ball et al., 2012; Dalby & Noyes 2018; Hardy 2015; Keddie 2019; Maguire et al., 2015; Sheikh & Bagley 2018; Spillane et al., 2002). These considerations apply just as much to moderation policy as they do to any other policy domain.

Mediating factor: The people involved

The people actively involved in the interpretation and translation of policies for a particular organisation—‘policy narrators’, as Ball and colleagues call them—strongly mediate policy enactment (Ball et al., 2012; Maguire et al., 2015; Sheikh & Bagley 2018; Spillane et al., 2002). In relation to the present study, the policy narrators are likely to be academic leaders with responsibility for moderation.

A policy narrator must interpret what a policy means for their organisation before they can determine how to apply it. These processes of interpretation and translation for a specific organisational context are influenced by the narrator’s existing knowledge and frame of reference—their attitudes, beliefs and understandings (Spillane et al., 2002) and also by their affective responses during the policy enactment process (Maguire et al., 2015; Sheikh & Bagley 2018). People tend to notice ideas that are similar to what they already know, and narrators have idiosyncratic understandings of policy requirements, resulting in differences between narrators in what they notice in—and take from—a policy (whether existing or new).

Mediating factor: Context

In addition to the process of policy narration, both the external and internal contexts within which an organisation operates play mediating roles in policy enactment (Ball et al., 2012; Dalby & Noyes 2018; Hardy 2015; Keddie 2019). The decentralised nature of the NZ education system—for example, most education organisations are self-governing, and NCEA is designed to enable choice in the assessment programmes for courses (Ministry of Education, n.d., 2020)—optimises opportunities for diverse educational offerings and organisational priorities and practices but also increases the risk of substantial deviation from the intent of policy makers. In future, such opportunities for diversity (and risk of divergence) may lessen somewhat, due to the current Reforms of Tomorrow’s Schools and of Vocational Education (NZ Government, n.d., 2019b) and review of NCEA (NZ Government, 2019a). Of particular relevance here are aspects of the reforms that will introduce more centralisation to schools and ITPs: intended changes in the governance and administration of schools to generate more networking and interdependence between schools (NZ Government, 2019b), and of ITPs, all of which have become subsidiaries to the newly-established New Zealand Institute of Skills and Technology, *Te Pūkenga* (NZ Government, n.d.). The changes associated with the NCEA review will primarily affect schools (as opposed to TEOs); for example, the replacement of existing achievement standards with fewer standards of larger credit values (NZ Government, 2019a), will reduce the choices available in terms of assessment.

Other contextual factors that play a mediating role in policy enactment include the regulatory and quality assurance regimes within which organisations operate; these differ between the secondary and tertiary sectors in NZ. For example, the Education Review Office (ERO) is the agency responsible for evaluating the educational quality of schools (ERO, 2020), whereas NZQA is responsible for evaluating the quality of assessment in schools as well as both educational and assessment quality in ITPs and PTEs (NZQA, n.d.b, 2020c). At the time of data collection, NZQA used different moderation systems for each

sector, and—for schools but not TEOs—conducted review audits to assure assessment quality (NZQA, n.d.a, 2014a). To date, these differences have largely continued, although since then, NZQA has had an increased focus on compliance and monitoring across the tertiary sector and of individual TEOs (e.g., NZQA, 2017), and increased the demand for accountability from TEOs with ongoing compliance and performance issues (e.g., Gerritsen, 2017).

Mediating factor: The policy itself

The nature of a policy—including the degree of freedom afforded to organisations in its enactment—also plays a mediating role in how it is enacted (Ball et al., 2012; Maguire et al., 2015). NZQA moderation policy differed by sector at the time of data collection and, for the most part, continues to differ. Both school-sector moderation systems had tightly prescribed requirements for sample selection (NZQA, 2014a), while in the TEO-check moderation system, organisations had freedom in the sample-selection processes used (NZQA, 2014b).^{vi}

It seems likely that the quasi-autonomous nature of internal assessment and moderation, as well as the different moderation policies applied to schools and TEOs, contributes to substantial intra- and cross-sector differences in policy enactment. To shed light on this, in the present study we sought to ascertain whether there were observable differences in the perceptions of leaders in respect of the functions of moderation, according to organisation type and sector.

Methods

We used a pragmatic mixed methods research design incorporating quantitative and qualitative approaches to compare participants' perceptions across organisation types (Teddlie & Tashakkori, 2009). The study was conducted with ethics approval (reference: 0000020332) from the Victoria University of Wellington Human Ethics Committee.

The target population comprised academic leaders with responsibility for moderation in NZ ITPs, registered PTEs, and state schools that include senior secondary levels. Using a census approach (Creswell, 2012), we invited—via email to the publicly available contact address for each organisation—everyone in the target population to participate in an online survey. Where the email recipient was not a member of the target population, we relied on our invitation being forwarded to the relevant person.

The resulting sample comprised 215 leaders, 21 percent of the overall target population (Table 1). We were unable to ascertain how representative the sample was because not all respondents provided demographic data.

Table 1. Target Population and Sample Details

	TP	<i>n</i>	% TP	% total <i>n</i>
Total	1032	215		
ITPs	18	13	72	6
PTEs	575	128	22	60
Schools	439	74	17	34
Notes: TP, target population; <i>n</i> , sample				

We used a bespoke survey instrument for data collection (Williams, 2019) addressing the current emphasis of NZQA moderation through 26 closed-response items. Each had a four-point rating scale with response options of *No-*, *Low-*, *Medium-* or *High emphasis* to allow response differentiation (Cohen et al., 2011). The survey also contained four open-field items, interspersed with those aforementioned, which solicited other functions of NZQA moderation not covered by the closed-response items. Data were collected over nine weeks, May–July 2014.

Data analysis

We first conducted quantitative analyses, with subsequent qualitative analyses informed by the quantitative findings (Williams, 2019). We then brought the quantitative and qualitative findings together to identify the perceived functions of moderation.

Quantitative approaches

To ascertain the broad functional domains of NZQA moderation as perceived by our participants, we ran exploratory factor analyses (EFAs) on the ordinal quantitative data collected via closed-response items. These analyses enabled us to identify the underlying latent structure—factors—and theorise the latent constructs underpinning the variables (Tabachnick & Fidell, 2013). We used a maximum likelihood extraction method and the oblique rotation procedure, Oblimin with Kaiser Normalisation. We then used a separate principal component analyses with varimax rotation on the response data for the set of items associated with each factor, to check that its structure was sufficiently unidimensional to justify subsequent calibration using a single-parameter graded item-response model (Samejima, 1969; see also Hambleton et al., 1991).

The calibration process involved fitting the set of ordinal responses to the items associated with each factor to an interval measurement scale; in this case model parameters were estimated using a maximum-likelihood iterative algorithm (for computational details, see Hambleton et al., 1991). Because the resulting measurement variables were equal interval scales, parametric statistical tests could validly be used to analyse them. We disaggregated these variables by organisation type and calculated the mean scale location of each type on each measurement scale. One-way analyses of variance were used to ascertain whether there were statistically significant differences among those mean locations (Tabachnick & Fidell, 2013). Where significant differences were found, we conducted independent sample *t*-tests on the three possible pairings of organisation types to identify between which organisation types there were significant differences (Manly, 1986).^{vii}

To ascertain associations between the different dimensions of leaders' perceptions of the functions of moderation and the extent of any variation in these associations according to organisation type, we calculated Pearson's correlation coefficients (2-tailed) between measurement scales for the aggregated data and for individual organisation types (Cohen et al., 2011).

