Teacher AfL perceptions and feedback practices in mathematics education among secondary schools in Tanzania

Florence Kyaruzia, Jan-Willem Strijbosb, Stefan Uferc, Gavin T.L. Brown

a LMU Munich, Department of Psychology, Germany & Faculty of Education, Dar es Salaam University College of Education, University of Dar es Salaam, P. O. Box 2329, Dar es Salaam, Tanzania
b LMU Munich, Department of Psychology, Germany & University of Groningen, Department of Educational Sciences, Netherlands
c LMU Munich, Department of Mathematics, Germany
d University of Auckland, Faculty of Education and Social Work, New Zealand

A R T I C L E   I N F O

Keywords:
Formative assessment
Perceptions of FA and AfL
Teacher feedback practices
Mathematics education
Mixed methods

A B S T R A C T

Feedback that monitors and scaffolds student learning has been shown to support learning. This study investigates the effect of mathematics teachers’ perceptions of Formative Assessment (FA) and Assessment for Learning (AfL) and their conceptions of assessment on the quality of their feedback practices. The study was conducted in 48 secondary schools in Tanzania with 54 experienced mathematics teachers teaching Grade 11 (Form three in the Tanzanian system). Validated questionnaires were combined with interviews to investigate mathematics teachers’ perceptions, conceptions, and feedback practices. Data were analysed by structural equation modeling and content analysis techniques. Results from the structural equation model indicated that mathematics teachers’ perceptions of FA and AfL and their conceptions of assessment purposes positively predicted the quality of their feedback practices. Interview results illustrated that mathematics teachers used their students’ assessment information for both formative and summative purposes. Future interventions for improving the quality of mathematics teacher’s feedback practices are proposed.

1. Introduction

Formative assessment (FA) and Assessment for Learning (AfL) are widely acknowledged as powerful tools for effective instruction (Black & Wiliam, 1998; Ecclestone, 2012). Assessment as a formal or purposeful attempt to determine students’ performance during and/or after a learning phase can be used formatively for improving the teaching and learning process, certifying students, placement of students in tracks, or for curriculum improvement (Brown, 2008; Pellegrino, 2014). Based on the ten principles of AfL first drafted by the Assessment Reform Group (ARG, 2002), the most important practices that guide teachers’ implementation of AfL are: rich (classroom) questioning, feedback, peer assessment, self-assessment, and sharing learning goals and criteria of quality (Black & Wiliam, 1998, 2009; James & Pedder, 2006).

FA and AfL practices are assumed to serve two core functions: monitoring to track student progress and scaffolding to help students improve their learning (Pat-El, Tillema, Segers, & Vedder, 2013; Stiggins, 2005). Monitoring allows teachers to know the direction, speed, and quality of students’ learning progress so that supportive interventions can be put in place. Scaffolding is feedback that explicitly provides task, process, or self-regulatory information (Hattie & Timperley, 2007) so that students know how to proceed. However, recent research has shown that the implementation of FA and AfL is far from straightforward. For example, peer and self-assessment can be biased due to students’ intra- and interpersonal factors (Panadero, 2016), feedback is often ineffective (Lipnevich, Berg, & Smith, 2016), teachers do not always ask good questions (Airasian, 1997; Barnette, 1980), or actively promote feedback seeking (Winstone, Nash, Parker, & Rowntree, 2017). As a result, the nature of FA and AfL and how it leads to improved outcomes has been debated in FA and AfL literature (Bennett, 2011; Black & Wiliam, 2003).

1.1. What makes assessment formative?

The term ‘formative evaluation’ originates from Scriven’s (1967) distinction of formative evaluation to summative evaluation (Bennett, 2011; Black & Wiliam, 2003). According to Scriven (1967), summative evaluation provides information to judge the overall value of an educational program and formative evaluation refers to information provided early enough in the process so as to inform improvements. Bloom

* Corresponding author.
E-mail addresses: florence.kyaruzi@duee.ac.tz, sakyaruzi@gmail.com (F. Kyaruzi), j.w.strijbos@rug.nl (J.-W. Strijbos), ufer@math.lmu.de (S. Ufer),
gt.brown@auckland.ac.nz (G.T.L. Brown).

https://doi.org/10.1016/j.stueduc.2018.01.004
Received 24 October 2017; Received in revised form 22 January 2018; Accepted 24 January 2018
Available online 08 February 2018
0191-491X/ © 2018 Elsevier Ltd. All rights reserved.
(1969) shifted the initial focus of formative evaluation from ‘program evaluation’ to ‘student evaluation’. The purpose of formative evaluation was to provide feedback and corrections at each stage in the teaching and learning process (Bloom, 1969; Bloom, Hastings, & Madaus, 1971). Later on, ‘formative evaluation’ which was about testing and assessment, evolved into what is now referred to as ‘formative assessment’ (FA).

In our view, formative assessment is the thoughtful application of a purposefully selected methodology or instrument that fosters the interpretation of student performance to inform teachers and students about the learning progress (Bennett, 2011; Popham, 2014). Tests can be used to collect evidence that helps teachers or students to adjust their actions accordingly; however, tests themselves are neither formative nor summative (Popham, 2014). Nonetheless, rather than focusing on the evaluative or grading function of assessment, the present study focuses on the pedagogical use of FA and AfL as an instructional strategy to regulate both the teacher’s teaching processes and the students’ learning tactics and active use of feedback to improve outcomes. In this framework, the decisions made by the teacher and their students’ regarding assessment-elicited evidence make assessment formative. Decision-making depends in part upon beliefs that frame and guide thinking about a phenomenon (Fives & Buehl, 2012). Beliefs about assessment, which are rooted in their experience of assessment, influence the degree to which teacher assessment practices act formatively (Barnes, Fives, & Dacey, 2016; Fulmer, Lee, & Tan, 2015).

