

FORMATIVE ASSESSMENT IN AN INTERNATIONAL SCHOOL

# FORMATIVE ASSESSMENT IN AN INTERNATIONAL SCHOOL

## Policy, Attitudes & Practice

Xenia Iacob

1454699

ITEps NHL Stenden University

Grada Okken

Word count: 10242

## Abstract

Formative assessment has been rapidly gaining world-wide popularity due to its widely acknowledged potential to enhance student learning. However, change in education has often proved to be slow and impeded by many factors and studies show that formative assessment often fails to yield the promised results, most likely due to poor implementation. Therefore, this case study aimed to further the current understanding of the complexities of implementing formative assessment by exploring the alignment between (i) the school policy, (ii) teachers' attitudes and (iii) their practices (both self-reported and observed). A mixed methodology and various instruments (questionnaire, interviews, observations) was employed to collect in-depth, qualitative and quantitative data on the three variables. Data was collected and analysed inductively based on a list of 7 principles for the implementation of formative assessment which emerged from the literature review:

1. Learning intentions and success criteria
2. Frequency
3. Feedback
4. Eliciting data
5. Interpreting data
6. Classroom culture
7. Teacher attitude, competency, collaboration and development

The results revealed that the three variables aligned with regards to principles 1, 3 and 4. Two additional themes emerged from the policy which did not align with the other 2 variables. Based on the findings, a set of recommendations was provided to maximize the alignment of the variables within the school, which might also prove relevant to other contexts.

**Keywords:** *formative assessment, Assessment for Learning, teacher attitudes, policy, practice*

# FORMATIVE ASSESSMENT IN AN INTERNATIONAL SCHOOL

## Table of Contents

Introduction.....	1
Rationale .....	3
Aim .....	3
Motivation.....	3
Research Questions .....	4
Significance.....	5
Literature review .....	7
I.    What is formative assessment? .....	7
Origins.....	7
Definitions.....	7
Principles for implementation.....	9
II.   What is the case for formative assessment in primary education? .....	14
Research design .....	16
Methodology .....	16
Paradigm .....	16
Setting .....	16
Participants.....	16
Ethical issues.....	17
Data gathering tools .....	18

## FORMATIVE ASSESSMENT IN AN INTERNATIONAL SCHOOL

Document analysis .....	18
Questionnaire .....	19
Observations .....	20
Interviews.....	20
Data analysis .....	21
Questionnaire .....	21
Interviews.....	22
Observations .....	22
Quality.....	23
Results.....	25
Policy .....	25
Attitudes.....	27
Practices .....	29
Discussion .....	34
Conclusion .....	39
Recommendations.....	40
Limitations .....	42
Appendices.....	49
Appendix I – Key Terms.....	49
Appendix II – Interview Questions.....	50

## FORMATIVE ASSESSMENT IN AN INTERNATIONAL SCHOOL

Appendix III – Observation Form.....	51
Appendix IV - Attitude maps.....	52

## **Introduction**

The rapid pace of change across the spectrum of society is reflected in education and consequently a myriad of reforms are being adopted world-wide (Priestley & Sime, 2005). With the rise of constructivist learning theories, increasing emphasis is being put on student-centred approaches (Taylor, 2017), yet the economic pressure and expectations of schooling have also led to a growing standardisation of curricula, teaching and assessment (Robinson & Aronica, 2015). As a reaction to this standardisation tendency, more student-centred approaches to learning, teaching and particularly assessment emerged, one of which is formative assessment, aiming to extend the function of assessment from merely measuring, to enhancing learning (Miller, 2015).

In practice, formative assessment is a form of assessment which happens during instruction and aims to (i) inform teachers and students on learning progress, and (ii) teachers on how to adapt instruction to meet students' needs (Andrade & Cizek, 2010; Heritage, 2010). It then appears that two conflicting trends -on the one hand a standardisation of assessment and, on the other hand, expectations to use student-centred assessment approaches- are unfolding simultaneously and teachers, who are ultimately the enactors of any such reform (Vähäsantanen, 2015), must efficiently integrate this apparent dichotomy in their daily practice (Cheng & Yan, 2015).

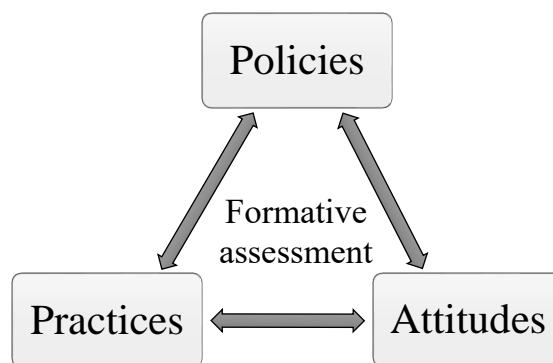
The implementation of formative assessment, like any other reform, involves the interplay of three major factors: the policy underpinning the reform, teachers' attitudes regarding the reform, and teachers' practice in the classroom and in the school. To ensure successful implementation, it is vital that these factors align. When educational reforms are enacted on a top-down basis, with little consideration for teachers' role, agency and beliefs, as it often

happens, their success is jeopardised (Hasim, Di & Barnard, 2019; Vähäsantanen, 2015).

Teachers' beliefs, attitudes, and professional identity are considered to be predicting factors of whether a reform will be (successfully) implemented (Cheng & Yan, 2015; Fives & Gill, 2015).

Teachers' beliefs are known to act as an interpretative filter, transforming curricular intentions developed elsewhere (Bryan, 2012). Sustainable implementation of any educational change requires change beyond surface structures or procedures, focused on altering the knowledge, skills, attitudes and beliefs of all stakeholders, particularly teachers (Heitink, Van der Kleij, Veldkamp, & Schildkamp, 2016). Consequently, when teachers' beliefs are not aligned with the beliefs underpinning a reform, its successful implementation is at risk (Bryan, 2012). In the case of formative assessment, teachers' beliefs of learning should reflect constructivist views (Heitink et al., 2016). Moreover, teachers' supportive attitude towards formative assessment is known to predict their actual intentions to implement it and the level and quality of their implementation (Heitink et al., 2016).

Therefore, the present study aims to help broaden the current understanding of the complex process of implementing formative assessment by exploring the alignment between (i) the school policy, (ii) teachers' attitudes and (iii) teachers' practices in an international, primary school in Malaysia (see *Figure 1*).



*Figure 1. Conceptual model of variables alignment*

## **Rationale**

### **Aim**

This study aimed to explore the alignment between the school policy, teachers' attitudes and teachers' practices regarding the implementation of formative assessment in an international, primary school in Malaysia. Moreover, based on the findings, a list of recommendations was provided to maximize the alignment for the successful implementation of the reform.

### **Motivation**

Two main aspects led the researcher to the decision to commence this study. Firstly, a personal interest in the areas of teacher identity and its relation to change, particularly educational reforms. Having conducted a previous study on international teachers' identity formation, the researcher has gained substantial knowledge and a personal fascination regarding its role in shaping classroom practice. Therefore, this study was a natural next step in a journey of increased understanding and awareness of the impact of a teacher's identity, with all its facets (e.g. beliefs, attitudes, values, self-image etc.) on the classroom practice.

Secondly, this study's heuristic value was enhanced by the fact that the researcher was herself a pre-service teacher, soon to be teaching her own class and to make sense of such challenges as educational reforms and conflicting beliefs and values. Understanding how other, more experienced teachers deal with such challenges on the one hand, and acquiring the knowledge and vocabulary to reflect on, discuss and manage educational change on the other hand, were seen as essential pre-requisites in the researcher's toolkit to positively and successfully deal with this important part of a teachers' life.



Lastly, formative assessment is increasingly being used internationally (Birenbaum, et al., 2013), hence a sound understanding of the concept, both theoretically and in practice, were deemed as a valuable asset for a beginning teacher to effectively implement it and thereby support children's learning.

## Research Questions

This study aimed to answer the main research question:

*How do school policies, teachers' attitudes and teachers' practices align in the implementation of formative assessment in an international, primary school in Malaysia?*

To effectively answer the main question, the following sub-questions were addressed:

- (i) *What is formative assessment?*
- (ii) *What is the case for formative assessment in education?*
- (iii) *How is formative assessment formulated in the school policy?*
- (iv) *What are teachers' attitudes towards implementing formative assessment?*
- (v) *How do teachers implement formative assessment in their teaching?*

The first two sub-questions were explored and answered in a thorough Literature review following this chapter, where the key term *formative assessment* was defined at length, given its complexity. The remaining sub-questions were examined in a field study and answered in the Results, Discussion and Conclusion chapters.

The three remaining key words in the main question -*policy*, *attitudes* and *practice*- were defined as follows.

Rokeach (1969) defines an *attitude* as “a relatively enduring organization of beliefs around an object or situation, predisposing one to respond in some preferential manner” (p. 112). Attitudes are held towards particular things referred to as ‘attitude objects’ and are evaluative in

the sense that they reflect the degree of positivity or negativity a person feels towards that ‘attitude object’ (Eaton & Visser, 2008). Attitudes have the potential to shape our interpretations and motivate, guide and predict our behaviour (Eaton & Visser, 2008).

According to Cambridge Dictionary (n.d.), a *policy* is “a set of ideas or a plan of what to do in particular situations that has been agreed to officially by a group of people, a business organization, a government, or a political party”.

*Practice* is defined as “the actual application or use of an idea, belief, or method, as opposed to theories relating to it” (Oxford Dictionaries, n.d.).

