Evaluating Assessment Moderation: Towards Best Practice

Christine Shobana Arthur
Sunway University, Malaysia
christines@sunway.edu.my

ABSTRACT

Moderation of assessment is done to ensure consistency of marking across academicians, subjects, and programmes. The concerns that have always been raised by many academicians from the tertiary education setting are the difficulties of attaining consistency, the apprehensiveness of new academic members and the demanding nature of the assessment moderation process which increases the workload of academicians. Using a Theory of Change (Weiss, 1998), this paper provides a brief case-study description by evaluating the intended outputs, assumptions and activities of assessment moderation. In particular, this study aims to identify and depict the assessment moderation practice, elicit views on its effectiveness and assess the extent to which improvements can be attained in the higher education (HE) setting. The academic staff members of University A were interviewed to obtain their views and comments on their assessment moderation practice. The participants were selected based on the following characteristics: 1) Must have completed assessment moderation for at least 10 rounds (have at least 5 years and above of experience in moderation), 2) Must have carried out assessment moderation once or twice (have less than one year of experience). The outcome of this study challenges University A to improve its assessment moderation practice by introducing other relevant and suitable methods of moderation. The paper also proposes applicable solutions for the institution to address the lack of training and support amongst new academic staff members and part-timers. The need for establishing better coordination between moderators and markers as well as establishing a formal buddy system were highlighted in this case study. This paper empahsises continous quality improvement at the HE setting because universities allow for substantial discretions and local adaptation (Weick, 1995). All in all, this study adds further weight to the existing body of literature on moderation practices in HE institutions.

KEYWORDS: Assessment Moderation, Theory of Change, Shared Understanding

1.0 INTRODUCTION

In the higher education (HE) setting, assessment is always considered a tool to shape and influence students' experience (Bloxham & Boyd, 2007). Another crucial aspect of assessment is the assessment moderation practice which ensures that the marker has awarded grades fairly and consistently while adhering to marking criteria. This practice is crucial in order to maintain the quality of HE. As Holroyd (2000) mentioned, "there are pressures for accountability, transparency, and consistency from government and from potentially litigious students as consumers which make reliable marking more important now than ever before". If this was mentioned 18 years ago, the importance of maintaining appropriate and effective assessment moderation system has also increased significantly more so now because of the demanding nature of HE.

The governing body in Malaysia, MQA, has clearly stated the need for assessment moderation. It is mentioned that there is a need for HE providers to establish a system for moderation of assessment in order to ensure consistency and quality (MQA, 2013). However, how the practice is established by every institution may differ. For that reason, it is interesting to look at Chart 1 which shows the discretion framework of evaluative practices in institutions by Bamber, (2011, p.129).

Connecting the model proposed by Bamber (2011) to the assessment moderation practice in Malaysia, the practice basically falls under the quadrant of tightly controlled focus with a loosely controlled process. The focus of moderation has been established by the governing body, while the mechanism of achieving its goals relies on the institution itself. When different institutions use different mechanisms of evaluating reliability and consistency, there is definitely a need to evaluate this evaluative practice in order to identify and depict assessment moderation practice, elicit views on its effectiveness and assess the extent to which improvements might be sought after. This study also suggests that the Bamber's model (Figure 1) serves as a valuable analytical tool in elucidating the control University A has over the way it carries out its moderation of assessments practice. This model aptly divides the discretion capacity and control the institution has over assessment moderation.

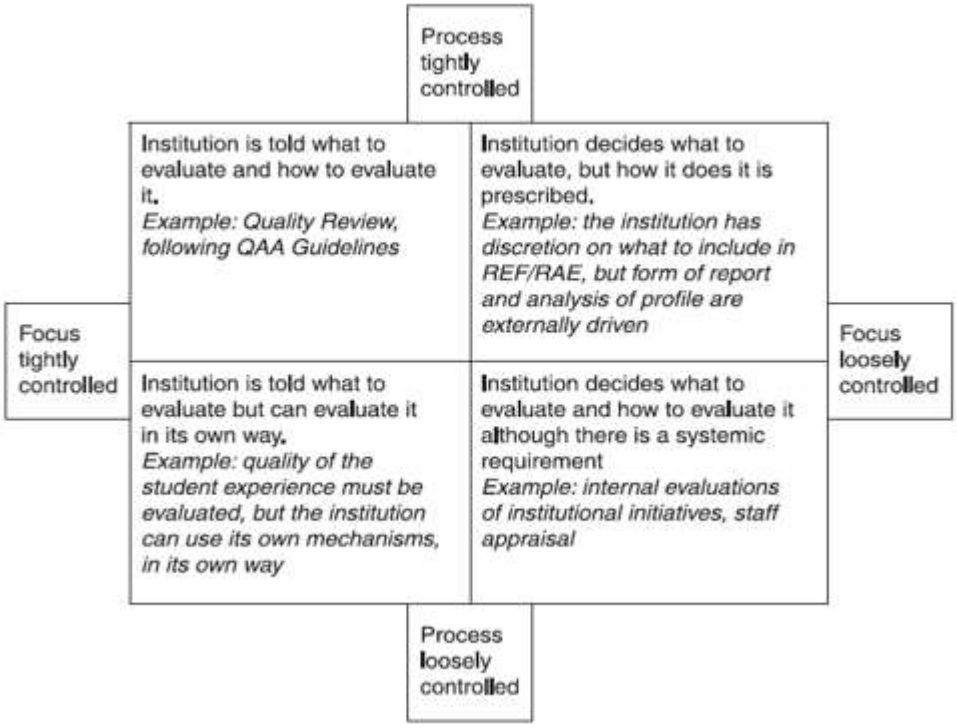

Figure 1: Discretion framework of evaluative practices in institutions (Bamber's Model, 2011)