Qualitative approaches

Qualitative data were collected via open-field items. Our qualitative analysis process was iterative, cyclical and systematic (Newby, 2010). We initially used a predetermined coding framework based on the findings from the EFAs in a deductive approach before adopting a semi-inductive approach and modifying several predetermined codes to better 'fit' some data (Teddlie & Tashakkori, 2009). Throughout the process we used coding guidelines to support consistent and transparent interpretation (Mertens, 2010). We maintained matrices of the codes assigned to individual respondents' data to support process monitoring and review (Newby, 2010).

Results

Quantitative analyses

The EFA identified five factors reflecting participants' views of the current emphasis of NZQA moderation: *Checking internal assessment quality*, *Improving internal assessment quality*, *Providing opportunities to learn from NZQA moderation reports*, *Organisational quality assurance*, and *Maintaining public confidence*. Tables 2–6 show the segment of the pattern matrix associated with each of these factors. For each factor, the associated items are conceptually coherent. For example, the three items associated with the *Checking internal assessment quality* factor address checking different aspects of the quality of internal assessments: that the assessment materials and practices meet the Standard requirements, as do the assessor judgements (refer Table 2).

Table 2. Pattern Matrix of Items Associated with the Checking Internal Assessment Quality Factor

		Percentage of variance (Extraction Sums of Squared Loadings)
		5.2
		Factor
Items		<i>Checking internal assessment quality</i>
Q21.3	To check that assessor judgements are accurate against the requirements of the Standard	.764
Q21.2	To check that assessment materials are fit for purpose in terms of the requirements of the Standard	.617
Q21.1	To check whether or not the internal assessments of an organisation meet the requirements of the Standards, as set nationally	.493
Note: coefficients less than 0.30 suppressed.		

Table 3. Pattern Matrix of Items Associated with the Improving Internal Assessment Quality Factor

		Percentage of variance (Extraction Sums of Squared Loadings)		
		7.2	8.8	5.2
		Factor		
Items		<i>Improving internal assessment quality</i>	<i>Organisational quality assurance</i>	<i>Checking internal assessment quality</i>
Q27.3	To assist in the development of a shared understanding among relevant people	.784		

Q27.2	To stimulate professional and collegial conversation	.744		
Q27.4	To build the confidence of assessment designers or assessors	.735		
Q27.1	To give guidance regarding how to interpret the requirements of a Standard	.708		
Q26.6	To give feedback about how to improve assessor judgements	.605		
Q26.4	To give feedback about how to amend the assessment materials in order to better assess achievement against a Standard	.524		
Q26.5	To give feedback about the quality of the assessor judgements	.509		.380
Q29.2	To give feedback about whether an assessment supports teaching and learning	.388	.351	
Note: coefficients less than 0.30 suppressed.				

Table 4. Pattern Matrix of Items Associated with the Providing Opportunities to Learn from NZQA Moderation Reports Factor

		Percentage of variance (Extraction Sums of Squared Loadings)	
		28.0	5.2
		Factor	
Items		<i>Providing opportunities to learn from NZQA moderation reports</i>	<i>Checking internal assessment quality</i>
Q26.2	To provide opportunities for learning from the feedback in moderation reports	.921	
Q26.1	To provide opportunities for learning from the moderation results in moderation reports	.697	
Q26.3	To give feedback about how well the assessment materials assess achievement against a Standard	.377	.368
Note: coefficients less than 0.30 suppressed.			

Table 5. Pattern Matrix of Items Associated with the Organisational Quality Assurance Factor

		Percentage of variance (Extraction Sums of Squared Loadings)
		8.8
		Factor
Items		<i>Organisational quality assurance</i>
Q22.1	To provide evidence for consideration in an organisation's Self-assessment (ITPs and PTEs), or Self-review (Schools)	.765
Q22.2	To provide an evidence source that contributes to an organisation's External Evaluation and Review, and Provider Category (ITPs and PTEs), or Managing National Assessment review (Schools)	.643
Q22.3	To monitor an organisation's compliance with NZQA's rules and requirements	.622
Q22.5	To provide information for performance appraisals	.581
Q22.4	To give confidence to the management team or governing body about an organisation's teaching and assessment	.567
Q21.5	To provide information about whether or not content that is assessed is appropriate	.535
Q21.4	To check that assessments are fair to all students	.492
Note: coefficients less than 0.30 suppressed.		

Table 6. Pattern Matrix of Items Associated with the Maintaining Public Confidence Factor

		Percentage of variance (Extraction Sums of Squared Loadings)
		11.8
		Factor
Items		<i>Maintaining public confidence</i>
Q24.2	To maintain public confidence in the internal assessment conducted by organisations	.888
Q24.1	To help to ensure that education and qualifications are seen as credible by parents, employers and other education organisations	.873
Q24.3	To maintain public confidence in national qualifications	.737
Note: coefficients less than 0.30 suppressed.		

Respondents indicated that NZQA moderation has its greatest emphasis on the items associated with *Checking internal assessment quality*. There was almost unanimous endorsement of the rating of *High emphasis* across the three items. Figure 1 shows the rate and distribution of responses for item 21.1. These were closely mirrored by the response rates and distributions for the two other items associated with the factor (21.2 and 21.3; refer Table 2). Across the three items, 80–90 percent of respondents rated each as having *High emphasis*, around 10 percent rated each as having *Medium emphasis*, and barely any indicated that the items had *No* or *Low emphasis*.

To check whether or not the internal assessments of an organisation meet the requirements of the Standards, as set nationally (Item 21.1; Response rate 87%; n = 188)

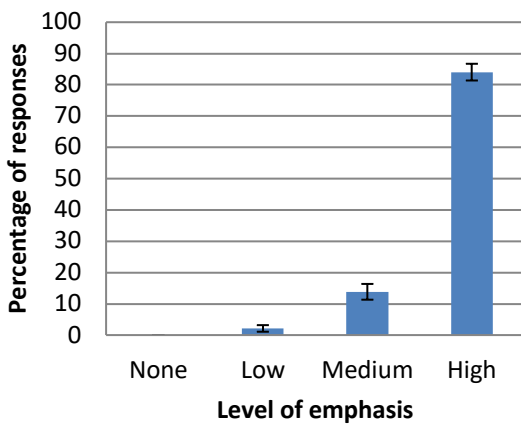


Figure 1. Distribution of responses for one of the three items contributing to the *Checking internal assessment quality* factor. Error bars denote standard error of the percentage of responses.

The perceived emphasis by NZQA moderation on *Improving internal assessment quality* was lower than it was on *Checking internal assessment quality*. The proportion of respondents who rated the items associated with the factor *Improving internal assessment quality* (see Table 3) as receiving *Medium* or *High emphasis* ranged from just under 60 percent (for item 27.2: ‘To stimulate professional and collegial conversation’) to approximately 90 percent (item 26.5: ‘To give feedback about the quality of assessor judgements’). Refer Figure 2 for representative response rates and distributions. The response rate for each of the eight items associated with this factor ranged from $n = 175$ to $n = 177$.

Figure 2 shows the distribution of responses for two of the items associated with the factor *Improving internal assessment quality*. The response distribution of item 27.3 was closely mirrored by those for items 27.1 (‘To give guidance regarding how to interpret the requirements of a Standard’), 27.2 (‘To stimulate professional and collegial conversation’), and 27.4 (‘To build the confidence of assessment designers or assessors’). Across the three, the only material variation in perceived emphasis from item 27.3 was for item 27.1, where around 10 percent more respondents rated the item as receiving *Medium emphasis* than did for 27.3. A corresponding several percent fewer respondents rated item 27.1 as receiving each of *No*, *Low* and *High emphasis*.