## **Significance**

This study examined the three central factors in any educational reform: the policy, the teacher and the classroom practice. Most studies focus merely on one or two of these factors (Cheng & Yan, 2015) yet this study acknowledged that analysing their interaction within reform implementation, especially of the role of teachers’ attitudes, presents the potential to deepen our understanding of the complexities of implementing formative assessment.

Moreover, despite its largely acknowledged potential to improve student outcomes, formative assessment has been repeatedly found to fail to do so due to its often ineffective implementation (Hendriks, Scheerens, & Slegers, 2014). Therefore, to meet the commitment to improve students’ learning and teachers’ teaching, it is imperative to understand the multiple facets of implementing formative assessment on a school level in order to unlock its potential. Hence, this study followed the recommendation of Heitink and colleagues (2016) to focus on the *context* in order to understand the complexities of implementing formative assessment.

This study had the potential to further contribute to the existing epistemology as it examined both teachers' reported as well as observable practice while considering the contextual factors involved, which was a missing piece in the puzzle of formative assessment implementation (Cheng & Yan, 2015).

Lastly, the results had the potential to benefit the school (teachers, students and leaders) which acted as the setting of this study and which had been focusing on implementing formative assessment; the comprehensive overview and discussion of the data could be used to evaluate the current implementation and consider areas of improvement.

## **Literature review**

This chapter provides a critical overview of the current literature and research concerning formative assessment. The first sub-section concerns the origins, definitions and conceptualisations of formative assessment, followed by a discussion of the rationale for implementing formative assessment in education in the second sub-section.

### **I. What is formative assessment?**

#### **Origins**

The term ‘formative evaluation’ was first proposed by Scriven in 1967 and it was used to describe the role of educational programme evaluation, as distinct from ‘summative evaluation’ (Bennett, 2011). Scriven maintained that summative evaluations provided information to judge the overall success of an educational programme, whilst the results of formative evaluations were used to improve the programme (Bennett, 2011). Two years later, Benjamin Bloom made the same distinction regarding the evaluation of student performance (William & Leahy, 2015). Ever since, although formative assessment gained a great deal of popularity and is now being implemented across the educational system in many countries (Birenbaum, et al., 2013), scholars have yet to reach a consensus on its definition (Dunn & Mulvenon, 2009).

More recently, a new term has emerged: Assessment for Learning (AfL) and some suggest it differs from formative assessment in that it places more emphasis on students’ active role in assessment. Nevertheless, William (2011) maintains that precise terminology is irrelevant, and the two terms continue to be used interchangeably.

#### **Definitions**

According to William (2011), formative assessment is generally defined as a *process*. Kahl (2005,) however, defines it as a “tool that teachers use to measure student grasp of specific topics

and skills they are teaching” and to “identify specific student misconceptions and mistakes while the material is being taught” (p. 11). Although much of the literature on formative assessment is concerned with demarcating it from summative assessment (Dunn & Mulvenon, 2009), more recently scholars agree that any assessment can be used both summatively as well as formatively (William, 2011). What distinguishes then one form of assessment from the other are not the *tools* used or the *processes* involved, but the *purpose* to which the data resulting from any assessment is put (William, 2011).

In their seminal work, Black & William (1998) defined formative assessment as “encompassing all those activities undertaken by teachers and/or by their students, which provide information to be used to modify the teaching and learning activities in which they are engaged” (p. 8). However, William (2011) later identified a possible flaw in this definition, in that data does not always have to result in changes in the instruction: for example, analysing assessment data, the teacher might, in some cases, conclude that all students are making the desired progress and that the choices the teacher has made were the right ones; hence no changes are required yet (Heritage, 2010; William, 2011). Therefore, William (2011) proposes a new, now widely popular definition:

an assessment functions formatively to the extent that evidence about student achievement is elicited, interpreted and used by teachers, learners or their peers to make decisions about the next steps in instructions that are likely to be better, or better founded, than the decisions they would have made in the absence of that evidence. (p. 43)

Although many scholars advance the idea that the purpose of formative assessment-to continuously inform instruction- is what distinguishes it from summative assessment-its purpose being to inform stakeholders on students' mastery- (Andrade & Cizek, 2010), Dunn and Mulvenon (2009) warn against such a viewpoint. Accordingly, they state that what truly differentiates formative from summative assessment is the methodology, data analysis and the use of the results (Dunn & Mulvenon, 2009). They advance the idea, supported by others as well (e.g. William, 2011) that summative assessment can also serve formatively, to inform instruction (e.g. results from mid-term assessments can determine teachers to adapt their planning for the next term). Moreover, the 'purpose' narrative also absolves summative assessment of any role in supporting student learning as its purpose is seen to merely measure it; this view is discharged by the studies proving the benefits of summative testing on student achievement (Dunn & Mulvenon, 2009).

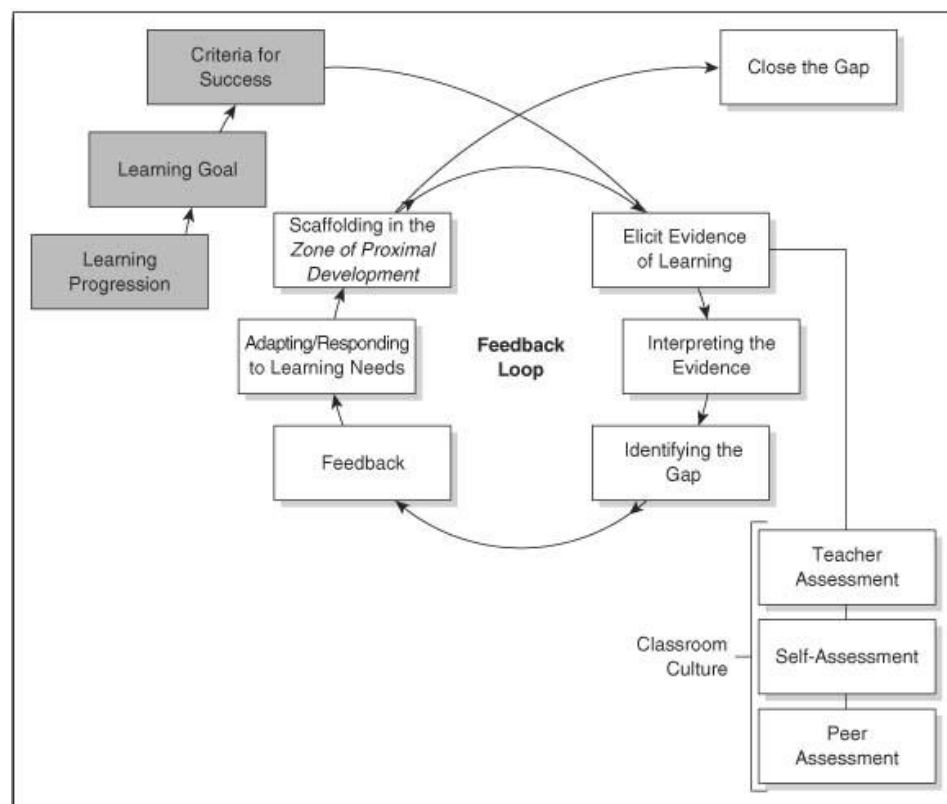
Despite the plurality of definitions, one can conclude that formative assessment happens during instruction and that it informs teachers and students about current learning progress and next steps. In distinguishing formative assessment from summative assessment, the former can generally be thought of as prospective, whilst the latter is retrospective (Heritage, 2010).

### **Principles for implementation**

There are several principles which guide the implementation of formative assessment. Firstly, like any other form of assessment, it should be *valid*, i.e. measure what is intended to be measured (Murchan & Shiel, 2017). When this is not the case, the inferences drawn from it are likely to be erroneous and lead to inappropriate decisions (Heritage, 2010). Secondly, *reliability* must be ensured, in other words: how consistently does an assessment measure what is intended to measure, i.e. are the results repeatable? (Heritage, 2010). No inconsistencies should result

from changes to *how* the assessment is carried out (e.g. time of administration, day and time of scoring, who scores the assessment etc.). Thirdly, all assessment practices should be fair towards all students (Murchan & Shiel, 2017).

Alongside these general principles, there is a number of other principles which specifically address formative assessment. Heritage (2010) proposed a comprehensive conceptualisation of formative assessment in practice (see *Figure 2*), which bears a close resemblance to that advanced by Leahy, Lyon, Thompson, & William (2005). From these two models, the following list of principles of implementing formative assessment has resulted: learning intentions and success criteria, frequency, feedback, eliciting data, interpreting data and classroom culture. Additionally, taking account of the extensive research (Heitink, et al., 2016; Heritage, 2010; William & Leahy, 2015) emphasising the role of teacher attitudes, competency, collaboration and professional development, this has been included as a 7<sup>th</sup> principle. Each principle is elaborated upon in this sub-section.



*Figure 2. The process of formative assessment (from Heritage, 2010, p. 11)*

### *1. Learning intentions and success criteria (LISC)*

Clear learning intentions and success criteria are considered the starting point of formative assessment in practice. Teachers should share these with their students and ensure they are well understood, but also that students are convinced of the importance to achieve the set learning goals (Heritage, 2010; William & Leahy, 2015). Moreover, teachers should also encourage and guide students to set their own goals. Once the learning intentions and success criteria are understood by both teacher and students, formative assessment can be planned and carried out to identify where the students are at a given moment and how to proceed from there to reach the learning goals. For this reason, many refer to the role of formative assessment as identifying and helping to ‘close the gap’ between children’s current level and the target one (Heritage, 2010; William & Leahy, 2015).