According to Bloxham, Hughes and Adie (2016) assessment moderation is described as a crucial, critical yet problematic component of quality assurance in the HE context. In reality, the purpose of moderation is to ensure the validity and reliability of grading judgments and decisions. However, the concerns that have always been raised by many academicians are the difficulties of attaining reliability (Watty et al., 2014; Sadler, 2013; Annetts, Jones & Deursen, 2013), the apprehensiveness of new academic members or sessional staff (Beutel et. al, 2017; Grainger, Adie & Weir, 2016) and the demanding nature of the assessment moderation process which eventually increases the workload of academicians (Bloxham, 2009).

For these reasons, the evaluation for assessment moderation practice is devised with the intention of developing a better moderation practice for University A, a private HE institute in Malaysia, and for the academic community.

In order to achieve the objective of this paper, the following research questions are posited :

1. To what extent does University A provide professional support for the development of knowledge, understanding and practice in assessment moderation?
2. How was the assessment moderation practice delivered and received?
3. To what extent are the assessment moderation practice goals and objectives achieved?

Having established the research questions, this study sets out to evaluate the moderation practice of University A by identifying the extent the institution provides professional support in terms of assessment moderation especially for the new teaching staff. This study also seeks to carry out the investigation in an inclusive manner whereby the input of new teaching staff is also taken into consideration. Besides that, the study also takes into account the details of assessment moderation activities in order to identify whether the goals of assessment moderation have been met. Hence, whilst the objective is to achieve validity and consistency of marking judgements, the argument that is raised is whether the objective is being achieved.

2.0 SITUATING THE STUDY

The purpose of evaluation is to assess the goals of a programme in relation to its effectiveness which eventually contributes to the improvement of the programme (Weiss, 1972). This evaluation can be associated with continuous quality improvement (CQI) in HE. CQI refers to the understanding of the function of a system and improving the performance of the system.

2.1 Evaluation Approach

This evaluation adopts the Theory of Change (ToC) model. According to Clark & Anderson (2004, p.12), "ToC links outcomes and activities to explain how and why the desired change is expected to come about". Besides that, Weiss (1998, p.55) also pointed out that ToC not only helps to answer the question "Did the program work?" but also "What made it work?", "Why was it successful or unsuccessful?" and "How can we make it better?" Msila & Setlhako (2013) described ToC as a necessary tool in evaluating for CQI:

> "Every community needs a roadmap for change. Instead of bridges, avenues and freeways, this map would illustrate destinations of progress and the routes to travel on the way to achieving progress. The map would also provide commentary about assumptions, such as the final destination, the context for the map, the processes to engage in during the journey and the belief system that underlies the importance of travelling in a particular way. This type of map is called a *theory of change*".

Therefore, in order to understand the complexity of the assessment moderation practice in University A, the process is illustrated in Chart 2. The intended outcome of the practice is described first. It is then followed by the underlying assumptions regarding this practice. Then the activities that takes place to fulfil the assumptions are described as well. Later on, Chart 2 will be relooked at to observe if there are any other emerging outcomes or even if any interventions are needed for the assessment moderation to be improved.

2.2 Assessment Moderation

Analysing moderation of assessment using Bamber's discretion framework of evaluative practices in institutions *(refer to chart 1)* provides the evaluator with a perspective to understand the control the institution has over assessment moderation practice. It seems that the governing body – in this case MQA – has delegated the responsibilities to the institution to carry out assessment moderation practice according to the institution's circumstances. The focus is tightly controlled, however, the process is not. Therefore, the institution has been given the autonomy to carry out moderation of assessment as it seems fit to the institution. Having the autonomy, the institution supposedly should be free in carrying out the assessment moderation practice without much restriction as long as the practice is evident and effective. This paper offers a critical analysis of assessment moderation practice of University A by using ToC. The intended outcome of assessment moderation in line with the assumptions and activities of the practice is analysed. Therefore, ToC not only helps to evaluate to the effectiveness of this evaluative practice but also helps to purposefully identify the pathway to change if it is necessary.