### 1.2. Teacher perceptions of FA and AfL practices and conceptions of assessment

While multiple terms are used to refer to the mental representations humans have for a phenomenon (e.g., perception, conception, belief, attitude, etc.), it seems likely these terms refer to the same thing (Brown, 2008). Hattie (2015) suggested that while Australian scholars use the term ‘beliefs’, those in the United States commonly use the term ‘epistemology’, while those in Europe prefer the term ‘conception’. In the present study, we consider the terms perceptions and conceptions separately—in particular because of the terminology used in survey instruments to refer to different content. However, it is acknowledged that others may consider perceptions and conceptions as largely synonymous.

In this study, teacher perceptions of FA and AfL are concerned with the way teachers evaluate their own practices to perform the core functions of monitoring and scaffolding student learning (e.g., “I adjust my instruction whenever I notice that my students do not understand a topic”). Monitoring practices entail analysing student learning progress to foster students’ self-monitoring by finding challenges and opportunities to optimise teaching and learning. Meanwhile, scaffolding involves teachers helping students to improve their learning by controlling elements of the task that are essentially beyond the student’s capacity (Pat-El et al., 2013). According to Wood, Bruner, and Ross (1976) scaffolding permits learners to concentrate upon and complete only those elements that are within their range of competence. Kyaruzi, Strijbos, and Ufer (2016) showed that students’ perceptions of their mathematics teacher’s monitoring and scaffolding practices were significantly related to their mathematics achievement.

In contrast, conceptions of assessment refer to the more general representations teachers have concerning the purposes of assessment (e.g., “Assessment improves learning”). It has been shown that teachers’ conceptions about the nature and purposes of assessment strongly influence how they teach and what students can actually learn or achieve (Dacey, 2015, 2017; Pajares, 1992). Three conceptions of assessment have been generally attested to, including (1) assessment improves teaching and learning, (2) assessment evaluates and holds accountable students, schools, and teachers, and (3) assessment is irrelevant (Bonner, 2016).

The conception that assessment improves teaching and learning is the central argument for FA and AfL (Black & Wiliam, 1998; Popham, 2014) and requires teachers to use evidence about student progress to support their learning (Brown, 2004). The conception that assessment makes schools and/or teachers accountable is the rationale behind accountability policies that use student assessment results to judge and reward or punish schools and teachers (Brookhart, 1994; Nichols & Harris, 2016). Student accountability is evidenced in the assignment of grades, checking off student performance against criteria, placing students into classes based on performance, as well as various qualification examinations for graduation or placement into further opportunities (Brown, 2008). The conception that assessment is irrelevant regards assessment as a negative practice, which, because it is unfair to students or inaccurate, can be ignored—even if it is imposed. If assessment cannot help teachers improve student learning, then teachers may choose to ignore it (Deneen & Brown, 2016).

Although studies into the role of teacher conceptions of assessment and assessment perceptions on teaching and learning processes have proliferated in the last two decades (e.g., Brown, 2004; Brown, Chaudhry, & Dhamija, 2015; Gibbs & Simpson, 2003; Macellans, 2001; Pat-El et al., 2013), few studies have examined African educational systems (e.g., Gebrig & Brown, 2014; Kitta, 2014; Ndichukwe, 2015). Furthermore, comparatively few studies provide accounts of teachers’ assessment perceptions and their determinants in mathematics education (e.g., Adams & Hsu, 1998; Al Duwairi, 2013; Ginsburg, 2009; Rach, Ufer, & Heinz, 2013).

### 1.3. Teacher feedback practices

FA and AfL literature provides extensive evidence that, if well implemented by teachers and well perceived by students, FA and AfL have the potential to improve student learning (Njabbi, 1999; Willam, Lee, Harrison, & Black, 2004; Willam, 2011), and especially for struggling learners (Black & Wiliam, 1998). Specifically, the quality of how teachers deliver feedback and how it promotes students to seek feedback (i.e., ‘feedback delivery’ and ‘promote feedback seeking’) is essential in contributing to both student outcomes and increased student regulation of their own learning. In fact, the more considerate a feedback source is when providing feedback, the more likely an individual is to accept and respond to the feedback provided (Strijbos, Pat-El, & Narciss, 2010; Dujinhouwer, Prins, & Stokking, 2012; Gregory & Levy, 2015; King, Schrodt, & Weisel, 2009). Considerate feedback, among other things, maximises clarity of information (Winstone et al., 2017).

An important goal of FA and AfL is to have students become active participants in assessment and active seekers of feedback. Feedback seeking requires students to identify areas in which they need help and seek feedback that aligns with their learning needs (Carless, Salter, Yang, & Lam, 2010). However, students are likely to seek feedback only if the social dynamics of the classroom or the teacher promotes feedback-seeking behaviours (Neitzel & Davis, 2014). Unfortunately, there is limited information on how teachers can promote students’ becoming active feedback seekers instead of being passive feedback receivers (Winstone et al., 2017). Thus, teacher feedback delivery and promoting students to seek feedback are important aspects of feedback practices that require further scrutiny.

