CHAPTER SIX

DESIGN, DEVELOPMENT AND EVALUATION OF THE FIRST AND SECOND PROTOTYPES

6.1 INTRODUCTION

The aim of the study is to understand and explore the characteristics and quality processes needed in the performance assessment of Agriculture Form Four students to ensure valid and reliable examinations in Botswana. Findings from Chapter 5 and guidelines from the literature review (Section 3.3) guided the design and development of the intervention as outlined in Section 6.3 taking into consideration the philosophy of learner-centred approach. The literature revealed how performance assessment should be conducted to improve students' performance (Wiggins, 1998). The review of studies conducted in Africa employing design-based research to develop exemplar curriculum materials emphasised the design of the instructional materials (Mafumiko 2006; Tecle 2006; Tilya, 2003), with little or no emphasis on assessment for summative purposes, with the exception of Januario (2008) and Motswiri (2004).

While Motswiri's and Januario's studies investigated assessment, they concentrated on improvement of formative assessment for learning which did not contribute to certification. This study therefore sought to design an assessment intervention which served the dual purposes of improving learning (formative) and evaluating learning (summative). For such an intervention to be widely adopted, it was designed and developed iteratively in collaboration with practitioners and stakeholders (Dick, Carey & Carey, 2009), and involved the development of both the task and the assessment instrument.

In this chapter, product design specifications are outlined in Section 6.2. The description of the first prototype development is presented in Section 6.3, while Section 6.4 delineates formative evaluation of the first prototype. Section 6.5 presents experts views regarding the first prototype. Conclusion of the first prototype is presented in Section 6.6 while implication for design of the next prototype is outlined in Section 6.7. Based on the experts review, conclusions and implications of the first prototype, the design of the second prototype is described in Section 6.8, while its formative evaluation is presented in Section 6.9. The results of the evaluation of the second prototype are presented in Section 6.10. Sections 6.11

and 6.12 delineate the conclusion and implications for the subsequent design of the prototype respectively.

6.2 PRODUCT DESIGN SPECIFICATIONS

The baseline survey established the needs for stakeholders in performance assessment (Rainey, 2005), to be integrated in the design and development of quality assurance processes (Abramowich, 2005). It emerged from the baseline findings that there is a need to develop standardised tasks together with their assessment instruments to be used by all teachers throughout the country. When developing tasks, consideration should be given to the current state of resources provision in schools as well as teachers' training in performance assessment. Teachers' workload was high due to large class sizes hence time was an important factor to be factored in the design of the materials. The tasks to be developed should be student-centred to engaged students in the meaning creation of their learning with the hope to improve their negative attitude towards performance assessment. The design also drew heavily on the literature review regarding performance assessment best practices internationally. The conceptual framework provided the roadmap which outlined guidelines and specifications for design of formative evaluation of the assessment prototypes. The product design specifications are delineated below:

1. The tasks should be complex and engaging

The tasks developed should address content of importance and substance, and designed in the form of investigations, portfolios and performances, which involved problem-solving that in turn results in report-writing, (Ariev, 2005; Diez, 2002; Macmillan, 2004; Maxwell, 2004: Rennert- Ryan, 2006), rather than just using traditional paper-and-pencil tests (Macmillan, 2004). Such tasks encourage divergent thinking resulting in multiple correct answers to real-world problems. Complex and demanding tasks allow the fulfilment of the primary purpose of improvement in student learning leading to excellence (Wiggins, 1998). Complex tasks would be developed to last longer, encompassing many domains of cognitive, psychomotor and affective skills (Nitko & Brookhart, 2007), for students to use varied multiple skills (Airasian & Russell, 2008; Gardner, 2006).

2. Assessment should be integrated into instruction

The traditional standardised testing in which students answer uniform questions in an artificial environment (Wiggins, 1998) are normally designed to audit learning (McMillan 2000; Shepard, 2000), and as such could not be integrated into instruction. School-based assessment of performance tasks has the potential to improve learning if conducted properly and integrated into instruction, as it reveals students' strengths and weaknesses which served as inputs in designing appropriate remedial actions by the teacher. . Performance assessment would be infused in a normal lesson, allowing for assessment to be done when students are ready (Harlen, 2006) and reassessment whenever they did not do well. To assist in having meaningful participation by students, they would be alerted to the teachers' expectations and given assessment materials in advance to familiarise themselves.

3. Assessment should be aimed at both processes and products

Product assessment is a crucial aspect of performance assessment, especially where the procedure has been mastered by students (Gronlund, 2006). However, bias towards assessment of the product has the potential to conceal students' capabilities in other domains, such as manipulation. Dick et al. (2009) posit that to determine if learners have achieved an attitude, they have to do something, namely a psychomotor, intellectual or verbal skill. For those skills of which there is only temporary evidence, it is important that one assesses the processes, as it is believed that the repeated use of these improves the product (McMillan, 2004). To effectively assess thinking processes the students undergo in constructing their responses (Airasian, 2005), performance assessments are designed for use under varying contexts, to present all students with the opportunity to showcase their skills. Sometimes it is difficult to prescribe whether to carry out product or process assessment. In such cases it is left to the teacher to use his/her professional judgement as to when to assess product and when to assess the processes to balance the two (see Gronlund, 2006).