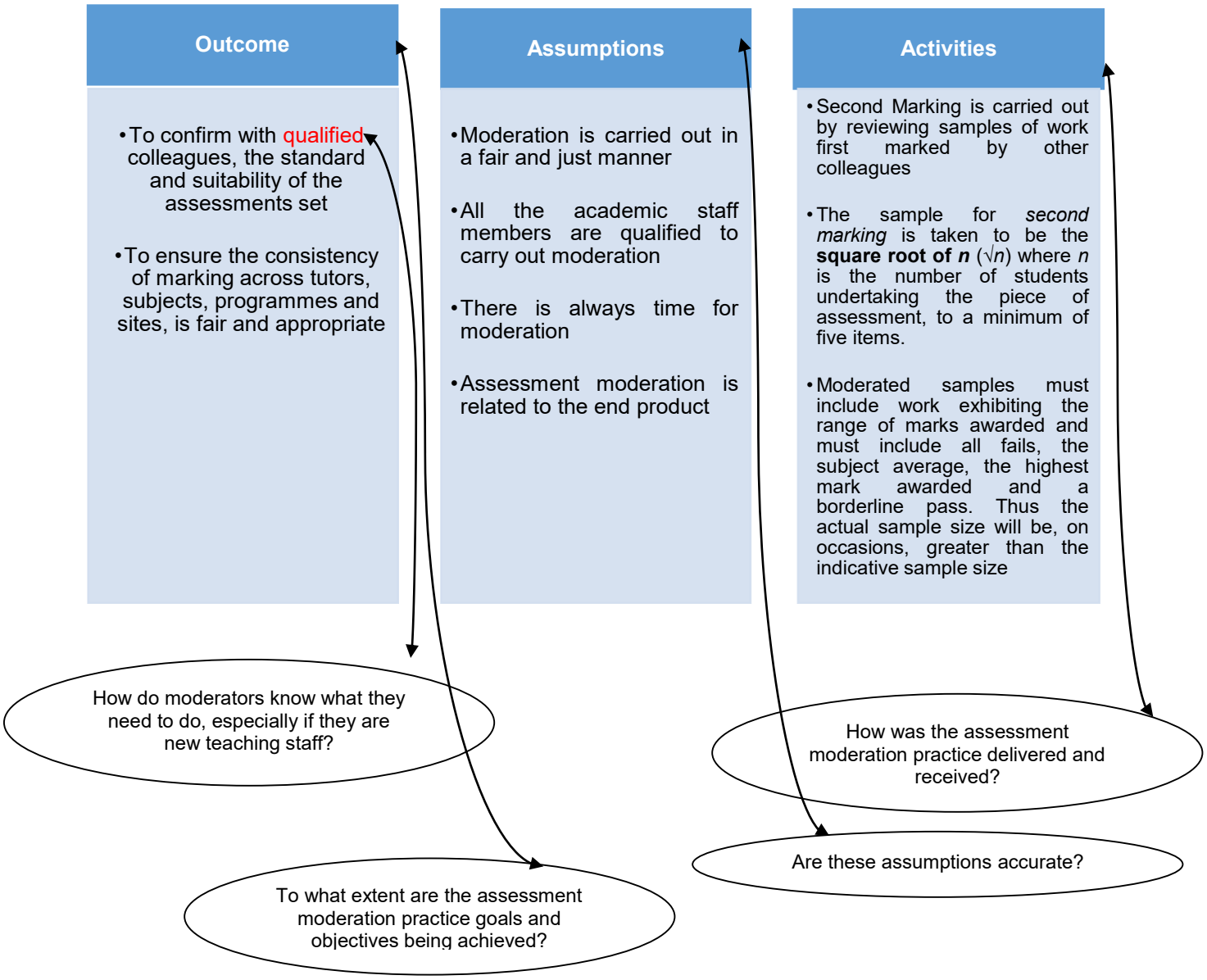

Figure 2: Theory of Change Model prior to the evaluation

3.0 METHODOLOGY

3.1 Research Method

The evaluation was administrated from August to September 2018. The evaluation assessed the assessment moderation practice of University A, particularly the Business School (SUBS) and the Centre for English Language Studies (CELS). This study employed a qualitative, case study approach to provide a comprehensive and holistic analysis of assessment moderation practice that takes place in these two departments. Yin (2018) highlighted that case study approach helps to obtain more extensive and in-depth description of a social phenomenon.

The researcher takes the perspective of interpretivist. Mack (2010, p.8) mentioned interpretivists believe that "knowledge is gained through personal experiences". This study attempts to capture the perspectives of different participants and focuses on how the participants' different meaning contributes to different perspectives of the topic. Although the benefits of the case study research have been mentioned, the concern that may arise is the generalisability of the research findings to the entire community. However, as Yin (2018, p.20) mentioned, "case studies are generalisable to theoretical prepositions and not to populations or universes". This study also falls under exploratory case study as this study can be seen as a

preliminary study that can be used to be tested in a larger scale, even evaluating all the departments in University A.

3.2 Data Collection

In this study, the participants were selected through purposive sampling. The participants should fall into any one of the following characteristics: 1) Must have completed assessment moderation for at least 10 rounds (have at least 5 years and above of experience in moderation), 2) Must have carried out assessment moderation once or twice (have less than one year of experience). The researcher obtained the information on academics that fit the criteria from programme leaders of each department. Once this was obtained, an email was sent to the academics to invite them to take part in the study. This email was accompanied by a participant information sheet, consent form and ethical approval letter from University A. Those who gave consent to be part of the study were then contacted for a semi-structured, face-to-face interview session. Each interview session lasted around 15-30 minutes and was undertaken at mutually convenient times. All the interviews were audio recorded and transcribed to analyse and identify themes.

The sample was composed of fourteen academic staff, seven of which were drawn from SUBS and the other seven were from CELS. Overall, eight participants fit the criteria of having completed more than 10 cycles of assessment moderation. Another six participants fulfilled the criteria of having carried out assessment moderation once or twice at University A. It is crucial to include the voices of new academic staff as well in this research to promote inclusivity. Out of the fourteen participants, four of them had responsibilities either as Heads of Department or as Programme Leaders. It is important to identify the responses of top management because they play a major role in managing the quality of assessment moderation practices.