It is also noteworthy that previous research has examined teacher perceptions of FA and AfL independent of their conceptions of assessment. While perceptions and conceptions may overlap considerably, there is a possibility that teacher perceptions of feedback practices may vary systematically with their conception of assessment. For example, Brown & Harris (2009) showed that when New Zealand primary teachers conceived of assessment as a measure of student accountability they perceived assessments as formal tests and measures of surface cognitive processes. Thus, it may be that teachers, who have a strong conception of assessment as an evaluation of students (perhaps consistent with systemic use of external examinations), would perceive feedback as accuracy on task-focused skills rather than supportive of
self-regulation skills. It may also be that such a conception would be associated with more self-oriented feedback practices in which unsuccessful students are blamed and successful ones praised. This would stand in contrast to a teacher who conceives of assessment as oriented towards improving teaching and thus uses student failure as a way to reflect on pedagogy and curriculum rather than the student. The present study is an early attempt at bringing together perceptions of feedback with conceptions of assessment.

1.4. Tanzanian context

The education system in Tanzania is characterized by high stake examinations which hold long-term implications to students’ lives. At the end of each instructional cycle of primary and secondary education levels, students participate in an external summative examination which is centrally administered by the National Examinations Council of Tanzania (NECTA). However, to overcome the overreliance on summative examinations Tanzania introduced in 1976 a Continuous Assessment (CA) program in secondary schools. It was intended to serve as a formative practice in secondary schools and to partly contribute to student’s final national examinations (NECTA, 2004; Njibili, 1999). The CA program emphasized that students should be continuously assessed and that the combined result should constitute a student’s success or failure (United Republic of Tanzania, 1974). It is implied that even school-based-assessment have impact on students’ final summative examination. Ottevanger, Akker, and Feiter (2007) pointed out that although most Sub-Saharan African countries – including Tanzania – have integrated school-based continuous assessment, testing at the school level was mainly summative and hardly used formatively, that is, for instructional purposes or to provide feedback to students. Zembazemba (2017) argued that most assessment practices in Tanzanian secondary schools are teacher-centred with low student involvement.

Teachers as key implementers of CA in schools are supposed to provide feedback to their students and help them bridge the gap between current performance and the desired standard. Teachers who are qualified to teach in secondary schools in Tanzania are supposed to possess a Diploma or Degree in Education or above (Ministry of Education and Vocational Training, MoEVT, 2014). Diploma in Teacher Education course comprises two years of pre-service teacher education and a Bachelor degree in Education comprises at least three years of pre-service teacher education. Teacher training colleges prepare Diploma teachers while Universities prepare Degree teachers. Teacher education programmes provide pre-service teachers with subject related courses and didactical courses including a teaching practicum.

Despite school improvement programmes such as the Secondary Education Development Programme (MoEVT, 2008), mathematics education in secondary schools in Tanzania has suffered from low passing rates (BEST, 2014; Kitta, 2004). Several studies have examined general educational challenges in Tanzania that might explain this: (a) the transition from Swahili as the language of instruction in primary schools to English in secondary schools (Brock-Une, 2007; Qorro, 2013; Vuzo, 2007), (b) large class (BEST, 2014) curriculum content overload (Kitta & Tilya, 2010), and (d) lack of in-service teacher professional development (Komba, 2007). Further challenges includes the lack of assessment skills to implement effective school based assessment (Osaki, Hosea, & Ottevanger, 2004). Since it seems logical to consider that teacher conceptions and/or perceptions will be consistent with the contextual demands in which the teacher is operating (Brown & Harris (2009)), these challenges may influence their conceptions and perceptions of their own teaching and assessment practices, including the quality of feedback practices.

1.5. The present study

Formative assessment practices ought to be well perceived by teachers because conceptions and perceptions of assessment frame and guide teacher assessment practices (Ajzen, 1991; Fives & Buehl, 2012). Hence, the present study investigates the relationship of Tanzanian secondary school mathematics teachers’ perceptions of FA and AfL, their conceptions of the purposes of assessment, and the quality of their feedback practices. Four research questions are posed:

1) To what extent do Tanzanian secondary school mathematics teachers perceive their own assessment practice in terms of the monitoring and scaffolding functions?
2) What conception of assessment do Tanzanian secondary school mathematics teachers report?
3) To what extent do Tanzanian secondary school mathematics teachers’ perceptions of their FA and AfL practices (monitoring and scaffolding) and their conceptions about the purposes of assessment (assessment improves learning, school accountability) relate to the quality of their self-reported feedback practices?
4) For what purposes do Tanzanian secondary school mathematics teachers typically use students’ assessment information (such as student’s scores in terminal and mid-term tests)?

2. Method

2.1. Participants

The study was conducted in 48 Tanzanian secondary schools. Just over half (n = 25) of the schools were in the mostly urban Dar es Salaam region and the remaining schools (n = 23) were in the mostly rural Kilimanjaro region. Based on national educational statistics (MoEVT, 2013) the mean GPA for the sampled schools in mathematics (M = 4.63, SD = 0.69) did not deviate statistically from the overall schools’ mean GPA (M = 4.85, SD = 0.70). Three criteria were used to achieve a representative sample of teachers from these 48 schools: school mathematics performance (high, medium, low) according to school ranking (MoEVT, 2013), average class-size (< 40, ≥40), and school-type (private, government). Within the 48 randomly sampled schools, which were drawn from schools of varying mathematics performance (N_high = 8, N_middle = 19, N_low = 27), 54 mathematics teachers in Grade 11 (Form three in the Tanzanian system) participated. The participants included nine women and 45 men, whose overall mean age was 37.26 (SD = 10.96; range 23 to 66). Teachers had on average close to 11 years of teaching experience (M = 10.87, SD = 10.39, range 1–36 years). In terms of highest qualification: 19 held a Diploma, 33 held a Bachelor degree and two held a Master degree. The average class size for the participating teachers was large (M = 49 students, SD = 20.49) and the average number of 40-min teaching periods per week (M = 22.05, SD = 7.33, range = 6–38 periods) meant that the average work-load was about 15 h per week. Grade 11 was selected because it should contain a wide variety of assessment and feedback practices relatively unconstrained by the official public examination.