### *2. Frequency*

Frequency refers to how often formative assessment is carried out in the classroom. William and Leahy (2015) distinguish between three types of formative assessment depending on frequency: long-cycle (termly), medium-cycle (weekly) and short-cycle (daily, minute-by-minute). There is widespread agreement that, were formative assessment to truly improve instruction and learning, it is essential it is implemented ‘minute-by-minute, day-by-day’ (Heritage, 2010; Marzano, 2006; William, 2011)

### *3. Feedback*

Feedback given by teachers to their students and by students to their peers is intrinsically linked to the practice of formative assessment (Heritage, 2010). In 1992, John Hattie found that feedback was the most powerful modification to enhance learning (Marzano, 2006). However, feedback has been proven to be harmful when it does not fulfil certain criteria or is shared with



the students in ways which are negative or discouraging (Heritage, 2010; Marzano, 2006). In contrast, feedback given by students to their peers is reported as a means to increase motivation and enhance learning (Marzano, 2006).

#### *4. Eliciting data*

Teachers must carefully plan and carry out activities which will elicit data on students' progress towards the specific learning goals and criteria previously set. To ensure validity, teachers should have a very clear image of what the formative assessment activities aim to reveal (William & Leahy, 2015). Moreover, they should involve the students in this process by planning and implementing self- and peer-assessment during the lessons (Heritage, 2010; William & Leahy, 2015), as this will lead to students' increased motivation and ownership over their learning.

#### *5. Interpreting the data*

The data resulting from frequent, valid, fair and reliable assessment practices must be interpreted and used by teacher and students alike. For this interpretation to result in appropriate, correct inferences, it is essential that the success criteria have been clearly shared with and understood by all students (Heritage, 2010). Furthermore, teachers should also seek to use corroborating evidence for the inferences they make, instead of relying on just one piece of data (Heritage, 2010). Teachers can also increase the quality of their inferences by asking colleagues to interpret parts of the data and check for confirmation. It is imperative that the interpretation of assessment data is as accurate as possible, otherwise the basis for adjusting instruction is weakened and the practice of successful formative assessment jeopardised (Bennett, 2011).

### 6. *Classroom culture*

Feedback is central to formative assessment, yet it can only be valued if the teacher establishes a classroom culture which is conducive to giving and receiving feedback (Heritage, 2010). In practice, that means a re-distribution of power: from being held solely by the teacher, power is shared with the students and is accompanied by responsibility for learning (Heritage, 2010). The classroom must be a 'safe' place where students can ask for help, admit mistakes without fear and regard their mistakes as opportunities for learning (Heitink et al., 2016; Heritage, 2010; William, 2011). Lastly, relationships in the class should be supportive and collaborative and be built on mutual trust (Heitink et al., 2016; Heritage, 2010). Teachers are responsible with establishing values, standards and practices which will lead to such a classroom culture (Heritage, 2010; William, 2011).

### 7. *Teacher attitude, competency, collaboration and development (TACCD)*

Successful implementation of formative assessment, like any other educational approach or reform, ultimately depends on teachers' positive attitudes and willingness to implement it, their individual competency, and the competency and culture within the team. Therefore, it is imperative that teachers have a positive attitude towards formative assessment, understand the theoretical rationale and values underpinning it, have room for agency in implementing it, receive continuous support from their colleagues and school leadership as well as opportunities for professional development, and lastly, that they collaborate with their colleagues within a professional learning community (Heitink et al., 2016; Heritage, 2010; William & Leahy, 2015).

## **II. What is the case for formative assessment in primary education?**

Formative assessment has gained attention from a variety of stakeholders (teachers, policy makers, school leaders, educational researchers, teacher educators etc.) and is now being implemented in many countries across the educational spectrum, mainly due to the widespread belief that it holds a great potential for enhancing learning (Birenbaum, et al., 2013). The idea that formative assessment greatly increases student achievement was mostly proliferated by Black and Wiliam's (1998) meta-analysis, *Inside the Black Box*, in which the authors review approximately 20 studies on the impact of various formative assessment practices on student achievement. They concluded that formative assessment has an effect size ranging between 0.40 and 0.70, one of the highest ever reported for any educational intervention (Black & William, 1998, 2009). Ever since, there has been little doubt that formative assessment positively influences learning, and the narrative shifted towards *how* to implement, rather than *whether* to implement it.

Despite this apparent conviction regarding the reported effect sizes of formative assessment, some scholars (e.g. Bennett, 2011; Dunn & Mulvenon, 2009) have raised questions regarding the accuracy of the conclusions drawn by Black and William (1998) and the quality of the studies upon which they were founded. The main issues being raised are that the studies reviewed are too disparate (some concern the impact of feedback, others of self-assessment and so forth) to be summarised meaningfully, involve populations or characteristics which cannot be generalised to the large student population and education system (e.g. students with disabilities, disadvantaged students or concerning only particular subjects such as Science) and have considerable methodological flaws (Bennett, 2011; Dunn & Mulvenon, 2009). Since Black and Wiliam's work is almost universally cited to support the implementation of formative

assessment, these shortcomings place the foundation of formative assessment under critical scrutiny.

Additionally, Dunn and Mulvenon (2009) warn that the lack of a comprehensive, concrete and widely accepted definition of formative assessment renders the concept virtually impossible to accurately study and quantify since a myriad of practices fall under the umbrella term ‘formative assessment’ (Bennett, 2011). Hence Black and Wiliam’s (1998) review has been referred to by some as a good qualitative review of the literature on an ill-defined, amorphous intervention type (Hanover Research, 2014). Nonetheless, some experts argue that this vagueness is precisely what ensures efficacy as formative assessment is intended to be used in a variety of contexts and classrooms and therefore cannot be confined to rigid parameters (Hanover Research, 2014). Whilst this might be the case, there is no doubt that the elusive nature of the concept poses considerable methodological challenges (Hanover Research, 2014).

Notwithstanding, Clark (2011) emphasises that quantitative ambiguities are not sufficient to obscure the deep cognitive and metacognitive processes resulting from high quality feedback and interaction. Similarly, Wiliam (2011) maintains that the question is not what works in education, as most interventions work to varying degrees and under different circumstances. Notably, neither Bennett (2011) nor Dunn and Mulvenon (2009) dismiss the possibility that formative assessment can significantly and positively impact learning, yet they maintain that so far, studies failed to conclusively demonstrate this fact and ask for high quality, rigorous and focused research which can be compared and generalised (Hanover Research, 2014).

## **Research design**

### **Methodology**

The present project represents a *case study* which explored, in-depth, the alignment between the three variables- school policy, teachers' attitudes and teachers' practices- through the collection and analysis of *qualitative* and *quantitative* data (Hope, 2016).

### **Paradigm**

An interpretative paradigm was employed considering (i) the volatile, complex nature of the concept of attitudes and the observable practice, and (ii) the multitude of existing definitions and conceptualisations of formative assessment and the ambiguity still surrounding this term (Fives & Gill, 2015; Hope, 2016).

### **Setting**

The setting for this project was an international, primary school in Malaysia. The school included a primary and secondary department in the same campus. The international primary department, following the British Curriculum and the International Primary Curriculum, had been making efforts over the last three years to implement formative assessment.

### **Participants**

The participants in the research were class teachers employed in the upper and lower primary department at the time the study took place. Twelve teachers took part in the online questionnaire; Table 1 provides information on their background gathered through the Demographics section of the questionnaire. Of the 18 teachers working in the international

primary department at the time of the study, 17 were Malaysian, therefore the team was highly homogenous in terms of nationality and cultural background. To ensure anonymity in the likelihood that the study would be shared with the school, the respondents were categorised based on age and years of experience rather than collecting specific data on their background.

Table 1  
*Demographics*

Age	Respondents	Years of experience	Respondents
20-30	2	1-5	3
31-40	5	6-10	5
32-50	5	11-20	3
		21-30	1

Two teachers (one from KS1 and 1 from KS2) were invited to individually and anonymously take part in 1 interview and 2 lesson observations. Throughout the study they will be referred to under a pseudonym: Alice and, respectively, Emma.

### **Ethical issues**

Throughout the process of data collection, analysis and dissemination, the Ethical Guidelines in Educational Research established by the British Educational Research Association (BERA, 2011) were followed. In line with these guidelines, anonymity has been ensured to all participants and at all stages of the project. Menter et al. (2011) stress that observations in particular can pose ethical challenges as they might constitute a breach of participants' privacy, therefore it is vital that anonymity is ensured. Therefore, following the recommendation of Menter et al. (2011), participation only took place with the participants' written consent, after they have been informed on the purpose, conditions and their rights regarding participation. All

participants have been made aware that they can withdraw from the study at any point with no consequences. A hard copy of the written consent forms has been kept by the researcher and a soft copy has been shared with the participants.

Since the researcher was an intern in the school at the time of the study, the participants have had the chance to get to know her and did not seem to perceive her as a threatening authority. This, alongside the ensured anonymity, has helped to balance the power relations between the researcher and the participants.

### Data gathering tools

*Figure 3* offers an overview of the data gathering tools used for each variable. The selection of tools was closely considered to ensure the study is achievable with the limited resources available.

Policy	• Document analysis
Attitudes	• Questionnaires
Practice	• Lesson observations • Interviews

*Figure 3. Data gathering tools*

#### Document analysis

The school's written policy on assessment was analysed; permission and access to the document was verbally requested from the Head of Primary who has provided a hard copy. The policy was titled Primary Assessment and Reporting Policy and was comprised of 13 sections

amounting to 7.5 A4 sheets. Of the 13 sections, 5 sections included data relevant to the study which is outlined in the Results chapter.