3.3 Ethical Consideration & Data Analysis

Prior to the face-to-face interview session, participant information sheet, consent form and ethical approval letter from University A were emailed to all participants. This research has been approved by both University A (Malaysia) and Lancaster University (United Kingdom) research ethics committees. BERA (2011) guidelines were adhered to throughout the research by fully informing the participants the research details, providing the participants the rights to withdraw from the study and treating the data with confidentiality. The participants of the research were treated anonymously and were given a pseudonym. The interviews were audio recorded after obtaining consent and were subsequently transcribed. The transcripts were imported to the qualitative software package NVIVO (QSR International, 11). Keywords-in-context analysis of the interview was carried out to evaluate if there are any recurring themes to understand the depth of the issues discussed. The findings and discussions are presented according to common themes. Each theme is described in greater detail below, and is illustrated by interview quotations to support the discussions. The quotations are italicized for easier identification. The numbers within parentheses after the excerpts represent the code numbers of the interviewees.

4.0 FINDINGS AND DISCUSSION

In this section, the most common themes that emerged from the findings are discussed while answering all the research questions of this study as well.

4.1 Assessment Moderation Process

It was found that SUBS and CELS carry out their assessment moderation process slightly differently although there are some similarities. The similarities of the assessment moderation process are that both the departments only carry out second or double marking. Double marking is only carried out if the course carries 6 credits or more (e.g. final year projects). Besides, the examiner individually marks and then the sample which is calculated by the square root of the total scripts is given to the moderator. In this study, the power issue with second marking is observed, especially among new academic staff who tend to agree to the veteran first markers' marks when they moderate or are being moderated. This is obvious especially in the first few semesters of the new academic staff's tenure. When asked why the new staff thinks that her standard of marking is not at the appropriate level. One of the new academic staff's answers was:

"my second marker was an experienced lecturer" (1).

Having said that, through this evaluation, it was found that CELS and SUBS are bounded by very limited options of moderation, particularly focusing on second marking, when there are other types of moderation which can be explored. For instance, Beutel et al. (2017) highlighted:

> "in conference moderation, each member of a teaching team marks all assignments in his/her tutorial group or allocated assignments and then brings a representative of each grade or agreed benchmarks to a teaching team moderation meeting in which consensus for the grades is the focus".

This type of moderation may be ideal when the course is taught by many academic staff. At the same time, the discussions from conference moderation may also contribute to better shared understanding among moderators. The next section will highlight further on shared understanding which is one of the emerged themes of this study.

Moving on to the differences that was discovered between these two departments are the types of assessment that is moderated. CELS moderates all the assessments even if the weightage of the assessment is only 5%. However, SUBS sends its assessments for moderation only if the weightage is more than 20%. The partcipants from both departments were unsure why such differences occur. Of course, the participants from CELS highlighted this difference because they are the ones who are required to do more moderations of assessments:

> *"The first one that I've noticed recently is that we moderate all the assessments regardless of weightage which to me is a bit strange because the amount of work that is involved in moderation and the discussion I had recently was whether or not we need to moderate all assessments or we should just focus on time and energy really on those that carry a large percentage" (5).*
> *"Here they came up with policy where all graded assessments should be moderated, just for consistency purposes. There must be a reason for this to happen. So far, in my previous institution, we only moderated the final exam, which carries the highest percentage, the other you can if you are co-teaching. That is okay to do it. So I guess every institution has different policies. Only thing is we have to ask ourselves if it is necessary to go for all" (6).*

It can be pointed out that the differences occur because of the tightly controlled focus but loosely controlled process *(refer to Figure 1)* which refers to the 'relative control an institution may have on the way it undertakes its internal evaluative practice' (Bamber, 2011, p.129). According to the 2018 Academic Regulation document of University A, all final examinations and coursework accounting for 50% or more of the total subject has to be sent for moderation. The discrepancy that can be noticed between the practice and the policy actually has a positive outlook. This is because the departments that have been evaluated actually do more moderation than is required. This is reflected by the academic staff from SUBS who mentioned the following:

> *"I will always give more samples for moderation, giving more is better than giving less" (14).*

Although the moderation process within University A is loosely controlled, the number of samples that is being moderated by the departments are always more. However, for CELS more could be an issue in relation to time as how participant number five has highlighted earlier. Regardless of the department the participants belong to, 70% of the participants found time to be an issue in moderating assessments. Some of the comments of the participants are as follows:

> *"Personally, I find having to moderate exam paper you need careful attention, I need to really pay attention to what I am reading, because it concerns not only me but students also. And then the challenges in there in terms of writing the report, it takes time. I find it is the most challenging, having read all these, spotting all the errors, sometimes no errors at all, having to put it down in words, it is challenging" (3).*

> *"The biggest issue with moderating or second marking is always time" (11).*

> *"We could probably improve the process but I think it's very rushed. The timing… The markings are very rushed..." (12).*