2.2. Design

A mixed-method research approach was used, combining quantitative (survey) and qualitative (interviews) methods (Leech & Onwuegbuzie, 2009). Specifically, we employed a concurrent embedded design where qualitative and quantitative data were simultaneously collected and analysed to complement each other (Creswell, 2009; Dinglouli & Strijbos, in press). We complemented the quantitative analyses of survey data with content analysis of qualitative data from teacher interviews.

2.3. Instruments

2.3.1. Questionnaires

We used previously validated questionnaire scales for the survey,
which were adapted to the mathematics context by inserting the word ‘mathematics’ to ensure that teachers would reflect on their mathematics students. First, we used the Teacher Assessment for Learning Questionnaire (TAFLQ) to measure teacher’s perceptions of their AFL practice in terms of ‘perceived monitoring’ and ‘perceived scaffolding’ (Pat-El et al., 2013). Second, mathematics teachers’ conceptions of assessment were measured using the Teacher Conceptions of Assessment survey (TCoA-III) (Brown, 2004). This measure had nine 3-item first order factors which aggregate into four different conceptions: school accountability, improvement, irrelevant, and student accountability. Third, we adopted the ‘feedback delivery’ and ‘promoting feedback seeking’ subscales from the Feedback Environment Scale (FES) to measure the quality of teachers’ feedback practices (Steelman, Levy, & Snell, 2004). The various scales differed in response options (i.e., 5, 6, or 7) and were adapted to a common balanced 4-point scale: fully disagree (1), somewhat disagree (2), somewhat agree (3) and fully agree (4). We deliberately refrained from a middle category due to its ambiguous perceptual meaning (Dunham & Davison, 1991; Kulas & Stachowski, 2009). It should be noted that the ‘teacher conceptions of assessment subscales’ and ‘feedback environment subscales’ were below the 0.70 threshold for Cronbach’s alpha, which could be due to low sample size, suggesting that robust analyses methods would be required to draw valid conclusions—therefore our results should be interpreted cautiously. Table 1 summarises the scales, number of items per scale, a sample item per scale, and the Cronbach’s $\alpha$ from the original studies and for this study.

### 2.3.2. Teacher interviews

The interview questions were specifically developed for the present study to gain in-depth understanding of the topics covered in questionnaire scales. The interview focused on two main goals: (a) teachers’ teaching, assessment, and feedback practices such as: teacher reactions to student errors, teacher perceptions of FA and AFL practices, and (b) teachers’ perception of student experiences with teaching, assessment, and feedback such as: perceptions of student reactions to teacher feedback. For this study, teachers’ responses to the question “For what purposes do you typically use students’ assessment information (e.g., student’s scores in terminal and mid-term tests)” were analysed. The average duration of the interviews was 27 min ranging from 17 to 54 min.

### 2.4. Procedure

The study was conducted with ethics clearance from the University of Dar es Salaam. All participating teachers signed an informed consent form. Questionnaires were administered by the researcher or by one of two research assistants. The researcher or assistant demonstrated how to use the rating scales prior to the teachers filling-in the questionnaire. The teachers needed approximately 20 min to do so.

Prior to data analyses, data screening was carried out to identify and address outliers (univariate and bivariate), as well as missing value analysis and recoding of all negatively phrased items. Data were considered to be missing completely at random (MCAR), because Little’s MCAR test was not statistically significant ($\chi^2 = 101.67$, $df = 1732$, $p = .00$) (Peugh & Enders, 2004). We imputed missing values using expectation maximization (EM), which is an effective imputation method when data are MCAR (Musil, Warner, Yobas, & Jones, 2002). Investigation of the EM estimated statistics such as item means showed minimal differences to the un-estimated data (i.e., differences in descriptive statistics were at the third decimal point).

### 2.5. Analyses

#### 2.5.1. Questionnaire analyses

Ideally, confirmatory factor analysis should be used to ensure that the factor structure of an existing measurement applies equally to a new sample (Brown, Harris, O’Quin, & Lane, 2017). However, with just 54 teachers and 64 items the study did not meet the expectation of 5–10 cases per variable (Costello & Osborne, 2005). Hence, the existing scale structure was tested with this sample using Cronbach’s alpha values to identify groups of items that replicated the original scales. Where the scale (e.g., TCoA Irrelevance) did not generate a plausible alpha value (i.e., $\alpha > 0.40$), information about the possible alpha value after deletion of an item was used to create defensible scales.

To identify the relationship of instruments to each other it was decided to parcel (Little, Cunningham, Shahar, & Widaman, 2002) the manifest item responses into factor mean scores, according to the CFA results, so that the ratio of cases to variables was 54:9 (i.e., 6:1) meeting conventional standards. This approach meant that structural equation modeling (SEM) of the relationship of factors to each other with a relatively small sample could be conducted. The application of SEM with small samples reduces the chance of correlated residuals and sampling errors (Marsh, Hau, Balla, & Grayson, 1998).