The remaining three items associated with the *Improving internal assessment quality* factor (refer Table 3) had response distributions which closely mirrored that of item 26.4. The distribution of responses for item 26.6 (‘To give feedback about how to improve assessor judgements’) and item 29.2 (‘To give feedback about whether an assessment supports teaching and learning’) were almost identical to that for item 26.4. A stronger level of emphasis was perceived for item 26.5. Compared to item 26.4,

just over 10 percent more respondents indicated that 26.5 received *High emphasis*, and just over 10 percent fewer respondents indicated that it received *Low emphasis*.

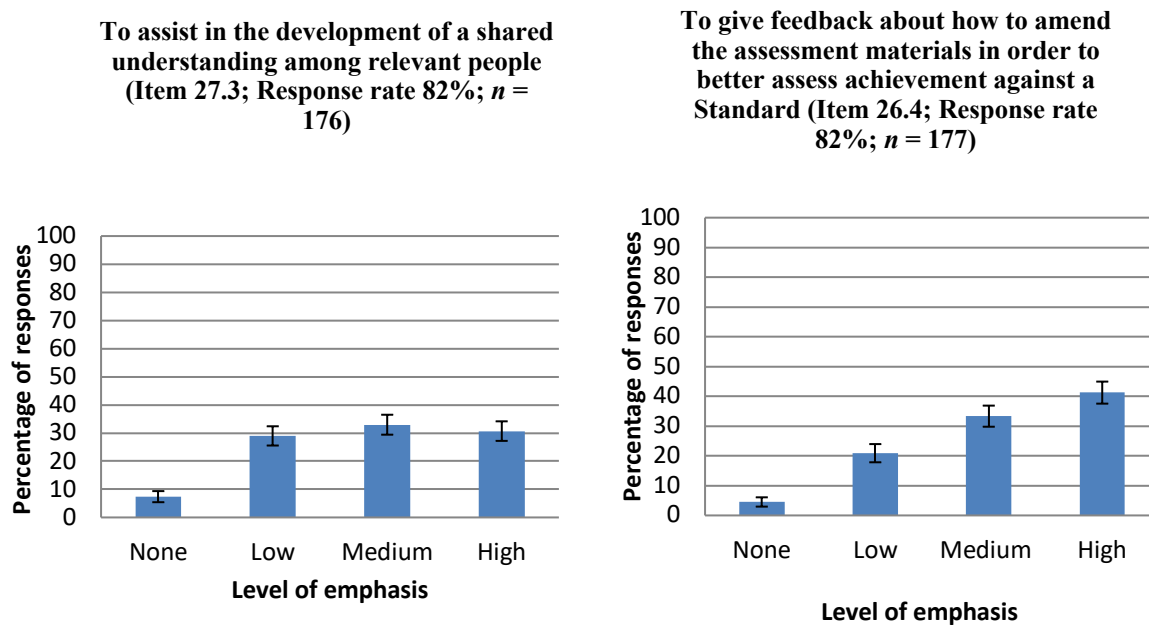


Figure 2. Distribution of responses for two of the eight items contributing to the *Improving internal assessment quality* factor. Error bars denote standard error of the percentage of responses.

The emphasis on *Providing opportunities to learn from NZQA moderation reports* was also relatively strongly endorsed, with more than 80 percent of respondents rating each associated item as receiving at least *Medium emphasis*. Figure 3 shows the rate and distribution of responses for item 26.2. The response rates for the other two items associated with this factor (26.1 and 26.3; refer Table 4) were almost identical to that of item 26.2.

The distribution of responses for item 26.1 ('To provide opportunities for learning from the moderation results in moderation reports') was also almost identical to that for item 26.2, albeit with a slightly weaker level of endorsement. Compared to 26.2, several percent fewer respondents rated item 26.1 as having *High emphasis*—and several percent more rated it as having *Low emphasis*.

The level of emphasis for item 26.3 ('To give feedback about how well the assessment materials assess achievement against a Standard') was perceived more strongly than for the other two items associated with this factor. Approximately 10 percent more respondents rated item 26.3 as receiving *High emphasis* than did item 26.2 (~55% cf ~45%; see Figure 3), and approximately five percent fewer respondents rated it as receiving *Medium* or *Low emphasis* than did 26.2 (~35% and 5%, respectively).

**To provide opportunities for learning from
the feedback in moderation reports
(Item 26.2; Response rate 82%; $n = 177$)**

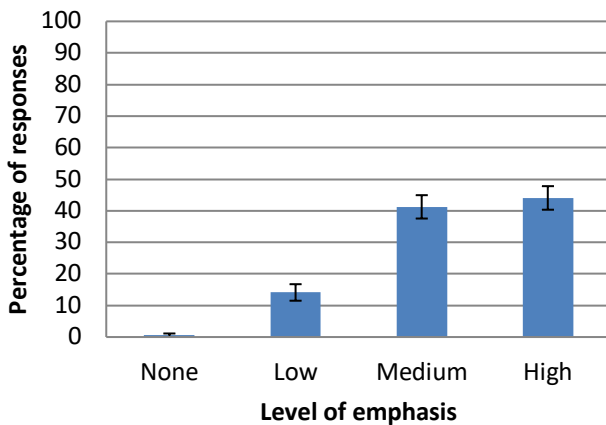


Figure 3. Distribution of responses for one of the three items contributing to the *Providing opportunities to learn from NZQA moderation reports* factor. Error bars denote standard error of the percentage of responses.

Respondents strongly endorsed the emphasis placed on *Organisational quality assurance* by NZQA moderation. A clear majority rated six of the seven items associated with this factor as receiving at least *Medium emphasis* (refer Table 5 for the items). The response rates for all of the items associated with this factor fell within the range of $n = 186$ -to- 189 .

The distribution of responses for item 22.4 is shown in Figure 4 and is representative of nearly all items associated with the factor (other than 22.5; see below). The response distributions for items 21.5, 22.2 and 22.3 mirrored that of item 22.4 almost exactly, with any variations for particular response categories being of only a few percentage points. In comparison, weaker endorsement was given to item 22.1 ('To provide evidence for consideration in an organisation's self-assessment or -review'). Half of the respondents rated item 22.1 as receiving *High emphasis* (compared with almost 60% who rated item 22.4 this way), and several percent more rated it as receiving *Low* and *Medium emphasis* than did 22.4. Item 21.4 ('To check that assessments are fair to all students') was slightly more strongly endorsed than was item 22.4. Almost 70 percent of respondents reported that item 21.4 received *High emphasis*, and slightly fewer reported it received either *Medium* or *Low emphasis*, as compared to 22.4.

In stark contrast, respondents felt that NZQA moderation placed the weakest emphasis on item 22.5 ('To provide information for performance appraisals'). One half of respondents rated this item as receiving *No* or *Low emphasis*. The ratings of *Medium* or *High importance* were each indicated by only one quarter of respondents.