Bloxham, Hughes & Adie (2016) highlighted that when there is little time available, as in university assessment timelines, the task of achieving appropriate, consistent and fair standards is enormously difficult. Other than that, Watty et al. (2014) highlighted that the teaching team that gives importance to greater shared understanding, could face time as an issue because it will take time for moderators to come to an agreement. However, this can lead to a greater agreement on grades and confidence in the alignment with established learning outcomes. Is that not the outcome that MOA intends to achieve? *(refer to Figure 2).* However, the bitter truth that has been revealed by one of the participants of this study with regard to time is the fact there could be a compromise in terms of quality:

> *"Sometimes the problem we face is the time is quick is tight. So after you finish creating the final exam papers and then you submit for the second moderation and then you have to start to prepare for the next semester coming in. So we are rushing and when you're rushing, sometimes you tend to compromise little bit of the quality doing the work. So I would say it's more about time constraints" (13).*

Therefore, in this case, a solution is needed to identify a better moderation practice that allows a better timeline approach. For this reason, it can be said that quality assurance could be related to time. One of the participants highlighted that sometimes, when discrepancy is observed between markers in assessment moderation process, it is only limited to that many sample scripts that is being moderated. Since there is time constraint, at times the entire cohort is not remarked which by right should be re-examined if quality assurance is taken into consideration. The following excerpt describes the participant's comment:

> *"So then I honestly think that this process is a bit futile because if you have hundred scripts and I'm only marking like ten if it's 10%, how would that benefit the entire cohort or maintain standardisation of marking whereas there should be a process where all this is done. Then, I think the Head, or whoever is in charge should meet, so that the committee will decide whether there is discrepancy here for all these students of this cohort. So I feel if you honestly ask me, it's a good process. Otherwise, it stops at the 10 scripts and it doesn't make any difference. Then I will say that the moderation process is ineffective. It does not serve its purpose" (7).*

Therefore, the suggestion given by participant 07 may be taken into consideration. Overall, there is still room for improvement in terms of the assessment moderation process carried out by University A. There is a need to identify further why there is a difference in samples of assessments that are being moderated according to the Academic regulation document of University A. Besides that, the suitable type of moderation assessment should also be chosen in relation to the timeline that is available more so if the department deals with a large cohort of students.

4.2 Training and Support

Another emerging theme from this study is training and support provided by the institution in terms of preparing the staff members for assessment moderation. If chart 2 is revisited, the outcome of assessment moderation is to confirm with qualified colleagues, the standard and suitability of the assessments set. The word 'qualified' has to be defined here in order for all the stakeholders to understand. It raises a point of consideration on how do the moderators know what they need to do. This evaluation intends to look at the support that the university provides for the academic staff in terms of assessment moderation. The responses of the participants are basically leaning towards negative reviews concluding that there is no proper training provided.

If proper training or support is not given to the staff members how can the institute measure that the assessment moderation practice is being carried out by qualified staff members. It can be said that the second marking process is one of the crucial aspects in achieving quality assurance. As such, this section is divided into two sub-categories under training and support.

4.2.1 Training and Support: New Academic Staff

Some of the participants mentioned that finding out what they need as a moderator is always an individual effort. It also appears that some Heads of Departments (HODs) take extra initiative in providing

some basic information in relation to assessment moderation. The following excerpts describe the participants' reflections:

> *"I actually asked personally, so I asked few questions before I started. I also took the subject file so I went through all of it and it helped." (1)*

> *"My personal experience has always been that as the head of department, What I do is try to brief them. In terms of what is the expectations and how you know what you should do. And personally I'm not sure whether we have any proper briefings or training in terms of exactly how you do the moderation. And I also don't think that the moderation document itself really guides people to look at key issues or rather than issues. So yeah. So you asked me whether people are aware how to do it. I'm not surprised if they are not sure. So a lot of it is as you do You sort of like and try there and sort of figure it out" (9).*

From the responses above, it can be noticed that the participants realise that there is no proper training and it always falls back on the individual's effort in asking questions or some HOD's initiative of providing information using their own effort. It is discouraging to realise that the institution does not provide some form of formal training relating to assessment moderation. One of the participants mentioned that training may not be possible because it involves many different subjects:

> *"I think it's quite difficult to have a training programme to standardise because it's different for every subject" (12).*

This may be true but it brings us to the question of how the institutional system is expecting the academic staff to know what they need to do. Is it just because they have some teaching experience for some time elsewhere? Of course, it is an advantage if the new teaching staff have experience in assessment moderation but what about understanding the system within the institution itself? According to Beutel et al. (2017), in order to gain a better shared understanding, the new staff has to purposefully inducted to the moderation of assessment process. The training and support becomes more and more important when part-timers are involved in the teaching system. One of the participants highlighted that this is a major issue in achieving standardisation of marking and moderation when the assessments are marked by part-time tutors but not the full-time examiners:

> *"It's going to be difficult and I'm going to probably have to hand it to a tutor to mark and I'm already struggling with the thought of how I'm going to tell her what I'm looking for and the rubrics are quite detailed that because they're all individual papers and there's so many different things… That's a big problem for students and I think the friction between lecturers and tutors is quite a big issue really and lots of just tend to be part time. You don't see so much of them. So the whole question of how you manage and standardise them…" (12).*

Therefore, the findings of this section lead the reader to the next category which is: what is then the proper way of support?