SEM was used to estimate the impact of mathematics teachers’ perceptions of their FA and AFL practices and conceptions of assessment on the quality of their feedback practices. The SEM approach was preferred over normal regression because it provides a stronger framework to account for response bias and takes into account non-random measurement errors (Comsa, 2010). Furthermore, SEM is preferred to path analysis since latent constructs can be retained in a model, rather than only manifest variables.

The evaluation of CFA and SEM model fit requires reporting

---

**Table 1**

<table>
<thead>
<tr>
<th>Scale</th>
<th>k</th>
<th>Sample item</th>
<th>Cronbach’s $\alpha$</th>
<th>Original Study</th>
<th>Present Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Assessment for Learning Questionnaire</td>
<td></td>
<td>Perceived monitoring 16 I ask my students to indicate what went well and what went badly concerning their assignments.</td>
<td>0.87</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>Perceived scaffolding 12 I adjust my instruction whenever I notice that my students do not understand a topic.</td>
<td>0.77</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>Teacher Conceptions of Assessment</td>
<td></td>
<td>School accountability 3 Assessment is a good way to evaluate a school.</td>
<td>0.77</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>Improvement 9 Assessment helps students improve their learning.</td>
<td>0.83</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Assessment is Irrelevant 3 Assessment interferes with teaching.</td>
<td>0.69</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Student Accountability 3 Assessment places students into categories.</td>
<td>0.55</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Ignore Assessment 3 Assessment results are filed and ignored.</td>
<td>0.69</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>Feedback Environment Scale</td>
<td></td>
<td>Feedback delivery 5 I am supportive when giving my students feedback about their mathematics performance.</td>
<td>0.86</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Promote feedback seeking 4 I encourage my students to ask for feedback whenever they are uncertain about their mathematics performance.</td>
<td>0.84</td>
<td>0.45</td>
<td></td>
</tr>
</tbody>
</table>

*Note. $k =$ number of items per scale.*
multiple indicators and determining whether fit is acceptable or good (Hu & Bentler, 1999). Unfortunately, not all indicators have been shown to be stable as model complexity, sample size, or model mis-specification occurs (Fan & Sivo, 2007). The chi-square statistic is overly sensitive in sample sizes above 250 (Byrne, 2010), the comparative fit index (CFI) rewards models with three or fewer factors, while the root mean square error of approximation (RMSEA) rewards complex models with more than three factors (Fan & Sivo, 2007). In contrast, the gamma hat and the standardised root mean residual (SRMR) have been shown to be stable estimators (Fan & Sivo, 2007).

Acceptable fit is indicated when CFI and gamma hat > .90, RMSEA and SRMR < 0.08, and the ratio of χ²/df has p > .05. Good fit is indicated when CFI and gamma hat > .95, RMSEA and SRMR < 0.05 and ≤.06 respectively. As recommended by Steiger (1990), the 90% confidence interval for RMSEA can be used to determine whether the observed value may be < 0.05.

2.5.2. Interview analyses

The interviews were content analysed (Braun, & Clarke, 2006; Strijbos, Martens, Prins, & Jochems, 2006). A data-derived coding scheme was developed using about ten percent of all interviews. The threshold for segmentation agreement was 80% (Strijbos et al., 2006) and a Krippendorff alpha of 0.80 or higher indicated good coding reliability (Krippendorff, 2013). Two independent coders were involved in all data analysis after a 50 min training session on the study rationale and the coding scheme. Four iterations of independent coding were performed. After these rounds, it was deemed that codes had been appropriately and reliably assigned to interview utterances (Table 2).

3. Results

3.1. Mathematics teachers’ perceptions of their FA and AFL practices

The first research question sought to describe mathematics teachers’ FA and AFL perceptions, conceptions of assessment, and feedback environment (Table 3). Scale analysis showed that the two long scales for TAFLQ had strong alpha values, the two shorter FES scales had moderate estimates, and three of the TCoA scales had moderate estimate values. The TCoA Irrelevance factor had to be split intoIgnore and Irrelevant scales, each with three items and moderate estimate values. Hence, in total, 58 items were used to create nine scales.

Table 2: Interview segmentation agreement and Krippendorff alpha.

<table>
<thead>
<tr>
<th>Trial</th>
<th>Size</th>
<th>Segmentation (%)</th>
<th>Krippendorff alpha</th>
<th>90% C.I</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Coder 1</td>
<td>Coder 2</td>
<td>Value</td>
</tr>
<tr>
<td>1</td>
<td>7 interviews</td>
<td>67</td>
<td>70</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>6 interviews</td>
<td>83</td>
<td>86</td>
<td>0.63</td>
</tr>
<tr>
<td>3</td>
<td>6 interviews</td>
<td>71</td>
<td>74</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>6 interviews</td>
<td>87</td>
<td>89</td>
<td>0.88</td>
</tr>
</tbody>
</table>

3.2. Teachers’ FA and AFL perceptions, conceptions of assessment, and quality of feedback practices

Table 3 reveals that mathematics teachers’ assessment perceptions were above ‘somewhat agree’ (3.00) for all scales except for conceptions of assessment as irrelevant and ignore assessment, suggesting that the mathematics teachers evaluated positively their own FA and AFL practices, conceptions of assessment, and the quality of feedback practices. The lower mean scores for irrelevant and ignore assessment conceptions indicates that teachers did not ignore assessment nor did they consider assessment to be irrelevant. Furthermore, the inter-correlations indicate that mathematics teachers’ FA and AFL perceptions had moderate positive correlations with feedback environment and conceptions of assessment scales (improvement, student accountability and school accountability), except for the relationship between perceived scaffolding, school accountability and student accountability—which was not significant. The association of AFL perceptions (monitoring and scaffolding) with the conception of assessment for improvement, suggests reasonable convergence around learning-oriented assessment practices and philosophy. However, the ignore assessment conception had moderate negative significant correlation with monitoring, and a small negative correlation with feedback delivery. This relationship indicates that for teachers to effectively monitor student learning and deliver feedback in a thoughtful manner they should not ignore assessment information.