To give confidence to the management team or governing body about an organisation's teaching and assessment (Item 22.4; Response rate 87%; $n = 187$)

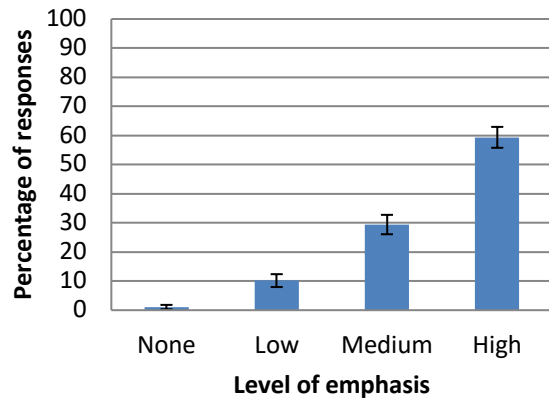


Figure 4. Distribution of responses for one of the seven items contributing to the *Organisational quality assurance* factor. Error bars denote standard error of the percentage of responses.

The emphasis of NZQA moderation on *Maintaining public confidence* was also strongly endorsed by respondents: At least 90 percent of them rated all associated items as receiving at least *Medium emphasis*, and more than 60 percent rated each as receiving *High emphasis*. Figure 5 shows the rate and distribution of responses for item 24.1, which are representative of those for the other two items associated with this factor (24.2 and 24.3; refer Table 6). The response rates for these two items were identical to that for item 24.1. The distribution of responses for item 24.2 ('To maintain public confidence in the internal assessment conducted by organisations') was essentially identical to that shown in Figure 5. The third item (24.3: 'To maintain public confidence in national qualifications') received slightly stronger endorsement than item 24.1, with several percent more respondents rating it as receiving *High emphasis*, and slightly fewer rating it as receiving *Low* or *Medium emphasis*.

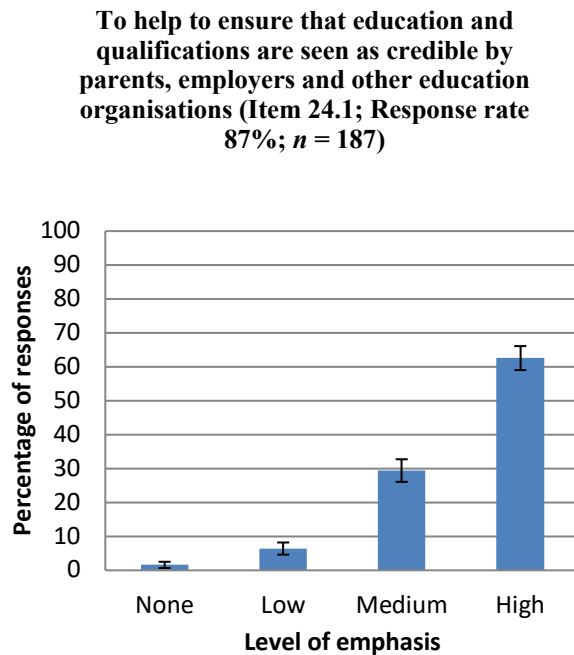


Figure 5. Distribution of responses for one of the three items contributing to the *Maintaining public confidence* factor. Error bars denote standard error of the percentage of responses.

One-way analyses of variance on the locations of the means revealed a significant difference among the three means on the *Organisational quality assurance* scale: $F(2, 184) = 11.39, p < .001$ (refer Figure 6). Independent samples t -tests showed that the difference between the means were significant for PTEs and schools: $t(170) = 4.40, p < .001$, and for ITPs and PTEs: $t(116) = 2.45, p = .02$, but not for ITPs and schools: $t(78) < 1, p = .85$.

No differences among the means for the organisation types on the other four scales were significant (Figure 6). *Checking internal assessment quality* scale: $F(2, 187) = 1.41, p = .25$; *Improving internal assessment quality* scale: $F(2, 174) = 1.73, p = .18$; *Maintaining public confidence* scale: $F(2, 184) = 1.91, p = .15$; *Providing opportunities to learn from NZQA moderation reports* scale: $F < 1$.

All correlations between the scales calibrated from the items associated with each of the five factors were positive and statistically significant. Disaggregated by organisation type, all correlations were similarly significant for PTEs, but not for ITPs or schools (see Table 7).

The strongest correlations for both the aggregated data and the individual organisation types was between *Improving internal assessment quality* and *Providing opportunities to learn from NZQA moderation reports* (aggregated: $r = .70, p < .01$). Also substantial was the correlation between *Improving internal assessment quality* and *Organisational quality assurance* (aggregated: $r = .58, p < .01$).

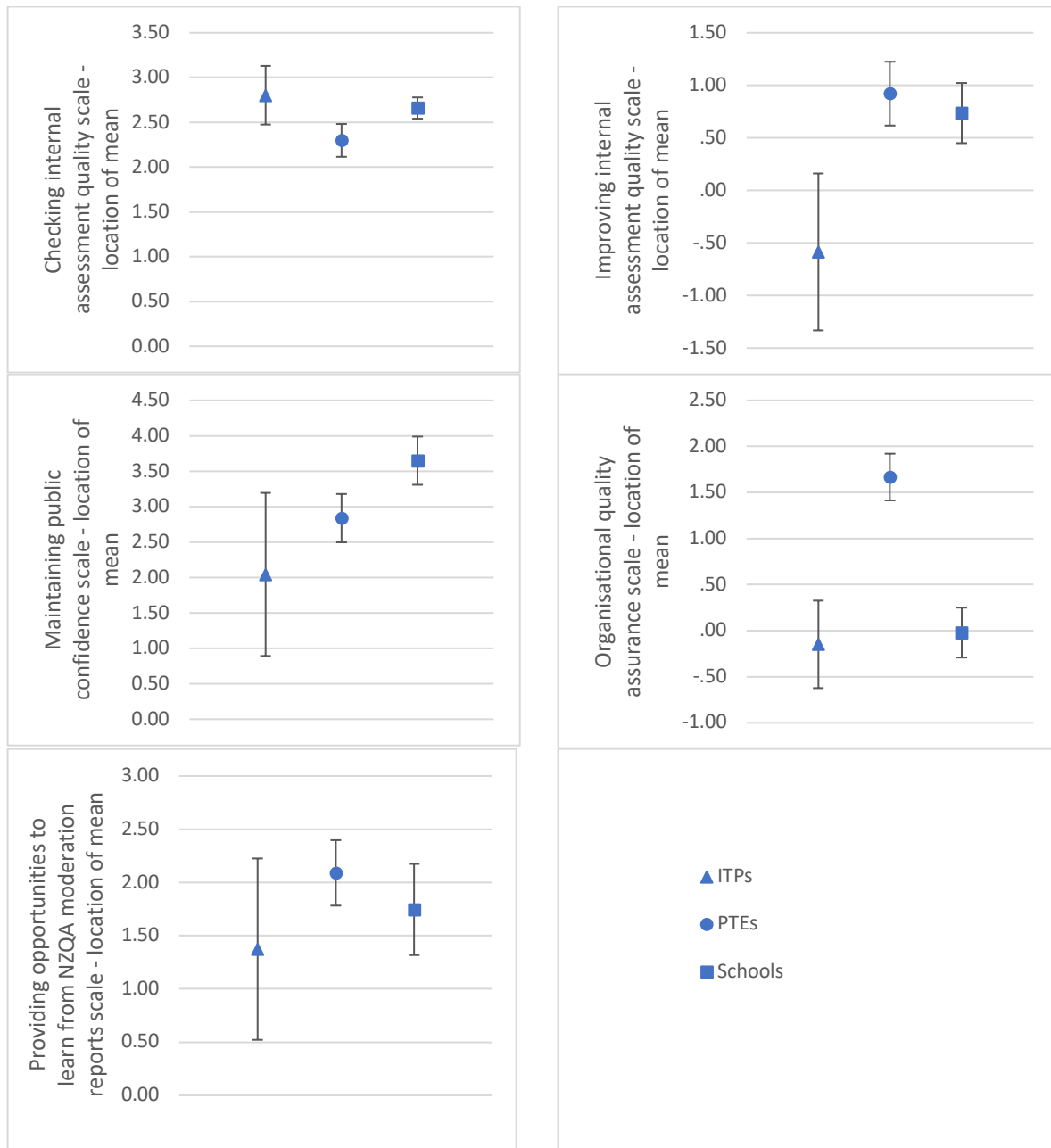


Figure 6. Measurement scale location of means for ITP, PTE and school respondents on each scale. Error bars denote standard errors of the means.