4.2.2 Training and Support: Shared Understanding

As mentioned earlier, the findings also revealed that there is a need for greater shared understanding between moderators and markers. Identifying the perception of the first marker is very crucial. The following excerpts explain that:

> *"I find it difficult if the internal moderator is not my area of expertise. So I have difficulties when the internal moderator just merely relies on answer scheme. So they might penalise the students unnecessarily. It's not fair" (10).*

> *"But when I was teaching on a programme that was moderated by UK University, this university was good in the sense that they included, I guess as a form of professional development for all the lecturers who were teaching on the programme. So what they did was they explained the philosophy*

of assessment for the University. They also talked about how for example year one modules are created and get to year three and how the assessments should be similar in a sense where there is continuity and increasing in difficulty, and then we actually had I think two sessions with senior lecturers and programme leaders from the University where they talked about how they use rubrics for example and how they then would moderate. So if this is a piece of work, we will actually sit down and mark it together as in time and have a discussion about why marks were given in a certain way. So to me that was interesting because that was more formalised process and that helped because then when you actually do the assessments you're marking assessments they are some guidelines and there's if there's any variation, it's less right?" (5).

"I believe different individuals will interpret differently, because I notice that some of us will have different expectations, so I guess the expectations are different way seeing things so maybe lecturer A will say that oh, I think we should give this mark to the students so I guess we establish the discussion in the very beginning before we do the marking scheme and so on then we are having things like in the marking scheme, like point to be taken. so if we put that done, that is easier for us" (6).

From the responses above, establishing a discussion among moderators and first markers is crucial. This will definitely help in achieving greater shared understanding among staff members. This is in view of Wallace et al. (2010) who suggested that moderation is not something that happens at the end of the assessment process but something that starts from the planning stage. It is also crucial to ensure that the moderators have understood the criteria involved so that discussion can be easily established.

As highlighted by participant 08 and 10, some moderators merely depend on the answer scheme as a point of judgement. However, they feel sometimes that other answers can also be accepted as long as students are able to justify their answers. The difference of perception occur because a less proficient marker may rely more on the marking criteria while the experienced marker may make more holistic judgements (Sadler, 2013). Such discrepancies can be avoided if both moderators and markers can be involved in the planning stage of assessment where more opportunities can be given for discussion. Active engagement of the academic staff in creating assessment criteria facilitates greater shared understanding instead of just being given the criteria to follow (Rust, O'Donovan, and Price, 2005).

This leads to another dimension of the findings of this paper. Some participants mentioned about introducing a formal buddy system which may help in achieving greater shared understanding. The following excepts describes the participants' ideas:

"The buddy system is lacking is under the HR department. I do notice is we don't have a checklist, because the checklist plays an important role for confirmation. With that checklist right, we are given mentor-mentee thing to look at, to say we have covered these things. the moderation can go in. The training part... I do see a gap there" (6).

"Some universities have a buddy system or mentor-mentee system. But unfortunately in this university I didn't get a buddy system. I mean to me I just rely on my own experience. So what I learned from here just pick it out from there but we don't really have a mentor system. Does it help to have a mentor? I'm not sure. I think to me it is the same thing across the board but maybe will help from younger groups" (8).

Buddy system may help to formalise the training and support received by academic staff members particularly the new staff and part-timers when involved in the moderation process. In the long run, proper measuring steps can be formalised if learning is taking place as participant number six suggested. Besides that, more discussions can be established for greater understanding between moderators and markers.

In summary, the findings of this study reveal that moderation of assessment is a complex process because the effectiveness of this evaluative practice is mainly influenced by the knowledge attained by carriers of this practice as well as the institution's choice of moderation method which eventually results in academic staff to be under tight timeline. This brings forth a point of discussion whether moderation is carried out for quality enhancement or is it just another compulsory task by the institution that needs to be completed. Since this practice is complex, it is only appropriate to justify the changes using ToC. Chart 3

shows a number of possibilities to achieve the objective of moderation of assessment. Special attention should be given to the fact that the outcome remains the same. This reflects the tightly controlled focus this evaluative practice has *(refer to Figure 3).* However, in order to achieve the outcome, the activities leading to the assumptions have to be changed. Connecting to Bamber's Model *(refer to Figure 3)*, the institute is free to change the process on condition the outcome is achieved. This study prompts University A to be creative with its process of assessment moderation.

In order to achieve the desired outcome, firstly the moderation of assessment need to be carried out in a just and fair manner. Therefore, there is a need for greater shared understanding between the moderators and markers. Besides that, since the institution is now compelled to move from its comfort zone, it has to explore other types of moderation in order to achieve quality in the process. Secondly, if it is assumed that all the academic staff members are qualified to carry out moderation then proper training and support should be provided by the institution. The support could even be formalised support like buddy system. Thirdly, it is only appropriate to begin moderation process at the planning stage of assessment development. This encourages more discussions among teaching team to achieve greater shared understanding.