Since both the TAFLQ and TCoA inventories related to perceptions or conceptions of assessment respectively, it was decided to model these self-reported beliefs as correlated predictors of the FES as the dependent variable. The structural model was motivated by the theoretical evidence underpinning that conceptions and perceptions of assessment have the potential to influence teacher assessment practices (Ajzen, 2005; Fives & Buehl, 2012). Hence, this model was tested in SEM (Fig. 1). The model had good fit: χ² = 40.86; df = 25; χ²/df = 1.63 (p = .024); CFI = 0.884, Gamma hat = 0.938, SRMR = 0.085, and RMSEA = 0.109 [0.040, 0.168]. Mathematics teachers’ conceptions of assessment purposes were moderately and positively correlated with their perceptions of AFL (r = 0.55). The regression from AFL practices to Feedback Environment was strong (β = 0.76), whereas the path from TCoA to FES was statistically not significant (and therefore not shown in Fig. 1). Combined, teacher’s conceptions of assessment and perceptions of AFL accounted for a large portion of variance in the Feedback Environment factor (SMC = 0.57, f^2 = 1.32; Cohen, 1992).

3.3. Mathematics teachers’ use of students’ assessment information

With the help of the interview, we aimed to investigate how the mathematics teachers used students’ assessment information. Essentially, this was motivated by the claim that it is the purpose for which the assessment information is used (by the teacher and students) that makes the assessment to be formative. Analyses of interviews resulted in six main themes on how mathematics teachers use their students’ assessment information: (1) to show students how to improve, (2) to devise their teaching approaches, (3) to categorise students into ability groups, (4) to compose accountability reports (to parents, school authority, etc.), (5) to motivate high achievers, and (6) to reprimand low achievers (see Table 4).

These six main themes reflect a mix of formative and summative purposes. Teachers reported formative use of student assessment information such as reflections on their teaching practices, improving their teaching approaches, correcting student errors and conducting remedial classes to support weaker students. They also reported a summative use of student assessment information such as ability grouping (if no specific support was provided to each ability group), accountability reports, and using assessment to reprimand low achievers.

Using assessment information as a motivation for high achievers may have positive impact on learning, if and only if, it leads to positive changes in students’ effort, engagement or self-efficacy (Kluger & DeNisi, 1996). Furthermore, it was noted that ability grouping can be a formative practice when intended to provide students with extra support as evidenced in the remark “In our school we identify and separate slow learners so that they can get a special attention; they are special classes” (Teacher 8). Ability grouping was a non-formative practice if it was used solely for ranking students as evidenced in the remark, “We normally use student’s scores first of all in ranking students” (Teacher 53).

In general, the interview results support the questionnaire results in Table 3 in various ways. First, teachers used assessment information to show students how to improve and to devise their teaching approaches,
Table 3
Descriptive statistics per scale and their inter-correlations.

<table>
<thead>
<tr>
<th>Descriptive</th>
<th>Scale inter-correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>AfL Perceptions</td>
<td></td>
</tr>
<tr>
<td>I. Monitoring</td>
<td>2.44 3.61 0.31</td>
</tr>
<tr>
<td>II. Scaffolding</td>
<td>2.75 3.64 0.32</td>
</tr>
<tr>
<td>Conceptions of Assessment</td>
<td></td>
</tr>
<tr>
<td>III. Improvement</td>
<td>2.33 3.40 0.40</td>
</tr>
<tr>
<td>IV. Student Accountability</td>
<td>2.33 3.45 0.49</td>
</tr>
<tr>
<td>V. School Accountability</td>
<td>2.00 3.49 0.51</td>
</tr>
<tr>
<td>VI. Irrelevant</td>
<td>1.33 2.55 0.70</td>
</tr>
<tr>
<td>VII. Ignore Assessment</td>
<td>1.00 2.00 0.61</td>
</tr>
<tr>
<td>Feedback Environment Perceptions</td>
<td></td>
</tr>
<tr>
<td>VIII. Feedback delivery</td>
<td>2.50 3.49 0.43</td>
</tr>
<tr>
<td>IX. Promote Feedback seeking</td>
<td>2.50 3.54 0.44</td>
</tr>
</tbody>
</table>

| Note. | Values in bold are within inventory correlations; N = 54; Maximum score = 4.00; **p < .01, *p < .05. |

Fig. 1. Teacher's FA and AfL perceptions and assessment conceptions as predictors of the quality of their feedback practices.

Table 4
Mathematics teachers’ (N = 54) use of their students’ assessment information.