The weakest correlations involved the *Checking internal assessment quality* scale (although not for ITPs, for two correlations). In particular, with *Improving internal assessment quality*, *Providing opportunities to learn from NZQA moderation reports*, and *Maintaining public confidence*. For the aggregated and individual organisation types, for the most part, these correlations were weak or with lower levels of significance.

The correlations between *Improving internal assessment quality* and *Maintaining public confidence*; *Improving internal assessment quality* and *Providing opportunities to learn from NZQA moderation reports*; and *Maintaining public confidence* and *Organisational quality assurance* were markedly stronger for PTEs than for schools.

Table 7. Pearson's Correlation Coefficients Between Scales for the Aggregated Organisations and for the Individual Organisation Types.

NZQA moderation scales						
NZQA moderation scale		Organisation type	2	3	4	5
1	Checking internal assessment quality	Aggregate	0.24	0.28	0.35	0.29
		ITPs			0.59	0.57
		PTEs	0.26	0.28	0.41	0.29
		Schools		0.33	0.46	
2	Improving internal assessment quality	Aggregate		0.7	0.58	0.35
		ITPs		0.73	0.65	
		PTEs		0.77	0.62	0.41
		Schools		0.62	0.54	0.25
3	Providing opportunities to learn from NZQA moderation reports	Aggregate			0.43	0.3
		ITPs			0.56	
		PTEs			0.46	0.33
		Schools			0.39	0.27
4	Organisational quality assurance	Aggregate				0.35
		ITPs				0.62
		PTEs				0.45
		Schools				0.28
5	Maintaining public confidence					

Note: Correlations in boldface are significant at the 0.01 level (2-tailed), and correlations in standard font are significant at the 0.05 level (2-tailed). Non-significant correlations have been omitted.

Qualitative analyses

Notwithstanding a relatively low response rate, the qualitative data indicate that respondents perceive NZQA moderation to function in the areas of assessment quality, professional learning opportunities, organisational quality assurance, and maintaining public confidence (Table 8). The first of these relates closely to the first two factors identified in the factor analysis: *Checking internal assessment quality* and *Improving internal assessment quality*. The remaining three themes have direct correlation amongst the statistical factors.

Table 8. Areas in Which Leaders Perceived NZQA Moderation to Function, According to the Qualitative Survey Data (N denotes the number of respondents identifying each theme; percentages show the organisation types of those respondents for each theme).

	Assessment quality		Learning opportunities		Organisational quality assurance		Public confidence	
	<i>n</i>	%	<i>N</i>	%	<i>n</i>	%	<i>N</i>	%
Total	10		18		7		24	
ITPs	2	20	2	11	1	14	1	4
PTEs	7	70	11	61	6	86	20	83
Schools	1	10	5	28	0	0	3	13

Note: *n*, number of respondents who identified the theme; %, organisation types of those respondents for each theme

Ten respondents (5% of the sample) provided comments on **assessment quality** as a function of NZQA moderation, suggesting that most had nothing further to add to the closed-response items, or that they did not feel strongly enough about those items to respond. Eight (2 ITPs and 6 PTEs) agreed that NZQA moderation had a role in assessment quality, specifically in its consistency and robustness. For example, “to ensure fairness and consistency in assessment methodology used by different providers across the country—particularly where the providers are offering the same course” (PTE).

Eighteen respondents (8% of the sample) commented on the **professional learning opportunities** provided by NZQA. Of these, 12 (2 ITPs, 8 PTEs, and two schools) referred to the potential for opportunities to be provided by NZQA moderation, although 10 of them (both ITPs, 6 PTEs, both schools) disagreed that this potential was met. Those who disagreed cited the type of feedback as well as its paucity and variability as reasons. For example, “The quality of the moderator’s feedback would need to be more consistent for [the provision of professional learning opportunities] to be high emphasis [for NZQA moderation]” (School).

Ten respondents (6 PTEs and 4 schools) referred to other professional learning opportunities in relation to moderation, provided by NZQA, with eight expressing dissatisfaction with them. Six (2 PTEs, 4 schools) referred to NZQA workshops, but the nature of comment differed according to organisation type: PTE respondents acknowledged that NZQA does run provider workshops; school respondents complained about access or cost.

Seven respondents (3% of the sample) commented on **organisational quality assurance** as a function of NZQA moderation, also suggesting that most had nothing to add to the closed-response items, nor did most feel particularly strongly about them. Those who commented mainly did so to refer to compliance with NZQA requirements—either by organisations or as a focus for NZQA itself.

Twenty-four respondents (11% of the sample) addressed **maintaining public confidence** as a function of NZQA moderation. More commented on this function than did on any other, suggesting that respondents felt more strongly about it than they did about other functions. Ten (1 ITP, 7 PTEs and 2 schools) agreed that NZQA moderation has such a role. For example, “If assessment is not robust, we are doing all stakeholders an injustice. [NZQA moderation] should ensure assessment is robust, and if it is the system will produce fit-for-purpose graduates who will instil confidence in the system” (PTE).

A further nine respondents (8 PTEs and 1 school) expressed dissatisfaction with—or only qualified agreement that—NZQA moderation has this role, citing shortcomings in the moderation process, or that it led to organisational responses which adversely impact on educational quality. For example,

NZQA places a lot of emphasis on the ‘perception’ of assured quality but appears to be hampered by their systems. For example, I would be calling for ‘specific’ reported assessment scripts for moderation rather than any three assessing a unit. (PTE)

Five respondents (1 ITP, 3 PTEs, 1 school) opined that NZQA moderation is not publicly visible, which, in their view, adversely impacts on the function of maintaining public confidence, and four PTEs expressed dissent from the notion that NZQA moderation even has such a role.

Discussion

Perceived functions of NZQA moderation

The quantitative and qualitative findings largely aligned: NZQA moderation was perceived as functioning in the domains of assessment quality (checking and improving), professional learning opportunities, organisational quality assurance, and maintaining public confidence. The emphasis that respondents thought that NZQA moderation placed on each function varied from an almost unanimous endorsement of checking internal assessment quality, to a much lower endorsement of stimulating professional and collegial conversation.

The functions also differed in breadth: checking assessment quality was the most narrowly focused; other functions were broader in focus in that they address aspects of education that are peripheral to assessment quality. Furthermore, whereas perceptions of emphasis on checking internal assessment quality and maintaining public confidence were seen as fairly 'stand-alone', perceptions of moderation improving internal assessment quality, providing learning opportunities, and providing organisational quality assurance were mutually correlated.