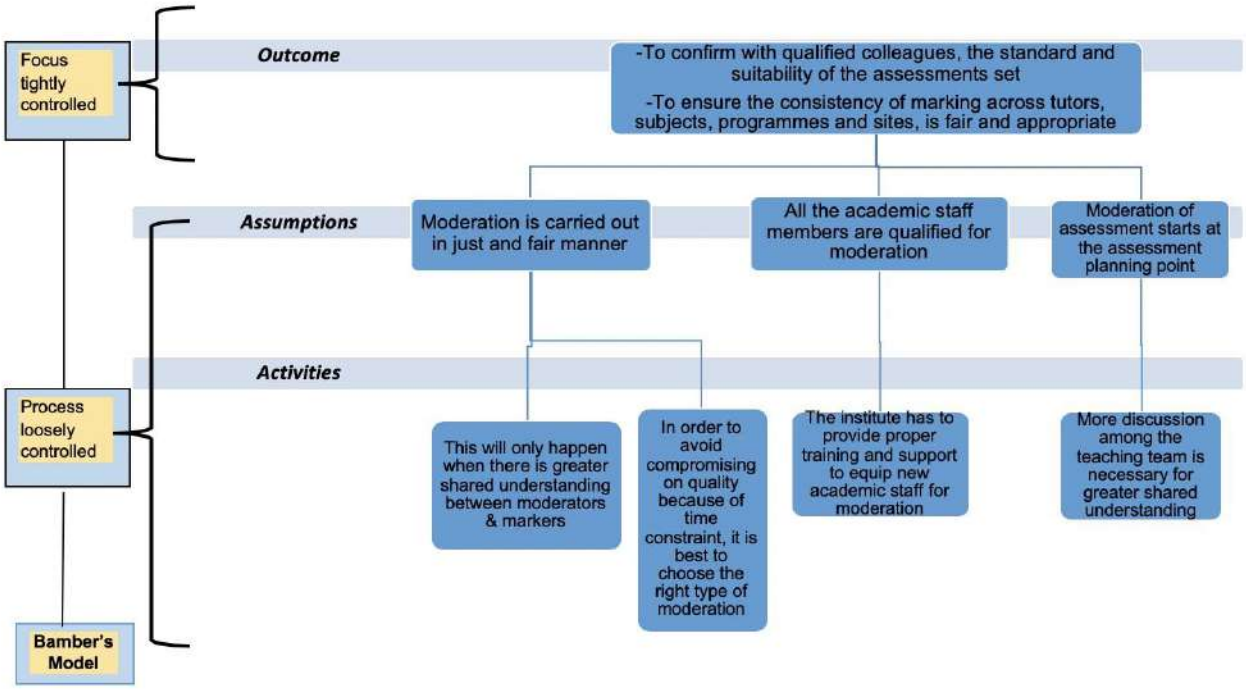

Figure 3: Linking Bamber's Model to the proposed Theory of Change Model

5.0 CONCLUSION

This study aimed to assess assessment moderation practice in order to identify and depict assessment moderation practice, elicit views on its effectiveness and assess the extent to which improvements might be sought using ToC. Limitations are acknowledged in that this study was confined to only two departments within one institution, and therefore cannot be representative of the whole system. However, the findings add further weight to the existing body of literature on moderation practices in HE institution by incorporating the voice of new teaching staff members in this evaluation.

One of the focuses of this evaluation is to identify if University A provides the right support especially for the new teaching staff members. It was found that the new academic staff members have to be equipped with proper support either through training or buddy system for greater shared understanding between moderators and examiners. Only then the objective of the moderation can be achieved. The next focus of the evaluation is to understand the assessment moderation process of University A. It was found that the type of assessment moderation has to be re-evaluated by the institution. There is no need to just rely on second marking or double marking as stated in Academic Regulation of University A. Since the

process of this evaluative practice is loosely controlled, the academic staff can be innovative in choosing the appropriate assessment moderation mechanism that suits their environment. The decision could be made based on the number of students in a cohort, the involvement of sessional staff or even by taking into consideration the timeline available. One of the recommendations proposed is to carry out further research in the area of identifying the appropriate assessment moderation method especially if the moderation process involves a large cohort and a tight timeline approach.

This exploratory case study which is evaluated using ToC certainly proves that assessment moderation practice is a complex initiative. Chart 2 is revisited with some new possibilities in achieving the objective of moderation which is shown in Chart 3. According to Connell & Kelm (2000) the benefit of using ToC in an evaluation is the fact that the stakeholders are able to receive meaningful continual feedback from evaluators. In view of this, the findings that emerged from this study can be continuously used for the future improvement of the institution and the academic community.

Referring to Figure 3, the initial ToC model is re-examined with improved activities in order to achieve the initial outcome. By providing a clear and justifiable ToC, it is only beneficial for all the stakeholders because they will be able to understand the change needed. In this case, the Academic Standard Quality department of University A, SUBS and CELS will be informed of the need for such changes for continuous quality improvement.