<table>
<thead>
<tr>
<th>Key themes</th>
<th>Interview excerpts</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Show students how to improve (44%)</td>
<td>I use assessment information to: Do corrections to all students in the class and I normally involve students who are doing better to do corrections on the board so that other students can be encouraged (Teacher 38).</td>
</tr>
<tr>
<td>2. Devise teaching approaches (30%)</td>
<td>I use assessment information to: Evaluate myself if what I taught was understood by my students or not. If students perform poor, I prepare remedial classes so that I can re-teach students who scored below the average (Teacher 25).</td>
</tr>
<tr>
<td>3. Ability grouping (20%)</td>
<td>In our school we normally use students’ scores first of all in ranking students. Secondly, student’s scores are bases for student promotion or retention in the same class (Teacher 53). In our school we identify and separate slow learners so that they can get a special attention; they are special classes (Teacher 8).</td>
</tr>
<tr>
<td>4. Accountability reports (17%)</td>
<td>Normally assessment analysis goes far to inform parents (Teacher 5). I use assessment to: Collect marks for Continuous Assessment (CA) in order to meet our school development (Teacher 17).</td>
</tr>
<tr>
<td>5. Motivate high achievers (17%)</td>
<td>Sometimes we award best students and try to assist slow learners. In awarding the best students it helps the slow learners to work hard (Teacher 52).</td>
</tr>
<tr>
<td>6. Reprimand low achievers (4%)</td>
<td>I always use assessment results to reprimand students who drop in their performances (Teacher 46).</td>
</tr>
</tbody>
</table>
which is supported by their high rating in the conception that assessment is to improve (teaching and learning). Second, the observed teacher use of assessment for ability grouping supports the assessment is for student accountability conception. Third, teacher use of student assessment information for accountability reports (such as student reports to inform parents) supports their high agreement with the conception that assessment is for school accountability. In general, the mathematics teachers self-reported uses of their students’ assessment information in the interviews corresponds to their conceptions of assessment in the questionnaire.

4. Discussion

This study investigated the effect of mathematics teachers’ FA and AFL perceptions and conceptions of assessment on the quality of their feedback practices. The first research question sought to investigate the extent to which Tanzanian secondary school mathematics teachers perceived their assessment practice as formative in terms of the monitoring and scaffolding of student learning. The mathematics teachers had a positive perception, indicating that they perceived their own assessment practices as formative. Moreover, this indicates that Tanzanian mathematics teachers value their own FA and AFL practices, which replicates findings from previous studies on teacher perceptions of their own assessment practices (e.g., Rach et al., 2013; Pat-El, Tillema, Segers, & Vedder, 2015; Veldhuis & van den Heuvel-Panhuizen, 2014). The second research question investigated the conceptions of assessment reported by Tanzanian secondary school mathematics teachers. Our results indicate that mathematics teachers had positive conceptions of assessment, i.e. that the purpose of assessment was to improve student learning, promote student accountability and school accountability. However, teachers did not agree that assessment is irrelevant nor ignored assessment information. This is consistent with previous studies indicating that teachers consider the purpose of assessment to be that of improving student learning and promoting school accountability (Brown, 2004, 2006, 2008; Barnes et al., 2017).

The third research question investigated the extent to which Tanzanian secondary school mathematics teachers’ perceptions of their own FA and AFL practices and conceptions of assessment predicted the quality of their feedback practices. The structural equation model indicates that mathematics teachers’ perceptions of FA and AFL and their conceptions of assessment were highly correlated, and combined they strongly predicted the quality of their feedback practices. These findings support previous studies that perceptions of assessment are related to teacher assessment practices (Fives & Buehl, 2012). Moreover, these findings are consistent with Van de Pol, Oort, Volman, and Beishuizen (2014) who found that scaffolding is an important practice for improving teacher assessment practices and student learning. Furthermore, the findings on the impact of teacher perceptions of FA and AFL on their assessment practices is in line with recent findings that teachers’ perceptions of peer assessment (part of AFL) predicted how often teachers used peer assessment (Panadero & Brown, 2017). Likewise, our results are line with recent findings indicating that teachers’ positive perceptions and/or experiences with self-assessment (also part of AFL) influence their assessment practices (Panadero, Brown, & Courtney, 2014). Hence, consistent with previous studies, our findings support that teacher feedback practices (feedback delivery and promoting feedback seeking) are shaped by their conceptions about assessment purposes and perceptions of FA and AFL.

The fourth research question sought to identify typical uses of student’s assessment information by Tanzanian secondary school mathematics teachers. In line with Al Duwairi (2013), our interviews showed that mathematics teachers used students’ assessment information for both formative and summative purposes. The reported formative uses of student assessment information (i.e., improving student learning and instruction) are in line with how Canadian teachers in various instructional domains conceive the functions of assessment (Hunter, Mayenga, & Gambell, 2006). Furthermore, the interviews indicate that mathematics teachers’ assessment practices were rooted in their conceptions of assessment purposes. For example, the majority of mathematics teachers reported to use their student assessment information to reflect on their teaching approaches and to provide feedback on their students’ learning; both activities are considered core elements of a formative assessment practice (Black & Wiliam, 2003; Ginsburg, 2009; Hattie, 2009). The observed role of teacher conceptions of assessment adds to previous studies that assessment improves teaching and learning (Brown, 2004; Brown et al., 2015). This is also consistent with Ndaliachako (2015) who reported that 60% of 2047 Tanzanian secondary school teachers agreed that the purpose of classroom assessment was to improve teaching and learning processes. Finally, the reported summative uses of assessment information such as accountability reports to parents and students’ ability grouping, provide further support that conceptions of assessment promote student and school accountability (Brown, 2006; Firestone, Mayrowetz, & Fairman, 1998). Nevertheless, how mathematics teachers balance summative and formative uses of student assessment information is an important issue for future research.