### **Questionnaire**

Questionnaires have a well-established tradition in the research on attitudes (Maitland, 2008) and therefore were used to collect data on teachers' attitudes towards the 7 principles of implementing formative assessment. The questionnaire was created and administered online using Google Forms.

Statements were generated by the researcher for each principle (between 3 and 6) on which the respondents expressed their attitude using the following scale: 1-strongly disagree; 2-disagree; 3-neither agree nor disagree; 4-agree; 5-strongly agree. To render the study a clear and practically manageable focus and scope, a selection was made of what aspects of each principle were included in the questionnaire, based on the literature (Andrade & Cizek, 2010; Heitink et al., 2016; William, 2011) (see Table 2).

Table 2  
*Summary of Focus Areas per Principle*

<b>Principle</b>	<b>Area of focus</b>
<b>LISC</b>	Sharing and designing learning intentions and success criteria with students
<b>Frequency</b>	How frequently formative assessment should be implemented in lessons
<b>Feedback</b>	Importance of feedback in formative assessment
<b>Eliciting data</b>	Incorporating self and peer-assessment
<b>Interpreting data</b>	Involving the students in reviewing and using assessment data
<b>Classroom culture</b>	Preferring an authoritarian or democratic classroom management style; teacher-students power relations
<b>Collaboration</b>	Perception of current level of collaboration within the team and own attitude towards collaborating with colleagues
<b>Competency</b>	Teacher's perceived competency in implementing formative assessment



In designing the questionnaire, the recommendations provided by Cob (2016) served as guidelines to ensure high standards of quality. Lastly, the questionnaire underwent a pre-test in which 8 volunteers (student-teachers) completed the questionnaire and offered written feedback which was used to correct the identified shortcomings.

### **Observations**

Observations are highly recommended for studying behaviours and activities (Menter et al., 2011), therefore they were employed for gathering data on teachers' classroom practices. Four 1-hour, overt, non-participant observations (Menter et al., 2011) were conducted, each on a different day (2 of Alice and 2 of Emma). The data was recorded in the form of hand-written notes which were later transcribed using an observation form designed specifically to suit the purpose of the observation (see Appendix III).

### **Interviews**

Since what is observable is not always readily understandable, observations have been complemented with semi-structured interviews allowing the respondents to give their own view on their classroom practices, reducing bias on the part of the researcher (Menter et al., 2011). The interviews were conducted prior to the observations in a quiet setting (the school's library). The interviews were audio recorded and stored on the researcher's personal phone and later transcribed verbatim in a Word document, each within a week from the time they took place (see Appendix II for the list of questions).

## Data analysis

The resulting data was analysed deductively using thematic analysis (Blaxter, Hughes, & Tight, 2010), based on the *a priori* (Cohen, Manion, & Morrison, 2018) list of principles emerging from the Literature Review (Heritage, 2010; Leahy et al., 2005). Additionally, since the context was unique and of great importance to this case study, additional themes arising from the data were included under the category *Other* and considered of equal importance for the discussion, analysis and conclusion. To aid in categorising the data according to the 7 principles, the researcher has devised a list of key terms for each principle (see Appendix I).

## Questionnaire

The quantitative data collected through the online survey was encoded in attitude maps (Ogan-Bekiroglu, 2018) (see Appendix IV). The first 6 principles were encoded each in one attitude map, whereas the last principle (teacher attitude, competency, collaboration and development- TACCD) has been encoded in 2 maps (competency and collaboration) due to its complexity. For each principle, positive as well as negative statements have been included to validate the data. The scores of the negative statements (in italics, see Appendix IV) have been reversed during encoding (e.g. 5-strongly agree with a negative statement has been encoded as 1-strongly disagree in the positive attitude and vice-versa in the negative attitude). Each map offers an overview of the amount of participants (as *n* and percentages) holding a positive, negative or neutral attitude towards each principle. Additionally, the scale points selected by respondents have been added and their mean value shows how strongly a positive or negative attitude is held overall towards one particular principle (this is not applicable for the neutral attitude).

One unexpected limitation arose which is worth noting: due to an error which occurred at the moment of sending the URL of the questionnaire, the section on Interpreting data had been

saved as a tick question instead of a scale question. As a result, respondents could only tick one or more statements and no data could be obtained as to how strongly they feel about the statements. To overcome this inconvenience, the statements have been categorised into denoting a positive (2 statements), negative (1 statement) or neutral attitude (1 statement) towards involving students in the process of interpreting data (see Appendix IV). The mean amount of respondents for each attitude (in percentages) has been included in Table 5.

### **Interviews**

The data resulting from the interviews has been colour-coded and tabulated according to the 7 principles emerging from the Literature review. Additionally, comments have been inserted with the researcher's thoughts, interpretations and connections made to other data and findings from the Literature review.

### **Observations**

The typewritten notes of the observations were compared to the relevant categories of data on self-reported practice generated through the interviews (principles 1, 2, 3, 4 and 6). Discrepancies between the self-reported and observable practice of each relevant principle were marked with ✕ whilst correspondence was marked with ✓ in the Observations column of Table 7 and 8 (see Results).

### Quality

Throughout the study, efforts were made to maintain high quality and academic rigour to ensure that the conclusions drawn are valid. Hence *validity* -i.e. the extent to which (i) an instrument measures what it claims to measure and (ii) the meaning and interpretation of the results of the data collection and instrumentation are sound (Ary, Jacobs, & Razayieh, 2002)- was enhanced in several ways.

Firstly, the relatively long duration of the study (13 weeks) and the habitual presence of the researcher in the setting, enabled long-term, intense involvement and the collection of in-depth, rich data, which is a means to ensure validity (Maxwell, 2005; Menter et al., 2011)

Secondly, in line with the interpretative paradigm, open questions were included in the questionnaire, inviting respondents to define the key terms and provide examples (Hope, 2016); this enabled the researcher to make more informed and less biased interpretations of the data, thereby ensuring interpretive validity, a characteristic of qualitative research validity (Maxwell, 1992).

Additionally, two different instruments were used to collect data on teachers' practices, interviews (for self-reported) and observations (for observable practice) which according to Onwuegbuzie and Leech (2006) is a means of triangulation and enhances the validity of the study. This led to a clearer and less biased understanding of teachers' practices and aided the researcher in the interpretation and data analysis processes (Menter et al., 2011).

Furthermore, instrument validity of the questionnaire was maximized by following the guidelines set out by Magee, Rickards, Byars, & Artino (2013). Moreover, the questionnaire underwent a pre-test for general, as well as criterion validity (Cohen et al., 2018).

Lastly, several attempts were made to render the study *reliable*, meaning that the results are consistent and replicable over time, instruments and groups of respondents (Cohen et al., 2018). Although reliability is difficult to establish in qualitative research and particularly in case studies (Cohen et al., 2018), ensuring anonymity and piloting the questionnaire help attenuate issues which hinder reliability, such as respondents not understanding a question, being unwilling to disclose information or giving socially desirable answers (Fowler, 2009). Therefore, varied and continuous efforts were made to increase the validity and reliability of the study.

### **Results**

This section presents the main findings of the study, answering the following sub-questions:

- (vi) *How is formative assessment formulated in the school policy?*
- (vii) *What are teachers' attitudes towards implementing formative assessment?*
- (viii) *How do teachers implement formative assessment in their teaching?*

### **Policy**

Table 3 summarises the school policy regarding formative assessment in relation to the 7 principles discussed earlier.

Table 3

*Formative Assessment in the School Policy*

Principle	Present/absent	Primary Assessment and Reporting Policy
<b>LISC</b>	✓	TIPSI* believes that good assessment for learning: (...) involves sharing learning objectives (WALTs*) with students and aims to help students comprehend these objectives
<b>Frequency</b>	✓	Formative assessment should be on-going TIPSI believes that good assessment for learning: (...) uses self and peer assessment consistently.
<b>Feedback</b>	✓	TIPSI believes that good assessment for learning: (...) provides feedback which leads students to recognize the next steps in their learning and how best to proceed (see Feedback and Marking policy for more details).
<b>Eliciting data</b>	✓	TIPSI believes that good assessment for learning practice involves: questioning, giving meaningful written and oral feedback to students (...). TIPSI believes that good assessment for learning: (...) uses self and peer assessment consistently.
<b>Interpreting data</b>	✓	TIPSI believes that good assessment for learning: (...) involves both students and teachers reviewing and reflecting on assessment information. Students are exposed to data as a means to select appropriate targets and identify areas of weakness It is important that when teacher assessments are carried out, there is evidence recorded to justify judgements made. Writing moderation* is carried out termly by all year groups. The PLT* carries out Book Looks* (Work Scrutiny) termly, feeding back findings to teachers.
<b>Classroom culture</b>	✓	At TIPSI we aim to provide: (...) encourage dialogue between children and adults regarding the progress, success and areas to target.
<b>TACCD</b>	✗	
<b>Other</b>		Formative Assessment (Assessment for Learning) – This is the on-going assessment carried out by teachers both formally and informally during a unit of learning. The results of formative assessment have a direct impact on the learning and teaching strategies employed following the assessment.
<b>Definition</b>		

*Note.*

TIPSI- acronym denoting the name of the school

WALTs- “What Are We Learning To” equivalent to Learning Objective

WILF- “What Are We Looking For” equivalent to success criteria

Writing moderation- a termly meeting of teachers from KS1 and KS2 in which teachers mark and assess writing samples of students from classes other than their own with the purpose of discussing assessment practices to reach standardization and consistency across the school

PLT- Primary Leadership Team (Head of Primary and Deputy)

Book Looks- the termly practice of the PLT analysing workbooks of several random students from each class and for each major subject, to evaluate teachers’ practices of marking and writing feedback.