As a final point, it can be concluded that this evaluation is not an end by itself but rather a beginning to an improved future. Having said this, it is hoped that the changes suggested in this paper will be implemented by the institution and then subsequently the proposed ToC can be explored further for future improvement and intervention. This is because, evaluation should be ongoing and continuous for the purpose of quality enhancement particularly in the area of moderation of assessment in order to maintain the reliability, validity and consistency of assessment grading. *Be not afraid of going slowly, be afraid only of standing still. – Chinese Proverb.*

REFERENCES

Annetts, S., Jones, U. and Van Deursen, R. 2013. An innovative review of an undergraduate dissertation double marking policy. *Innovations in Education and Teaching International 50*(3), 308-317. DOI: 10.1080/14703297.2012.760771

Bamber, V. (2011). *Institutional evaluative practices: quality, enhancement and regulation.* In: *Reconceptualising evaluation in higher education: The practice turn* (pp. 127-132). SRHE and Open University Press: Maidenhead, ISBN 9780335241606

Beutel, D., Adie, L. & Lloyd, M. (2017). Assessment moderation in an Australian context: Processes, practices, and challenges, *Teaching in Higher Education, 22*(1), 1-14, DOI: 10.1080/13562517.2016.1213232

Beutel, D., Adie, L. & Lloyd, M. (2017) Assessment moderation in an Australian context: processes, practices, and challenges, *Teaching in Higher Education, 22*(1), 1-4, DOI: 10.1080/13562517.2016.1213232

Bloxham, S. (2009). Marking and moderation in the UK: False assumptions and wasted resources. *Assessment & Evaluation in Higher Education, 34*(2), 209–220. DOI: 10.1080/02602930801955978

Bloxham, S. and Boyd, P. (2007). *Developing effective assessment in higher education: A practical guide.* Maidenhead: Open University Press

Bloxham, S., Hughes, C. & Adie, L. (2016). What's the point of moderation? A discussion of the purposes achieved through contemporary moderation practices. *Assessment & Evaluation in Higher Education, 41*(4), 638-653, DOI: 10.1080/02602938.2015.1039932

Bloxham,S,., Hughes, C. & Adie, L. (2016). What's the point of moderation? A discussion of the purposes achieved through contemporary moderation practices. *Assessment & Evaluation in Higher Education, 41*(4), 638-653, DOI: 10.1080/02602938.2015.1039932

British Educational Research Association. (2011). *Ethical guidelines for educational research*. London: BERA. Retrieved from https://www.bera.ac.uk/wp-content/uploads/2014/02/BERA-Ethical-Guidelines-2011.pdf

Clark, H., & Anderson, A. A. (Nov, 2004). *Theories of Change and Logic Models: Telling Them Apart*. Paper presented at the annual meeting of the American Evaluation Association, Atlanta, Ga. Retrieved from http://www.theoryofchange.org/wp-content/uploads/toco_library/pdf/TOCs_and_Logic_Models_forAEA.pdf

Grainger, P., Adie, L. & Weir, K. (2016). Quality assurance of assessment and moderation discourses involving sessional staff. *Assessment & Evaluation in Higher Education, 41*(4), 548-559. DOI: 10.1080/02602938. 2015.1030333.

Holroyd, C. (2000) Are assessors professional? Student assessment and the professionalism of academics, *Active Learning in Higher Education, 1*(1), 28–44. DOI: 10.1177/1469787400001001003

Mack*, L. (*2010). The Philosophical Underpinnings of Educational Research*. Polyglossia, 19, 5 – 11. DOI: 10.5539/elt.v5n9p9*

Malaysian Qualifications Agency. (2013). *Guidelines to good practices: Assessment of students*. Retrieved from http://www2.mqa.gov.my/QAD/garispanduan/2013/ENGLISH%20GGP%20AS%20061113.pdf

Msila, V. & Setlhako, A. (2013) Evaluation of Programs: Reading Carol H. Weiss. *Universal Journal of Educational Research, 1*(4), 323-327, DOI: 10.13189/ujer.2013.010408

Rust, C., B. O'Donovan, and M. Price. (2005). A social constructivist assessment process model: How the research lliterature shows us this could be best practice. *Assessment and Evaluation in Higher Education 30*, 231–240. DOI: 10.1080/02602930500063819.

Sadler, D. R. (2013). Assuring Academic Achievement Standards: From Moderation to Calibration. *Assessment in Education: Principles, Policy & Practice*, *20*(1), 5–19. DOI: 10.1080/0969594X.2012.714742

Sunway University. (2018). *Sunway University Academic Regulation.* Academic Standards & Quality. Retrieved from https://vine.sunway.edu.my/vinecms/sites/default/files/APQA_67P3XJ5SQ7Rw6wVv2/Academic%20Regulations%202017-V3.pdf

Wallace, M., Mahmud, S., Sanderson, G., Yeo, S., Briguglio, C., and Hukam-Singh, P. (2010). "Moderation of Assessment in Transnational Higher Education." *In Collaborative Approaches to Moderation of Assessment in Transnational Education* edited by Gavin Sanderson, 1-11. Melbourne: RMIT University.

Watty, K., Freeman, M., Howieson, B., Hancock, P., O'Connell, B., de Lange, P. & Abraham, A. (2014). Social moderation, assessment and assuring standards for accounting graduates. *Assessment & Evaluation in Higher Education*, *39*(4), 461-478, DOI: 10.1080/02602938.2013.848336

Weick, K. (1995). *Sensemaking in Organizations*. Thousand Oaks, CA: Sage.

Weiss, C. H. (1972). *Evaluation Research: Methods of assessing program effectiveness*. Englewood Cliffs, N.J.: Prentice-Hall.

Weiss, C.H. (1998). *Evaluation: Methods for studying programs and policies*. (2nd ed.). Upper Saddle River, N.J.: Prentice-Hall.

Yin, R. K. (2018). *Case study research: Design and methods* (6th ed.). Thousand Oaks: Sage Publications.