4.1. Methodological limitations

Our results should be interpreted in light of some limitations. Firstly, we mainly used self-report data from surveys and teacher interviews, which might be limited in their scope—although our data provide evidence for their validity. Hence, future research could further substantiate our findings with other measures such as observational data. Secondly, the reliability of the ‘conceptions of assessment scales’ and ‘feedback environment scales’ were below the typical threshold (i.e., Cronbach’s α > 0.70), which indicates that our results should be interpreted cautiously. However, we applied structural equation modeling with parceling which is a robust technique for small sample sizes and takes into account random and non-random measurement errors. Thirdly, based on the relatively small sample, the models’ multiple fit indicators were acceptable to good; hence, we cannot generalise our findings beyond this sample. These results may be substantiated by future studies using observational and/or longitudinal data to examine other potential factors that might influence the quality of teacher feedback practices. Fourthly, we adopted the scales with different response options (i.e., 5, 6 or 7) to a common balanced 4-point scale which might have influenced item variance, resulting in low reliability for some scales.

4.2. Theoretical and practical implications

Our results showed that the mathematics teachers were aware that effective formative assessment demands both teachers and students to reflect on the assessment information. However, if mathematics teachers only reflect on this information but students do not utilize the feedback provided by their teachers, FA and AFL practices are apt to fail (Pat-El et al., 2013). Our results indicate that mathematics teachers had positive perceptions of their own FA and AFL practices and conceptions of assessment, and that combined they predicted the quality of their feedback practices. Thus, these results support the planned behaviour theory that conceptions (beliefs) influence behaviour (Ajzen, 1991). Furthermore, the interviews showed that the mathematics teachers reported various uses of student assessment information (Gronlund & Linn, 1990). These uses of assessment information aligned with established teacher conceptions of assessment purposes that assessment improves teaching and learning that promotes school and student accountability (Black & Wiliam, 1998; Brown, 2006, 2011). Surprisingly, a few mathematics teachers used their student’s assessment information to reprimand low achievers, which might be attributed to the high-stakes assessment culture in Tanzania. Hence, future research could
investigate how teachers can be encouraged to use student errors in mathematics tests or assignments to inform students on how to improve (Rach et al., 2013), or provide educational counseling instead of reprimanding low achieving students (Yaghame & Tshabangu, 2013). Likewise, interventions could be designed and implemented to improve the quality of teacher feedback practices to capitalize on teacher conceptions of assessment and perceptions of their FA and AF practices.

References


Barnette, J., Orletsky, S., & Sattes, B. (1994). Evaluation of teacher classroom questioning practices in mathematics tests or assignments to inform students on how to improve (Rach et al., 2013), or provide educational counseling instead of reprimanding low achieving students (Yaghame & Tshabangu, 2013). Likewise, interventions could be designed and implemented to improve the quality of teacher feedback practices to capitalize on teacher conceptions of assessment and perceptions of their FA and AF practices.

Likewise, interventions could be designed and implemented to improve the quality of teacher feedback practices to capitalize on teacher conceptions of assessment and perceptions of their FA and AF practices.

References


Barnette, J., Orletsky, S., & Sattes, B. (1994). Evaluation of teacher classroom questioning practices in mathematics tests or assignments to inform students on how to improve (Rach et al., 2013), or provide educational counseling instead of reprimanding low achieving students (Yaghame & Tshabangu, 2013). Likewise, interventions could be designed and implemented to improve the quality of teacher feedback practices to capitalize on teacher conceptions of assessment and perceptions of their FA and AF practices.

Likewise, interventions could be designed and implemented to improve the quality of teacher feedback practices to capitalize on teacher conceptions of assessment and perceptions of their FA and AF practices.

References


Barnette, J., Orletsky, S., & Sattes, B. (1994). Evaluation of teacher classroom questioning practices in mathematics tests or assignments to inform students on how to improve (Rach et al., 2013), or provide educational counseling instead of reprimanding low achieving students (Yaghame & Tshabangu, 2013). Likewise, interventions could be designed and implemented to improve the quality of teacher feedback practices to capitalize on teacher conceptions of assessment and perceptions of their FA and AF practices.

Likewise, interventions could be designed and implemented to improve the quality of teacher feedback practices to capitalize on teacher conceptions of assessment and perceptions of their FA and AF practices.

References


Barnette, J., Orletsky, S., & Sattes, B. (1994). Evaluation of teacher classroom questioning practices in mathematics tests or assignments to inform students on how to improve (Rach et al., 2013), or provide educational counseling instead of reprimanding low achieving students (Yaghame & Tshabangu, 2013). Likewise, interventions could be designed and implemented to improve the quality of teacher feedback practices to capitalize on teacher conceptions of assessment and perceptions of their FA and AF practices.

Likewise, interventions could be designed and implemented to improve the quality of teacher feedback practices to capitalize on teacher conceptions of assessment and perceptions of their FA and AF practices.

References


Barnette, J., Orletsky, S., & Sattes, B. (1994). Evaluation of teacher classroom questioning practices in mathematics tests or assignments to inform students on how to improve (Rach et al., 2013), or provide educational counseling instead of reprimanding low achieving students (Yaghame & Tshabangu, 2013). Likewise, interventions could be designed and implemented to improve the quality of teacher feedback practices to capitalize on teacher conceptions of assessment and perceptions of their FA and AF practices.