## **Attitudes**

The findings presented in this sub-section reflect teachers’ attitudes towards the 7 principles. Of special importance to the study is *how* teachers defined formative assessment, which is outlined in Table 4 and will later be compared to the definition set out in the policy. Table 5 summarises the results of the questionnaires and gives an overview of teachers’ positive, negative and neutral attitudes towards each principle.

Table 4

### *Teachers’ Definitions of Formative Assessment*

Category	<i>n</i>
Ongoing assessment	5
Assessment which impacts subsequent learning and teaching	4
Assessment which evaluates learning and teaching	2
Assessment which takes place during the learning process	4
Other	3

*Note.* The total *n* in the table amounts to more than 12 (the total of respondents) as some



respondents mentioned more than one category in their open answers.

Table 5

*Teachers' Attitudes towards Formative Assessment*

Principle	Positive	Negative	Neutral
<b>LISC</b>	74 % 40 p	21 % 1 p	13 %
<b>Frequency</b>	39 % 19 p	39 % 9 p	22 %
<b>Feedback</b>	86 % 46 p	11 % 2 p	3 %
<b>Eliciting data</b>	58 % 46 p	19 % 4 p	21 %
<b>Interpreting data*</b>	66 %	0 %	25 %
<b>Classroom culture</b>	61 % 28 p	17 % 3 p	19 %
<b>Collaboration</b>	58 % 29 p	28 % 6 p	14 %
<b>Competency</b>	42 % 21 p	17 % 4 p	42 %

*Note.* For practical reasons, the percentages presented have been rounded to the nearest whole, therefore the scores in each item add up to approximately, yet not exactly, 100%. The percentages show how many respondents hold each attitude, whilst the points (p) show how strongly the positive and negative attitudes are held (attitude strength is irrelevant for the neutral attitude). The attitude strength has been obtained by calculating the mean of the scale points, 4 (agree) and 5 (strongly agree) for the positive attitude and 2 (disagree) and 1 (strongly disagree) for the negative attitude. Therefore, in the case of negative attitudes, the value denoting the strongest attitude is where the *p* is proportionally very low in comparison to the % (see LISC).

## Practices

This sub-section presents the findings on teachers' self-reported and observable practice. Table 6 summarises teachers' answers to the open question: *What strategies do you use when implementing formative assessment?* from the questionnaire. The answers highlighted in grey have been categorised as not clearly relating to formative assessment practices. Tables 7 and 8 present the main findings on the self-reported and observable practice of Alice and Emma.

Table 6  
*Teachers' Self-reported Formative Assessment Strategies (Questionnaire)*

Strategy	<i>n</i>
Questioning	4
Feedback	4
Self-assessment	4
Peer-assessment	2
Thumbs	3
Exit card/ticket	2
2 stars and 1 wish	2
Quiz	1
KWL chart	1
Observation	1
Tickle pink and green for growth	2
Worksheet	1
Student activities	1
Group activities and interactive	1
Checking on topics that are important for their learning and understanding	1

Table 7

Teachers' Self-reported and Observable Formative Assessment Practices-Alice

Principle	Interview Alice	Observations
<b>LISC</b>	<b>x</b>	✓ Shared in written form and verbally discussed with the class
<b>Frequency</b>	Not for every lesson we can do formative assessment, and not every, all the formative assessment can be done in every lesson	✓ In each lesson the teacher assessed the students at the beginning, during and at the end of the lesson
<b>Feedback</b>	Formative assessment is, um is a way that we assess children's learning and where they are through a few ways: (...) <b>feedback</b> (...) and also, <b>when you talk to them, you know, is a two-way feedback- I</b> give them <b>feedback</b> and they give me <b>feedback</b> (...)	✓ Verbal (in plenary and individually) Written (comments on students' work)
<b>Eliciting data</b>	Thumbs up, thumbs down, traffic lights  (...) we assess children's learning and where they are <b>through</b> a few ways: questioning, observations, feedback, quizzes, their work, their worksheets, (...) and I ask them to do self-assessment  I think I need more <b>methods</b> (...) Yeah, <b>strategies</b> , because I feel I've been using the same strategy to the extent I'm kind of feeling bored of using the same strategy, so I would love some challenges, something new, and trying something new is always fun.	✓ Thumbs up and down, smiley face self-assessment, peer marking (tick or dot), questioning, mini-whiteboards
<b>Interpreting data</b>	(...) some can guess and correct the right answer (...) some of them copy their friends, (...) so you can't really buy and say that "ok, he's fine, he can move on" so that's why you need many other aspects (...) More observations, yes  If you go down to early years and you ask them if you are good put thumbs up if you don't know put thumbs down, nobody likes to put thumbs down (...) yes, it's very unreliable data, you cannot take	Not applicable

Principle	Interview Alice	Observations
	that data (...) definitely it has flaws, so that's why, at the end of the day it relies on the hand of the teacher to know your children and to know their strengths and weaknesses	
<b>Classroom culture</b>	There is a very good relationship between the teacher and the students, because when they communicate they feel it's <b>safe</b> and they <b>openly</b> share their thoughts when they need help so like, if we do only summative, like a, like an <i>exam</i> kind of test and if you don't know you just don't know, so they have that fear of saying "I don't know", but during formative assessment when they say "I don't know", it's okay to make <b>mistakes</b> and so they feel it's comfortable and they're not afraid of trying new things	✓ Students appeared comfortable asking for help; the teacher emphasised good struggle in learning ✕ Teacher focused on students getting the correct answer rather than explaining own thinking about the task
<b>TACCD</b>	Yeah, actually here you do receive good support from the primary head and the deputy head and the year leader, so we all help each other, and then we plan together Yes, so you know, you can always just ask the assistants "what do you think?", but not always because not all the assistants are open to share, (...) some, they will just tell you this is not right, but they are not constructive  Hm, I think maybe the school could send us out, somewhere, to go and have some training, intensive formative assessment training, rather than, usually, when you're in school it's like the head finds a set of, you know, some information from anywhere and then they share it with us and then we try that out; (...) at the end of the day you will have tried <i>all</i> of them	Not applicable
<b>Other Definition</b>	Formative assessment is, um is a way that we assess children's learning and where they are through a few ways: questioning, observations, feedback, quizzes, their work, their worksheets, and also, when you talk to them, you know, is a two-way feedback- I give them feedback and they give me feedback, and I ask them to do self-assessment (...)	Not applicable

Table 8

*Teachers' Self-reported and Observable Formative Assessment Practices-Emma*

Principle	Interview Emma	Observations
<b>LISC</b>	Yeah, you need to keep reminding them (...) to look at the WILF, or what is our outcome, have you reached that outcome and if not, how did you want to reach that outcome, you know	✓ Shared in written form and verbally discussed with the class
<b>Frequency</b>	<i>Do you think there are any possible negative aspects as well of implementing formative assessment (...)?</i> Maybe the main things would be time constraints that the teachers are facing, or over spilling into another time-taken or time eaten up to do formative assessment  I will start with the not satisfied yet, mainly because of time, because you want to do so much in a lesson and that's not possible	✓ In each lesson the teacher assessed the students at the beginning, during and at the end of the lesson
<b>Feedback</b>	Because it's kind of a two-way communication between the teacher and the students, especially for example one way of formative assessment is to get <b>feedback</b> from their learning, you know the teacher talks to them and they will give you the feedback to see where they are standing  (...) in marking their books as well I will usually give a lot of constructive <b>comments</b> , sometimes <b>written</b> or sometimes, a lot of times <b>verbal feedback</b> , you know, walking around the room	✓ Verbal (in plenary and individually) Written (comments on students' work)
<b>Eliciting data</b>	I do a lot of higher-order questioning, I try to do that, such as questioning about the How's and the Why's, you know, (...) and also I think that questioning helps to discern children's current level, where they are in now, so as from there we can really extend their understanding Popsicle sticks (...) Gallery walk (...) Thought alley (...) Plenaries (...) Tests Exit tickets (...) Quizzes (...) Thumbs (...) KWL (...) Posters end of topic Mini whiteboards (...) Think-pair-share (...) Two stars and a wish	✓ Thumbs up and down, peer marking (tick or dot), and assessment (2 stars and 1 wish), (higher-order) questioning, mini-whiteboards, plenary, quiz, think-pair-share, popsicle sticks
<b>Interpreting data</b>	✕	Not applicable

Principle	Interview Emma	Observations
<b>Classroom culture</b>	(...) the children have ownership of their own learning rather than being told to do so	✓ Students appeared comfortable talking about their thinking and learning; the teacher emphasised that there are no wrong answers and corrected students' mistakes without any visible negative connotations
<b>TACCD</b>	<p>On my part as a teacher I think I am better informed of the skills that need more improvement on so that eventually I can plan better to cater to their needs, yes, I feel that, yes I am more efficient, and more uhm I think I am higher in my <b>competency</b> as a teacher, I feel more rewarding as well</p> <p>Hm, the school has actually provided a lot of opportunities of <b>training</b> us, during uhm, what do you call that, term meeting and teacher training sessions, talking through about it and we're always encouraged to discuss to our <b>colleagues</b> (...) and share good practices</p> <p><i>How do you think still the support of the school could be better, if in any way?</i> Mmm, perhaps send us more frequently outside of the school, not only within</p>	Not applicable
<b>Other Summative vs Formative assessment</b>	<p>I think as long as it [formative assessment] doesn't sort of take over other forms of assessment, such as summative, if it goes hand in hand you know with the summative part of assessments, I think it will give a very good picture of the overall progression the students are making</p> <p>And the summative assessments should affect how the formative should go</p>	Not applicable
<b>Definition</b>	I think formative assessment stands for Assessment for learning, which speaks for itself, which is uhm, the teachers get to know, understand how students' understanding is about a certain topic, and then also looking at the students' work to have an insight into their knowledge and then from there the teacher is going to see the strengths and weaknesses and further on you could even analyse their learning ability (...)and whether they need assistance or not	Not applicable

## Discussion

This study examined the alignment between (i) the school's policy, (ii) teachers' attitudes and (iii) teachers' practices regarding formative assessment (as encompassing the 7 principles from the Literature review). This section discusses the main findings as outlined in the Results in relation to other relevant research, as well as the (mis)alignment of the three variables central to the study.