Assessment quality

The high level of agreement that NZQA moderation addresses the quality of internal assessment is unsurprising; indeed, moderation is frequently defined in such terms (Broadbent, 2018; Hipkins et al., 2016; Klenowski & Wyatt-Smith, 2014). The quantitative data revealed a distinction in the perceived role of NZQA moderation regarding assessment quality—evinced by a low correlation between the two factors from the EFA relating to this function: *Checking internal assessment quality* and *Improving internal assessment quality*—with the former relating to a relatively narrow quality-control mechanism, and the latter to a broader and longer-term process of improvement. The response distributions for the items associated with these two factors evinces a much stronger emphasis on the former than on the latter.

The correlations between *Checking internal assessment quality* and almost all other NZQA moderation factors were weak, suggesting that leaders see internal assessment quality control by NZQA moderation as a separate and essentially 'stand-alone' function. Checking internal assessment quality is a narrow function in that it focuses on assessment materials and assessor judgements meeting the requirements of each standard, without concern for whether the assessment supports learning. As such, it addresses only the *assessment* part of the teaching-learning-assessment cycle (Black & Wiliam, 2018). This conceptualisation thus positions moderation as a lens through which it is possible to consider internal assessment in isolation and ignore the contexts within which assessment occurs. When such a narrow focus is employed, assessment activities and practices that are deleterious to learning might be allowed to continue unchecked (Black & Wiliam, 2018; Crooks, 2011). However, given the quasi-autonomous nature of internal assessment for qualifications in NZ—conducted as it is by self-governing education organisations against the assessment standards they chose—a perception of the importance of checking on its quality by the agency responsible for maintaining that quality, is nonetheless unsurprising.

The function of moderation as a means of improving assessment quality has a broader focus. In addition to improvement of assessment materials and assessor judgements in themselves, moderation

with this focus also encompasses the extent to which assessment supports the teaching and learning parts of the cycle. It also includes aspects of professional learning for assessors and teachers, and thus has potential to inform aspects of educational practice beyond assessment—such as pedagogy or course design.

Professional learning opportunities

Leaders saw NZQA moderation as emphasising professional learning opportunities, mainly via moderation reports (which is one avenue for providing such opportunities, as Crimmins et al., 2016, found). This function aligns with an improvement purpose (Bloxham et al., 2016; Klenowski & Wyatt-Smith, 2014; Misko, 2015) and echoes one of the main roles of moderation (Grant, 2012). A clear majority of respondents rated NZQA moderation as placing at least a medium level of emphasis on the associated items. This finding coheres with the earlier findings of the Controller and Auditor-General (2012) and of Hipkins (2013), that most school teachers consider NZQA moderation feedback to be helpful or useful.

Professional learning also featured in the qualitative data, although most who commented did not agree that NZQA moderation serves such a role, due to a perception that feedback is of poor quality. This may reflect the *expert moderation* approach used by NZQA, which is conducted at distance. Many of the learning opportunities afforded by moderation rely on high-quality feedback (Bloxham et al., 2016; Crimmins et al., 2016). Feedback from NZQA moderation is contained in moderation reports and any dissatisfaction with that feedback would likely have an amplified effect on the recipients' perceptions of any learning opportunities it affords. Furthermore, leading up to and at the time of data collection, NZQA did not always provide feedback in moderation reports (NZQA, 2014b). Nor does it appear that such feedback is always currently provided (NZQA, 2019).

In the qualitative data, respondents also referred to learning opportunities provided by NZQA's moderators, albeit outside the direct ambit of the NZQA moderation process itself (e.g., via workshops). Again, however, most who mentioned these opportunities did so to express dissatisfaction with them.

Organisational quality assurance

The organisational quality assurance function of moderation encompasses a range of aspects, including satisfying external requirements, providing information for other quality assurance processes and maintaining the confidence of governance and management teams. Construed as such, moderation serves as a lens through which to view other organisational processes and factors, including those that impact only indirectly on internal assessment.

Respondents strongly endorsed six of the seven items associated with this factor. Many leaders consider this function of moderation to encompass *both* improvement *and* accountability—as Harvey and Newton (2007) advocated—instead of subscribing to the view that it addresses only one or the other. The strength of correlations between the factor *Organisational quality assurance*, and the factors *Improving internal assessment quality* and *Providing opportunities to learn from NZQA moderation reports*, indicate that this function aligns relatively closely with a broad improvement purpose. However, it was also moderately positively correlated (for PTEs) with both separate, 'stand-alone' functions: *Checking internal assessment quality*, and *Maintaining public confidence*; and for schools, with *Checking internal assessment quality*. Thus, at least as it relates to NZQA moderation, organisational quality assurance also appears to align with accountability to a certain extent. The relevant qualitative data supported an alignment with accountability-type functions, specifically in terms of compliance with NZQA's requirements.

Maintaining public confidence

A fourth function of NZQA moderation identified by respondents was the maintenance of public confidence, reflecting what the Controller and Auditor-General (2012) held to be a key function of NZQA moderation. This is a broadly-focused area that encompasses public confidence in the internal assessment that occurs within individual organisations, in the education and qualifications of graduates, and in national qualifications systems themselves. Maintaining public confidence is fundamentally aligned—and closely associated—with accountability (Controller and Auditor-General, 2012; Grant, 2012; Misko, 2015).

The *Maintaining public confidence* factor was correlated relatively weakly with other moderation functions, suggesting that, in the eyes of the respondents, it 'stands alone' to a certain extent (although a relative lack of variability in responses to the items associated with this factor might also have suppressed the correlations). Respondents strongly endorsed the view that NZQA moderation emphasises the maintenance of public confidence. Various perspectives were put forward in the qualitative data, in both agreement and disagreement that it serves this purpose, and dissatisfaction with its effectiveness. This function attracted the most comments, indicating that respondents felt more strongly about it than they did about other functions. This might, in part, be due to the importance of reputation and public confidence to organisations, and the impacts (including financial) that a change in either can have on an organisation.

Since our data collection there has been increased emphasis by NZQA on tertiary sector quality assurance and refinements to the NZQA moderation systems. Nonetheless, to date, the moderation systems implemented by NZQA remain fundamentally unchanged. Despite some commentators (e.g., (Darr, 2019; Darr & Hipkins, 2020) advocating for the opportunity to rethink assessment and moderation practices to be taken in the current climate, it would be a major shift for NZQA to materially alter their moderation practice or approach in response to the changes to NCEA. NZQA will continue to quality assure internal assessments against standards in order that the public can have confidence in NZ qualifications (NZQA, n.d.b).

Differences in perception according to organisation type

There were no observable differences between the organisation types in terms of respondents' perceptions of the functions of checking or improving assessment quality, professional learning opportunities, or maintaining public confidence. Leaders from the three organisation types tended to view the emphasis placed on these functions similarly. In contrast, respondents from PTEs perceived a higher emphasis on organisational quality assurance than respondents from either ITPs or schools; ITPs and schools rated the associated items similarly.

Those from PTEs also appeared to consider the broader and improvement-focused functions of NZQA moderation to be more closely associated or integrated than leaders from ITPs or schools. These inferences are supported by the magnitude and significance of several correlations: between *Providing opportunities to learn from NZQA moderation reports* and *Improving internal assessment quality*, and between *Maintaining public confidence*, *Improving internal assessment quality*, and *Organisational quality assurance*—all of which were markedly stronger (and in one case, more significant) for PTEs than for schools. Furthermore, school respondents appeared to see the function of maintaining public confidence as more 'stand-alone' than PTE respondents did. All correlations between this scale and others were weaker or less significant than the corresponding correlation for PTEs.

The differences between PTE and school respondents for the function of organisational quality assurance, and the perceived integration of the broader and improvement-focused NZQA moderation functions, may reflect differences in policy and contextual factors facing the different organisation types.

Compared with schools, PTEs are subjected to fewer and looser regulations and parameters, and are subject to fewer regulatory and quality assurance bodies (NZQA only, versus NZQA, ERO, and the Ministry of Education). PTEs are also required to engage in fewer quality assurance processes (ERO, 2020; Hipkins et al., 2016; NZQA, n.d.a, 2020c, 2020b). Such a situation may have resulted in PTEs viewing moderation—a process they are required to engage with—as a mechanism to use as broadly as possible and with which to inform other organisational processes. Conversely, the greater regulatory and quality-assurance oversight to which schools are subject might have narrowed the uses they make of each of the various quality assurance processes they are required to implement. In other words, greater regulatory oversight might result in narrower and more compliance-driven behaviour.

The similarity across organisation types in the perception that moderation is essential to the maintenance of public confidence is, perhaps, surprising, given the differences in policy pertaining to organisations in the secondary and tertiary sectors (e.g., different moderation systems and sample-selection processes, and for the secondary sector, the national agreement rate). A possible explanation is that there may not have been much awareness in our sample of the regulatory and quality assurance requirements that pertain to sectors other than those with which each respondent was familiar.

Multiple reforms affecting the secondary and tertiary sectors have been embarked upon in recent years. However, these reforms appear unlikely to decrease the current differences that exist between PTEs and schools in terms of regulations that apply, or regulatory or quality assurance bodies, or quality assurance processes, that the organisations are subjected to (NZ Government, n.d., 2019a). Nor do the reforms appear likely to have much influence on the governance or management of PTEs. Furthermore, while the scope for variation between schools or between (subsidiary) ITPs appears likely to reduce under the reforms (NZ Government, n.d., 2019b), there appears to be just as much potential as ever for variation between the different types of organisation under the reforms. As such, our findings illuminate key aspects of, and nuances within, a core mediating factor in this area of educational policy enactment.

Conclusions

The policy-enactment process in education organisations, including that relating to NZQA moderation, is mediated by many factors, including the people involved. The interpretations of those policy narrators are percolated through—and are thus shaped by—their current frames of reference (Ball et al., 2012; Maguire et al., 2015; Spillane et al., 2002). In the present paper, we have identified the functions of NZQA moderation as understood by the academic leaders who are likely to be in narrator roles for moderation policy in NZ education organisations. Notwithstanding changes in NZQA moderation policy and quality assurance activity in the tertiary sector, the contextual changes that will eventuate due to the Reforms of Tomorrow's Schools and of Vocational Education—and irrespective of whether material changes are made to NZQA moderation policy or practice as a result of the NCEA review—our findings have implications for education organisations and NZQA.

Implications for education organisations

Present findings could assist leaders to identify and examine their own perceptions of the functions of NZQA moderation. Doing so might help them to recognise the influence of those perceptions on policy enactment within their organisations and, in turn, identify how to make better use of opportunities provided by moderation. Engaging in such critical reflection may also assist leaders to read moderation policies more critically—including any amendments issued in the current climate—and to take a greater role in implementing those policies in ways that maximise the benefit to their organisations.

Implications for NZQA

A comparison of the present findings with existing—or new—moderation policy could ascertain the degree of alignment between the functions of NZQA moderation perceived by academic leaders and those intended by NZQA. Doing so may provide insights into current or likely policy enactment and any potential divergence from the intentions of policy makers. Because people are more likely to recognise and interpret familiar policy messages than unfamiliar ones (Spillane et al., 2002), the perceptions of leaders illuminated here could be used to inform policy and communications work undertaken by NZQA, to improve alignment between policy intention and enactment.

All respondents recognised the narrow, control-focused function of checking internal assessment quality. We can infer, from the lack of dissenting comment in the qualitative data, that respondents believed that NZQA moderation is effective in this function. It is therefore important that NZQA maintains and enhances this function to ensure the education and qualifications systems are robust and trustworthy—and retain legitimacy in the eyes of the public.

The leaders participating in our study perceived that NZQA moderation has broader, improvement-focused functions, such as organisational quality assurance. However, the perceived levels of both the emphasis and the effectiveness of NZQA moderation on these broader and improvement-focused functions were mixed. This result suggests that there may be opportunities to strengthen these broader functions—whether through NZQA moderation itself (e.g., feedback), or via other avenues (e.g., workshops)—and thereby to enhance the value that organisations can gain from engaging with NZQA and NZQA moderation.

National moderation, such as that conducted by NZQA, is critical to ensuring that high-stakes internal assessment occurring in the secondary and tertiary sectors is trustworthy, has integrity, and is conducted in such a way as to do no harm to—if not to enhance—teaching and learning. In doing so, such moderation supports the credibility and social legitimacy of the organisations, qualifications awarded, and of the education and qualifications systems themselves.

References

- Ball, S., Maguire, M., & Braun, A. (2012). *How schools do policy: Policy enactment in secondary schools*. Routledge.
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551–575.
<https://doi.org/10.1080/0969594X.2018.1441807>
- Bloxham, S., Hughes, C., & Adie, L. (2016). What's the point of moderation? A discussion of the purposes achieved through contemporary moderation practices. *Assessment and Evaluation in Higher Education*, 41(4), 638–653. <https://doi.org/10.1080/02602938.2015.1039932>
- Broadbent, J. (2018). Large class teaching: How does one go about the task of moderating large volumes of assessment? *Active Learning in Higher Education*, 19(2), 173–185.
<https://doi.org/10.1177/1469787417721360>
- Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7th ed.). Routledge.
- Controller and Auditor-General. (2012). *New Zealand Qualifications Authority: Assuring the consistency and quality of internal assessment for NCEA* [Performance audit report]. Office of the Auditor-General.
- Creswell, J. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson Education.
- Crimmins, G., Nash, G., Oprescu, F., Alla, K., Brock, G., Hickson-Jamieson, B., & Noakes, C. (2016). Can a systematic assessment moderation process assure the quality and integrity of