### *Policy*

The school Policy explicitly addressed the use of formative assessment and provided rather general (and in some cases more specific) guidelines on 6 of the 7 principles, excluding TACCD, despite an extensive body of research and literature stressing its importance in the successful implementation of formative assessment (Heitink, et al. 2016; Heritage, 2010; Priestley & Sime, 2005).

### *Attitudes*

Overall, teachers' attitudes towards the 7 principles were mostly positive, with Feedback and LISC scoring highest. This seems to be in line with the school's policy which repeatedly and explicitly emphasised the two principles. Additionally, literature on formative assessment also stresses the salience of shared (and especially co-created) learning intentions and success criteria as a prerequisite to formative assessment (see William, 2011) as well as the fundamental role of feedback in formative assessment (Andrade & Cizek, 2010; Hattie & Yates, 2014; Heitink et al. 2017; Marzano, 2006).

### *Definition of formative assessment*

Two thirds of the teachers did not give a definition of formative assessment which includes its purpose of informing and altering subsequent teaching and learning. This reveals a mismatch

between how the teachers defined and understood formative assessment and how the policy and, importantly, literature define it. Nearly half of the respondents (*n*5) however, did state that formative assessment is *on-going* which is in line with the policy, whilst 4 teachers correctly stated that it takes place during the learning process. Notably, 2 teachers even included terms such as *evaluate* when defining formative assessment, terms which are intrinsically linked to summative assessments (Dunn & Mulvenon, 2009). These results suggest an incomplete and/or erroneous understanding of the concept of formative assessment, which can have a significant impact on its implementation. In a meta-analysis of studies on the implementation of formative assessment in Asia (including Malaysia and neighbouring countries), Do Quyen & Khairani (2016) found that teacher knowledge was the biggest factor in implementing it; importantly, in most studies reviewed, teachers lacked a sound understanding of the concept and practice of formative assessment. Similar studies, focusing specifically on Malaysian (primary school) teachers, have also found a general lack of understanding of formative assessment (see Hasim, Di, & Barnard, 2019 and Talib et al., 2014); since 17 out of 18 teachers in the international primary department were Malaysian, this finding seems to confirm those of other studies. The assumption made here that teachers' understanding of formative assessment was somewhat erroneous/incomplete is partly supported by findings from the interviews (e.g. through Alice's definition and some of the strategies which Emma used yet which are not clearly related to formative assessment) and by the 4 responses of teachers' strategies highlighted in grey (see Table 6).

### *Frequency*

Interestingly, although five teachers defined formative assessment as on-going, the overall attitude toward Frequency was the least positive and most negative from all the 7 principles.



Lack of time might be a possible cause for this, as suggested by the data from the interview with Alice (see Table 8). Do Quyen & Khairani (2016) maintain that the time-consuming nature and the additional workload are two major disadvantages of formative assessment practice and found that therefore, “most Asian teachers in the studies reviewed were reluctant to integrate formative assessment into their daily teaching” (p. 168). This could explain why teachers held a negative attitude towards frequent implementation of formative assessment, despite being aware that it is an on-going process which can support students’ learning.

#### *Formative versus summative assessment*

A noteworthy finding is that teachers did not mention the use of summative assessments for formative purposes, neither in the survey nor in the interviews. Furthermore, one teacher juxtaposes the two forms of assessment in the interview and one respondent in the questionnaire (there is a possibility that the two respondents were, in fact, the same person). This is a misalignment with the policy, which mentioned the formative use of summative tests, as well as literature, which emphasises the need for congruency rather than juxtaposition between the two forms of assessment and maintains that there are no inherent aspects of summative assessment which can prevent teachers from implementing formative assessment (OECD/CERI, 2008). A distinct finding resulted from the interview with Emma, who emphasised that formative assessment should not take over summative assessment but should go hand-in-hand and even be affected by summative assessments. Nevertheless, this view seems to be isolated as all other remarks on this issue simply placed it at the opposite end from formative assessment.

#### *Interpreting data*

Another interesting finding is that Alice raised the issue of reliability of assessment data, particularly in the early years. According to her, students often copy each other’s answers or at

times might simply guess the correct answer, which leaves the teacher with an inaccurate picture of students' levels and needs. In her opinion, this can be overcome if the teacher knows her students very well and complements her formative assessment strategies with continuous observation. This shows that she was well-aware of the need to search for corroborating data and that she was actively engaged in this process, which according to Heritage (2010) is a cornerstone in interpreting assessment data.

### *Classroom culture*

The unreliability of assessment data due to students' copying their peers' answers raised the question of how safe it is for them to make mistakes and reach the 'wrong answer'. From the two observations of Alice's lesson it became visible that she focused on the product rather than the process of their learning, which is a trademark of summative and not formative assessment (Heitink et al., 2016). This might be partly explained by the focus and pressure on academic achievement which underpins the exam-driven approach to education in some Asian countries (Do Quyen & Khairani, 2016). In such a classroom culture, formative assessment becomes very problematic (Do Quyen & Khairani, 2016; Leahy & William, 2015) as students are likely to be reluctant to share their thinking, which undermines teachers' ability to gauge their understanding and provide feedback tailored to their needs. Alice, however, states that formative assessment has helped her create a safer and more open classroom atmosphere in which students dare to admit mistakes and ask for help. This might imply that although literature stresses that a conducive classroom culture is a pre-requisite for implementing formative assessment, in this particular case it became a result of its implementation.

### *Questioning*

Data on teachers' self-reported and observable practice clearly showed that questioning is one of the main strategies used to elicit formative assessment data (see Table 6 and *Eliciting data* in Tables 7 and 8). Not only is this in concert with the school policy, but also with the literature, as William and Leahy (2015) state that questioning is a staple strategy for eliciting evidence about student achievement world-wide. However, teacher questioning becomes more and more challenging as the size of the group increases; this could be particularly problematic in Emma's case, as her class had 27 students and no Teaching Assistant at that time. Interestingly, Emma did not mention the group size at all throughout the interview.

### *TACCD*

Regarding the 7<sup>th</sup> principle- TACCD- findings from the interviews revealed that teachers have been provided with plenty of opportunities of in-school training complemented by colleagues' observations and feedback, yet insufficient training outside of school. Both Emma and Alice stress a need to learn new strategies and have fresh input about formative assessment. The need for more professional development was also suggested by the fact that teacher competency scored highest in the neutral attitude and second lowest in the positive attitude (see Table 6), which implies a low self-efficacy -i.e. belief in one's ability (Suprayogi, Valcke, & Godwin, 2017). In a study on primary teachers' attitudes, intentions and practices regarding formative assessment, Cheng & Yan (2015) found that self-efficacy is one of the most influential factors in teachers' developing the intention to conduct formative assessment. Therefore, the overall low self-efficacy could be impeding the successful implementation of the reform.

### Conclusion

This case study examined the alignment between (i) the school policy, (ii) teachers attitudes and (iii) their self-reported and observable practice regarding formative assessment (see *Figure 1*). Overall, the findings suggest that the three variables are *partly* aligned. Table 9 provides a more detailed overview of the areas in which they align or not. The question mark refers to an insufficient amount of data in that category from which to draw a valid conclusion.

Table 9  
*Alignment between the School Policy, Teachers' Attitudes and Teachers' Practices*

Principle	Policy	Attitudes	Practices
Learning intentions and success criteria	✓	✓	✓
Frequency	✓	✗	✓
Feedback	✓	✓	✓
Eliciting data	✓	✓	✓
Interpreting data	✓	✓	?
Classroom culture	✗	✓	✓
TACCD	✗	✗	✓
<i>Other</i>			
<i>Definition</i>	✓	✗	?
<i>Formative use of summative tests</i>	✓	✗	?

### Recommendations

Based on the findings of this study, several recommendations were made to maximize the alignment of the three variables:

**1. Provide professional development opportunities outside the school-** interviewees

stressed a need for training regarding formative assessment which goes beyond the school team and grounds; this could include:

- Free, online training: courses available focusing on formative assessment and STEM subjects, use of technology and ELL students:

(<https://www.futurelearn.com/courses/introducing-assessment-for-learning>,

<https://www.futurelearn.com/courses/planning-for-learning>,

<https://www.coursera.org/lecture/ell-assessment/key-features-of-formative-assessment-uVNIW>,

<https://www.coursera.org/learn/assessmentforlearning>

<https://www.coursera.org/lecture/getinmooc/welcome-to-week-3-aRXEp>)

- Inviting teachers/experts from other schools to come in and observe lessons, give feedback and share their experience and strategies
- Sending teachers to other international schools to observe formative assessment practices

**2. Provide guidelines (in the policy), training, advice and support for building a**

**conducive classroom culture-** teachers need support and training in shifting their own and their students' mindset from one focused on the product of learning (correct answer) to the process of learning (reflecting on and improving learning).