- assessment practice while supporting the professional development of casual academics? *Assessment and Evaluation in Higher Education*, 41(3), 427–441.
<https://doi.org/10.1080/02602938.2015.1017754>
- Crisp, V. (2017). The judgement processes involved in the moderation of teacher-assessed projects. *Oxford Review of Education*, 43(1), 19–37. <https://doi.org/10.1080/03054985.2016.1232245>
- Crooks, T. (2011). Assessment for learning in the accountability era. *New Zealand. Studies in Educational Evaluation*, 37, 71–77. <https://doi.org/10.1016/j.stueduc.2011.03.002>
- Dalby, D., & Noyes, A. (2018). Mathematics education policy enactment in England’s Further Education colleges. *Journal of Vocational Education and Training*, 1–17.
<https://doi.org/10.1080/13636820.2018.1462245>
- Darr, C. (2019). Approaching classroom assessment after the NCEA review. *Set: Research Information for Teachers*, 3, 57–61. <https://doi.org/10.18296/set.0153>
- Darr, C., & Hipkins, R. (2020). Opportunities to reframe moderation practices in the wake of the COVID-19 pandemic. *Set: Research Information for Teachers*, 2, 49–51.
<https://doi.org/10.18296/set.0172>
- Education Review Office. (2020). *About ERO*. <https://www.ero.govt.nz/footer-upper/about-ero/>
- Gerritsen, J. (2017, July 26). NZQA finds 27 tertiary institutions not up to scratch. *Radio New Zealand*. <http://www.radionz.co.nz/news/national/335834/nzqa-finds-27-tertiary-institutions-not-up-to-scratch>
- Grant, L. (2012). Cross-sector moderation as a means of engaging staff with assessment and teaching pedagogy. *Policy Futures in Education*, 10(4), 434–446.
<http://dx.doi.org/10.2304/pfie.2012.10.4.434>
- Hambleton, R., Swaminathan, H., & Rogers, H. (1991). *Fundamentals of item response theory*. Sage.
- Hardy, I. (2015). Curriculum reform as contested: An analysis of curriculum policy-enactment in Queensland, Australia. *International Journal of Educational Research*, 74, 70–81.
<http://dx.doi.org/10.1016/j.ijer.2015.09.010>
- Harvey, L., & Newton, J. (2007). Transforming quality evaluation: Moving on. In D. F. Westerheijden, B. Stensaker, & M. J. Rosa (Eds.), *Quality assurance in higher education—Trends in regulation, translation and transformation* (pp. 225–245). Springer.
- Hipkins, R. (2013). *NCEA one decade on: Views and experiences from the 2012 NZCER National Survey of Secondary Schools*. New Zealand Council for Educational Research.
- Hipkins, R., Johnston, M., & Sheehan, M. (2016). *NCEA in context*. New Zealand Council for Educational Research.
- Keddie, A. (2019). Context matters: Primary schools and academies reform in England. *Journal of Education Policy*, 34(1), 6–21. <https://doi.org/10.1080/02680939.2017.1402959>
- Klenowski, V., & Wyatt-Smith, C. (2014). *Assessment for education; Standards, judgement and moderation*. Sage.
- Maguire, M., Braun, A., & Ball, S. (2015). ‘Where you stand depends on where you sit’: The social construction of policy enactments in the (English) secondary school. *Discourse: Studies in the Cultural Politics of Education*, 36(4), 485–499.
<https://doi.org/10.1080/01596306.2014.977022>
- Manly, B. (1986). *Multivariate statistical methods: A primer*. Chapman and Hall.
- Mertens, D. (2010). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative and mixed methods* (3rd ed.). Sage.
- Ministry of Education. (n.d.). *About NCEA*. Te Kete Ipurangi. <https://ncea.tki.org.nz/About-NCEA>
- Ministry of Education. (2020). *Briefing to the Incoming Minister*.
<https://www.education.govt.nz/assets/Documents/Ministry/Publications/Briefings-to-Incoming-Ministers/R-MOE-BIM-November-2020-Proactive-release.pdf>

- Misko, J. (2015). *Developing, approving and maintaining qualifications: Selected international approaches* (Research Report No. 2775). National Centre for Vocational Education Research.
- New Zealand Government. (n.d.). *Summary of change decisions: Reform of vocational education*. Education Conversation/Kōrero Mātauranga. <https://conversation.education.govt.nz/assets/RoVE/AoC/RoVE-Summary-of-Change-Decisions.pdf>
- New Zealand Government. (2019a). *NCEA change package 2019 Overview*. Education Conversation/Kōrero Mātauranga. <https://conversation.education.govt.nz/assets/NCEA/NCEA-Change-Package-2019-Web.pdf>
- New Zealand Government. (2019b). *Supporting all schools to succeed: Reform of the Tomorrow's Schools system*. Ministry of Education. <https://conversation.education.govt.nz/assets/TSR/November-2019/TSR-Government-Response-WEB.pdf>
- New Zealand Qualifications Authority. (n.d.a). *Managing National Assessment (MNA) reviews*. New Zealand Qualifications Authority. <https://www.nzqa.govt.nz/providers-partners/assessment-and-moderation-of-standards/managing-national-assessment-in-schools/managing-national-assessment-mna-reviews/>
- New Zealand Qualifications Authority. (n.d.b). *What we do*. New Zealand Qualifications Authority. <https://www.nzqa.govt.nz/about-us/our-role/what-we-do/>
- New Zealand Qualifications Authority. (2014a). *Assessment (including examination) rules for schools with consent to assess 2014*. <http://www.nzqa.govt.nz/about-us/our-role/legislation/nzqa-rules/assessment-and-examination-rules-for-schools-2014/>
- New Zealand Qualifications Authority. (2014b). *National external moderation manual for Tertiary Education Organisations 2014*. <https://www.nzqa.govt.nz/providers-partners/assessment-and-moderation-of-standards/tertiary-moderation/manual-for-teo/>
- New Zealand Qualifications Authority. (2017). *EQuate September 2017 (EQuate)* [Circular]. New Zealand Qualifications Authority. <https://www.nzqa.govt.nz/about-us/publications/newsletters-and-circulars/equate/september-2017/>
- New Zealand Qualifications Authority. (2018). *National external moderation guidelines for Tertiary Education Organisations*. <https://www.nzqa.govt.nz/assets/Providers-and-partners/Monitoring-and-Assessment/2018-TEO-NEM-guidelines.pdf>
- New Zealand Qualifications Authority. (2019). *Assessment (including examination) rules for schools with consent to assess 2020*. <https://www.nzqa.govt.nz/assets/About-us/Our-role/Rules/2020/CR19438-signed-Assessment-Rules-2020.pdf>
- New Zealand Qualifications Authority. (2020a). *Consent to assess against standards on the Directory of Assessment Standards rules 2011 v4 (Approved version 4)*. <https://www.nzqa.govt.nz/assets/About-us/Our-role/Rules/2020/CTA-Rules-with-all-amendments-April-2020.pdf>
- New Zealand Qualifications Authority. (2020b). *National external moderation guidelines for Tertiary Education Organisations (2020)*. <https://www.nzqa.govt.nz/assets/Providers-and-partners/Monitoring-and-Assessment/NEM-guidelines-for-TEOs-2020.pdf>
- New Zealand Qualifications Authority. (2020c). *Quality assurance (including external evaluation and review (EER)) rules 2016 v4 (Approved version 4)*. <https://www.nzqa.govt.nz/assets/About-us/Our-role/Rules/2020/QA-including-EER-Rules-with-all-amendments-to-April-2020.pdf>
- Newby, P. (2010). *Research methods for education*. Pearson Education.
- Newton, P., & Shaw, S. (2014). *Validity in educational and psychological assessment*. Sage. <http://ebookcentral.proquest.com>

- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika*, 34(1), 1–97. <http://dx.doi.org/10.1007/BF03372160>
- Sheikh, I., & Bagley, C. (2018). Towards a policy social psychology: Teacher engagement with policy enactment and the core concept of Affective Disruption. *British Educational Research Journal*, 44(1), 43–60. <https://doi.org/10.1002/berj.3316>
- Spillane, J., Reiser, B., & Reimer, T. (2002). Policy implementation and cognition: *Reframing and refocusing implementation research*. *Review of Educational Research*, 72(3), 387–431.
- Tabachnick, B., & Fidell, L. (2013). *Using multivariate statistics* (6th ed). Pearson Education.
- Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioural sciences*. Sage.
- Williams, A. H., (2019). *Perceptions of academic leaders in New Zealand regarding the functions of moderation of internal assessment: A mixed methods study* [Doctoral thesis, Victoria University of Wellington]. <http://hdl.handle.net/10063/8151>

^{vi} The sample-selection process that TEOs are required to use for achievement standards has since been tightened to more closely reflect that required of schools (e.g., NZQA, 2018).

^{vii} It is noted that the t-tests involving ITPs were of low statistical power, due to the few ITPs included in the sample compared with the size of the other two groups (Manly, 1986).