- Recommended reading: William & Leahy (2015)

3. **Establish a Professional Learning Community (PLC)**- this could help enhance teachers' self-efficacy as well as collaboration and development, which the results revealed to be rather weak.

- Recommended reading: Heritage (2010) and William & Leahy (2015)

4. **Give teachers more planning time and teaching time**- teachers acknowledged that formative assessment should be on-going, yet had the highest negative attitude towards Frequency due to time restrictions, therefore the school could:

- Reduce the number of school events which take up much teaching and planning time
- Increase the number of TAs in KS2, especially in classes larger than 25 students, to maximize the potential of formative assessment practices (particularly questioning, discussions and feedback).

### **Limitations**

Despite its promising heuristic value, this study also has several limitations which could negatively affect the overall quality of the results and the conclusions drawn from them.

The main limitation is the fact that it is a case study which examines the issue in-depth in a specific context and can therefore not be generalised to different contexts. Nevertheless, it can still bring about important gains in the study of formative assessment implementation as it might inspire other researchers to examine the alignment of the three main variables; as more studies on this issue will emerge, we will likely better understand the nuts and bolts of implementing formative assessment as a whole-school reform, leading to more concrete recommendations to support all stakeholders in this process.

Another considerable limitation is the fact that data gathered on formative assessment practices is mostly concerned with 2 participants while there are 18 teachers in the international primary department. It is quite reasonable to expect that observing other teachers might have revealed different practices which would have influenced the findings and conclusion of the study. Nevertheless, it was beyond the reach of a lone researcher to observe more than 2 teachers, as the study involved collecting and analysing other data as well in a tight, pre-set time frame. Therefore, an open question has been included in the online questionnaire inviting respondents to share which strategies they use within formative assessment, thereby providing a more general idea of some practices within the school. Furthermore, each teacher has been observed twice, on different days to get a more valid and comprehensive picture of their practice.

Additionally, the participants were informed of the observations several days beforehand which is likely to have led them placing more emphasis on formative assessment in the observed lessons than usually. This could have been prevented by simply observing a random class

spontaneously, yet the position of the researcher in the school (intern) could not allow such intrusion; moreover, this would not necessarily have led to more accurate data on teachers' general practice of formative assessment as due to various factors (e.g. time, focus of the lesson etc.) they might not have been able to incorporate formative assessment in those lessons, although they generally do so.

Lastly, neither the interviewees nor the researcher are native speakers of English, which is the language in which the interviews were conducted. Furthermore, the participants and the researcher come from very different cultures; these two aspects can make communication difficult and analysis of data unreliable as the researcher might be having a different perspective from that of the interviewees on basic concepts, such as education, children, learning etc. To minimise this limitation, data has been collected for both observable practice as well as teachers' self-reported practice to give a coherent overview of the participants' practice of formative assessment. In addition, during the interview the researcher has consistently paraphrased, summarised and asked checking questions to ensure that she understood the teachers' answers as they were intended.

To enhance the quality of further research in this direction, it is warranted to study the self-reported and observable practice of a larger population, possibly of all teachers in a school. To make this manageable, the observations could focus on obtaining quantitative data (e.g. using tally charts) by looking at the frequency of behaviours/activities associated with formative assessment. Similarly, a questionnaire focusing on self-reported practice could be administered to the whole sample, yielding more accurate, detailed and representative results than in the present study. Additionally, the reliability of the interpretations could be enhanced by including respondent validation, which in this study was not possible due to respondents' lack of time.



### References

- Andrade, H. L., & Cizek, G. J. (2010). *Handbook of formative assessment*. New York: Routledge.
- Ary, D., Jacobs, L. C., & Razayieh, A. (2002). *Introduction to research in education (sixth edition)*. Belmont: Wadsworth/Thomson Learning.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in education: Principles, Policies & Practice*, 18(1), 5–25.
- Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Looney, A., . . . Wyatt-Smith, C. (2013). International trends in the implementation of assessment for learning: Implications for policy and practice. *Policy Futures in Education*, 13(1), 117-140.
- Black, P. J., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.
- Black, P. J., & William, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 1(1), 1-41.
- Blaxter, L., Hughes, C., & Tight, M. (2010). *How to Research. Fourth Edition*. New York: McGraw Hill Education.
- Bryan, L. A. (2012). Research on science teachers' beliefs. In B. J. Fraser, K. Tobin, & C. J. McRobbie, *Second international handbook of science education. Vol. 1* (pp. 477-495). Dordrecht: Springer.
- Cheng, E. C., & Yan, Z. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teaching and Teacher Education*, 45(2015), 128-136.
- Clark, I. (2011). Formative assessment: Policy, perspectives and practice. *Florida Journal of Educational & Administrative Policy*, 4(2), 158-180.

- Cobb, W. (2016). Research methods: How will I collect the data? In R. Austin, *Researching primary education* (pp. 72-92). London: SAGE.
- Cohen , L., Manion, L., & Morrison, K. (2018). *Research methods in education. 8th edition.* Oxon: Routledge.
- Do Quyen, N., & Khairani, A. (2016). Reviewing the challengens of implementing formative assessment in Asia: The need for a professional development program. *Journal of Scoial Science Studies*, 4(1), 160-177.
- Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation*, 14(7), 1-11.
- Eaton, A. A., & Visser, P. S. (2008). Attitudes. In P. J. Lavrakas, *Encyclopedia of Survey Research Methods* (pp. 39-41). California: SAGE Publications Inc.
- Fives, H., & Gill, M. G. (2015). *International handbook of research on teachers' beliefs*. New York: Routledge.
- Fowler, F. J. (2009). *Survey research methods (fourth edition)*. Thousand Oaks: Sage.
- Hanover, R. (2014). *The impact of formative assessment and learning Intentions on student achievement*. Hanover Research.
- Hasim, Z., Di, S., & Barnard, R. (2019, May 1). Eliciting teachers' understanding and their reported practices on school-based formative assessment: Methodological challenges. *Indonesian Journal of Applied Linguistics*, 7(1), pp. 158-166.
- Hattie, J., & Yates, G. C. (2014). *Visible learning and the science of how we learn*. Abingdon, Oxon: Routledge.

- Heitink, M. C., Van der Kleij, F. M., Veldkamp, B. P., & Schildkamp, K. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational Research Review*, 17(2016), 50-62.
- Hendriks, M. A., Scheerens, J., & Slegers, P. (2014). Effects of evaluation and assessment on student achievement: A review and meta-analysis. In M. A. Hendriks, *The influence of school size, leadership, evaluation and time on student outcomes* (pp. 127-174). Enschede: University of Twente.
- Heritage, M. (2010). *Formative assessment. Making it happen in the classroom*. California: Corwin.
- Hope, G. (2016). Methodology: What approach should I take? In A. Rebecca, *Researching primary education* (pp. 55-71). London: SAGE.
- Kahl, S. (2005). Where in the world are formative tests? Right under your nose! *Cognition*, 25(4), 11.
- Leahy, S., Lyon, C., Thompson, M., & William, D. (2005). Classroom assessment: Minute-by-minute and day-by-day. *Educational Leadership*, 63(3), 18-24.
- Magee, C., Rickards, G., Byars, L. A., & Artino, A. R. (2013). Tracing the steps of survey design: A graduate medical education research example. *Journale of Graduate Medical Education*, 5(10), 1-5.
- Maitland, A. (2008). Attitude measurement. In L. P. J., *Encyclopedia of Survey Research Methods* (pp. 37-38). California: SAGE Publications Inc.
- Marzano, R. J. (2006). *Classroom assessment & grading that work*. Virginia: ASCD.
- Maxwell, J. A. (1992). Understanding and validity in qualitative research. *Harvard Educational Review*, 62(3), 279-300.

- Maxwell, J. A. (2005). *Qualitative research design: An interactive approach (second edition)*. Thousand Oaks: Sage.
- Menter, I., Elliot, D., Hulme, M., Lewin, J., & Lowden, K. (2011). *A guide to practitioner research in education*. London: SAGE.
- Miller, A. (2015, 9 2). *Using assessment to create student-centered learning*. Retrieved from Edutopia.org: <https://www.edutopia.org/blog/assessment-create-student-centered-learning-andrew-miller>
- Murchan , D., & Shiel, G. (2017). *Understanding and applying assessment in education*. London: Sage Publications Ltd.
- OECD/CERI. (2008). Assessment for learning. Formative assessment. *Learning in the 21st century: Research, Innovation and Policy* (pp. 1-24). Paris: Centre for Educational Research and Innovation.
- Ogan-Bekiroglu, F. (2018). Utilization of attitude maps in evaluating teachers' attitudes towards assessment. *Asia-Pacific Forum on Science Learning and Teaching*, 9(1), 1-13.
- Onwuegbuzie, A. J., & Leech, N. L. (2006). Validity and qualitative research: an oximoron? *Quality and Quantity*, 41(2), 239-249.
- Priestley , M., & Sime, D. (2005). Formative assessment for all: A whole school approach to pedagogic change. *The Curriculum Journal*, 16(4), 475-492.
- Robinson, K., & Aronica, L. (2015). *Creative schools. The grassroots revolution that's transforming schools*. New York: Viking Penguin.
- Rokeach, M. (1969). *Beliefs, attitudes, and values: A theory of organisation and change*. San Francisco: Josey-Bass.

- Suprayogi, N., Valcke, M., & Godwin, R. (2017). Teachers and their implementation of differentiated instruction in the classroom. *Teaching and Teacher Education*, 67, 291-301.
- Talib, R., Kamsah, M. Z., Naim, H. A., Latif, A. A., Abu Naim, H., & Abdul Latif, A. (2014). From principle to practice: Assessment for learning in Malaysian school-based assessment classroom. *International Journal of Social Sciences & Education*, 4(4), 850-857.
- Vähäsantanen, K. (2015). Professional agency in the stream of change: Understanding educational change and teachers' professional identities. *Teaching and Teacher Education*, 47, 1-12.
- William, D. (2011). *Embedded formative assessment*. Indiana: Solution Tree Press.
- William, D., & Leahy, S. (2015). *Embedding formative assessment. Practical techniques for K-12 classrooms*. Florida: Learning Sciences International.
- Yan, Z., & Cheng, E. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teaching and Teacher Education*, 45, 128-136.

## Appendices

### Appendix I – Key Terms

Principle	Key Terms
<i>Learning intentions and success criteria</i>	Learning objective, learning intention, WALT, WILF, success criteria, share(d), discuss(ed)
<i>Frequency</i>	Time(s), regular(ly), consistent(ly) on-going, daily, often
<i>Feedback</i>	Feedback, oral, verbal, written, comment(s), dialogue
<i>Eliciting data</i>	Method(s), strategy(ies), through, by means of, includes, involves students Self/ peer-assessment
<i>Interpreting data</i>	Students, corroborate, analyse(ing), evaluate, evidence, proof, infer(ence)
<i>Classroom culture</i>	Classroom culture, rules, safe, open, involve(s) students, mistakes, ask for help, share, dialogue
<i>TACCD</i>	Colleagues, collaborate, help, support, together, training, professional development, workshop, feedback, learn

**Appendix II – Interview Questions**

1. First of all, how would you define formative assessment in your own words?
2. How do you feel about having to implement formative assessment in your teaching as part of the whole-school approach?
3. What do you think are possible negative aspects of formative assessment as an approach?
4. Can you give examples of how you use formative assessment in your current class?
5. How satisfied are you with your current practice regarding formative assessment? Which factors bring you satisfaction? Which factors have the opposite effect?
6. How many opportunities have you had in the school (since you started working here) to receive feedback/advice or professional development regarding formative assessment?
7. How could the school help you improve your practice with formative assessment anything?
8. Do you think there is a great difference between the school's aims regarding formative assessment and the actual classroom reality? Can you further explain your opinion?

**Appendix III – Observation Form**

<b><i>Observed item</i></b>	<b><i>Frequency</i></b> 0-never; 1-at least once; 2-several times; 3-very often	<b><i>Notes</i></b>
<b>Learning intentions shared discussed</b>		
<b>Questions (closed and open)</b>		
<b>Self-Assessment</b>		
<b>Peer-assessment</b>		
<b>Feedback</b>		
<b>Classroom culture</b>		



### 1. Learning Intentions & Success Criteria

S1- It is vital that the teacher understands the learning objectives very well.

S2-Teachers must ensure that all children understand the learning objectives very well.

S3-There should be multiple learning objectives during the same lesson to ensure all students will reach them.

S4-The teacher should involve the students in designing the success criteria.

S5-Success criteria should always be shared with the students.

S6-Success criteria are unrelated to formative assessment.

#### Positive

S1- n11, 91,3% p. 49

S2- n10, 83% p. 44

S3- n8, 66,4% p. 36

S4- n9, 74,7% p. 39

S5- n9, 74,7% p. 41

S6- n7, 51,8% p. 32

**Mean**  
**p. 40,1**  
**73,5 %**

#### Neutral

S1- n0, 0% p. 0

S2- n1, 8,3% p. 3

S3- n3, 24,9% p. 9

S4- n1, 8,3% p. 3

S5- n1, 8,3% p. 3

S6- n3, 24,9% p. 9

**Mean**  
**p. 4,5**  
**12,5 %**

#### Negative

S1- n1, 8,3%, p. 1

S2- n1, 8,3%, p. 1

S3- n1, 8,3%, p. 1

S4- n2, 16,6% p. 2

S5- n2, 16,6% p. 2

S6- n2, 16,6% p. 3

**Mean**  
**p. 1.3**  
**20,8%**

## 2. Frequency

S1- Formative assessment should take place on a weekly basis.

S2-Formative assessment is not necessary in every lesson.

S3-Ideally, teachers should use formative assessment several times during a day.

### Positive

S1- n2, 16,6% p. 8

S2- n6, 49,8% p. 24

S3- n6, 49,8% p. 25

**Mean**

**p. 19**

**38,7 %**

### Neutral

S1- n3, 24,9% p. 9

S2- n3, 24,9% p. 9

S13 n2, 16,6% p. 6

**Mean**

**p. 8**

**22,1 %**

### Negative

S1- n7, 51,8% p. 12

S1- n4, 33,2% p. 8

S1- n4, 33,2% p. 7

**Mean**

**p. 9**

**39,4 %**

### 3. Feedback

S1- *The role of formative assessment is to collect data about students' progress and not to give feedback on it.*

S2- Feedback is central to formative assessment.

S3- Students should be given time to follow up on feedback they have received.

#### Positive

S1- n10, 83% p. 45

S2- n120, 83% p. 44

S3- n11, 91,3% p. 48

**Mean**  
**p. 45,6**  
**85,8 %**

#### Neutral

S1- n0, 0% p. 0

S2- n1, 8,3% p. 3

S3- n0, 0% p. 0

**Mean**  
**p. 1**  
**2,8 %**

#### Negative

S1- n2, 16,6% p. 3

S2 n1, 8,3% p. 1

S3- n1, 8,3% p.1

**Mean**  
**p. 1,6**  
**11,1 %**

#### 4. Eliciting data

S1- Self-assessment should be used regularly in formative assessment.

S2-Teacher-led formative assessment activities are the most beneficial for student learning.

S3-Peer-assessment is crucial in the practice of formative assessment.

#### Positive

S1- n8, 66,4% p. 38

S2- n7, 58,1% p. 29

S3- n6, 49,8% p. 27

**Mean**  
**p. 45,6**  
**58,1 %**

#### Neutral

S1- n2, 16,6% p. 6

S2- n3, 24,9% p. 9

S3- n3, 24,9% p. 9

**Mean**  
**p. 8**  
**21,1 %**

#### Negative

S1- n2, 16,6% p. 3

S2- n2, 16,6% p. 4

S3- n3, 24,9% p. 5

**Mean**  
**p. 4**  
**19,4 %**

## 5. Interpreting data

### Neutral

Frequent, interactive assessments of students' progress and understanding to identify learning needs and adjust teaching appropriately.

*n3*

### Negative

A tool that teachers use to measure student grasp of specific topics and skills they are teaching. It's a 'mid-stream' tool to identify specific student misconceptions and mistakes while the material is being taught. *n0*

### *Positive:*

*Encompassing all those activities undertaken by teachers and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged.* *n5*

*The process used by teachers and students to recognise and respond to student learning in order to enhance the learning, during the learning.* *n3*

## 6. Classroom culture

S1- Teachers are responsible with creating a positive classroom culture for giving and receiving feedback.

S2- I prefer a democratic classroom in which teacher and students share power.

S3- The teacher knows best what students need in order to learn.

### Positive

S1- n10, 83% p. 43

S2- n8, 66,4% p. 25

S3- n4, 33,2% p. 17

**Mean**  
**p. 28,3**  
**60,9 %**

### Neutral

S1- n1, 8,3% p. 3

S2- n2, 16,6% p. 6

S3- n4, 33,2% p. 12

**Mean**  
**p. 7**  
**19,4 %**

### Negative

S1- n1, 8,3% p. 1

S2- n1, 8,3% p. 1

S3- n4, 33,2% p. 7

**Mean**  
**p. 3**  
**16,6 %**

## 7. Collaboration

S1- I think within our team (the teachers' team) we are very supportive of each other in using formative assessment.

S2- I need more help from colleagues in implementing formative assessment.

S3- I am always open for feedback from my colleagues.

### Positive

S1- n9, 74,7% p. 37

S2- n1, 8,3% p. 4

S3- n11, 91,3% p. 46

**Mean**

**p. 29**

**58,1 %**

### Neutral

S1- n1, 8,3% p. 3

S2- n4, 33,2% p. 14

S3- n0, 0% p. 0

**Mean**

**p. 5,6**

**13,8 %**

### Negative

S1- n2, 16,6% p. 3

S2- n7, 58,1% p. 14

S3- n1, 8,3% p. 1

**Mean**

**p. 6**

**27,6 %**

## 8. Competency

S1- I feel competent in implementing formative assessment.

S2- *I often feel I lack the theoretical background to use formative assessment effectively.*

S3- I have enough experience with formative assessment to feel confident in implementing it.

### Positive

S1- n5, 41,5% p. 21

S2- n5, 41,5% p. 21

S3- n5, 41,5% p. 21

**Mean**

**p. 21**

**41,5 %**

### Neutral

S1- n5, 41,5% p. 15

S2- n5, 41,5% p. 15

S3- n5, 41,5% p. 15

**Mean**

**p. 15**

**41,5 %**

### Negative

S1- n2, 16,6% p. 4

S2- n2, 16,6% p. 4

S3- n2, 16,6% p. 3

**Mean**

**3,6**

**16,6 %**