4. Assessment should be authentic

Student-centred learning takes place in a context in which real life problems manifest themselves in varied forms, and require pragmatic approaches to their solution. Rather than crafting standardised practical tests to be administered to all students throughout the country (Resnick & Rescnick, 1992), a variety of authentic tasks would be developed to be applied in

the prevailing context (Johnson et al., 2009; Nitko & Brookhart, 2007). To judge the degree of authenticity of the tasks, Wiggins' (1998:22-24) six standards would be applied:

> *Are the tasks realistic*? Do the tasks replicate the ways in which a person's knowledge and abilities are tested in real world situations?
>
> *Do the tasks require judgement and innovations?* Does the student have to use knowledge and skills wisely and effectively to solve unstructured problems, and does the solution involve more than following a set of routine?
>
> *The tasks should involve the students doing something.* The student has to carry out the exploration and work within the discipline of the subject area, rather than restating what was already known or taught.
>
> *Replicates or simulates the context in which adults were tested in the workplace, in civic life, and in personal life.* Do contexts involve specific situations that had particular constraints, purposes, and audiences?
>
> *The tasks should require the student's ability to efficiently and effectively use a repertoire of knowledge and skills to do complex tasks.* Students are required to integrate all knowledge and skills needed, rather than to demonstrate competence of isolated knowledge and skills.
>
> *The tasks should allow appropriate opportunities to rehearse, practice, consult resources, and get feedback on and refine performances and products.*

5. The tasks should be feasible given the resources available in schools

Tasks would be designed to foster collaboration and cooperation to cater for both inadequate time and resources availability in schools (Subsection 5.3.4), as well as large classes (Subsection 5.2.4). Tasks should not only doable but also developmentally appropriate to develop thinking in a variety of ways (Wiggins, 1998). Some would be designed as simulations to serve as an intermediate step to performances that are complicated, involving a higher degree of realism, requiring expensive equipment, or those that put other people's lives in jeopardy (Popham, 2005). For example, instead of children under the age of 16 applying chemicals to crops, they would use water, as they are not legally allowed to use chemicals. Alternatively, they could collaborate with older students.

6. Tasks should be evaluated analytically and holistically

Complex performances require that several learning targets or several parts of the performance be assessed using several scoring rubrics consistently (Johnson, et al., 2009) to eliminate subjective scoring (Arter & McTighe, 2001). Some tasks would be crafted to be assessed holistically while others would be assessed analytically. A holistic task is one which is scored using a scale containing several criteria, yielding a single score that gives an overall impression or rating (McMillan, 2004), while a task requiring analytical scoring is one in which each scoring criterion receives a separate score (Nitko & Brookhart, 2007). Because holistic tasks offer little information that can be used for formative feedback, tasks which are scored analytically are inevitable. Analytic scoring separates the whole into parts but takes longer to create and score (McMillan, 2004; Popham, 2005). Each of the holistic and analytic tools would be applied in different situations. The assessment guidelines would be developed simultaneously with tasks (Johnson et al., 2009).

7. The assessment should be continuous and cumulative in nature

Evidence of learning is normally collected over time and in the form of a student portfolio (Maxwell, 2004). Assessment would be continuously conducted and integrated in the learning process (McMillan 2000), using multiple methods and raters (McIntire & Miller, 2007; Thorndike & Thorndike-Christ, 2010), with the opportunity for reassessment to approximate the student's true score (Raffan, 2000).

8. Self-Assessment

Performance assessment is an open activity (Njabili, 1987), hence assessment materials should be shared with students in advance, as well as parents and the public at large. Nitko & Brookhart (2007) assert that for the students to have their attitudes adequately assessed they should be provided with information about why they should act in a certain way. This helps in shaping an attitude and thereby increasing the chances that desired behaviour will be demonstrated. Having students to assess themselves (Harlen, 2006) helps them to reflect on their performance by applying established criteria to judge their own work. Their self-assessment also helps teachers in formulating sound corrective actions.

9. Assessment should produce a traceable evidence of assessing

Retrievable and traceable records should be kept (Le Grange & Reddy, 1998) for assessment to be used for different purposes, and different offices of the Ministry of Education make use of them. To increase the reliability and validity of records, the performance assessment tasks should be designed to lend themselves to minimal record keeping.

6.3 DEVELOPMENT OF THE FIRST PROTOTYPE

The description of the tasks that were developed is presented in this section. Tasks development was based on three content areas of *Preparing a plot and planting*, *Applying fertiliser as basal dressing*, and *Controlling weeds using chemicals*.

6.3.1 Description of tasks

The selection of subject content was based on what schools were offering. At the time of conducting this research, schools were offering *Field Crop production.* Tasks from this content were inevitable for the study to operate in synergy with the school's programme, so as to cause minimal disruptions as per the requirement of the Permission Letter from the Ministry (Appendix 4.12). Naturally, the three tasks varied, with *Applying fertiliser as basal dressing* entailing in essence activities with temporary evidence, dictating that assessment involve mainly observations of processes. Meanwhile, *Preparing a plot for planting* and Controlling *weeds using chemicals* involved both observation and product assessment. As a result, it was critical that the development of the intervention captured both.

Task 1: *Preparing the plot and planting*: The task has five skills, some of which could not be repeated once they had been performed (Appendix 6.3). Given the class sizes of 35 students (Section 5.2.4), it is difficult to observe all students in these skills, so those who could not be assessed could be assessed in others of similar demand.

Task 2: *Applying fertiliser as basal dressing:* The task has six skills which involve mainly assessment of activities (observation) and record keeping (see Figure 6.3 and Table 6.1). There is little or no product assessment and most of the skills could not be repeated once they had been completed. This creates a challenge for the teacher to assess as many students as possible, as the task could be completed in 120 minutes.

Task 3: *Controlling weeds using chemicals:* The task has six skills which involve assessment of activities (observation), the product, and record keeping (Appendix 6.4). This is a typical example of simulation. It could be carried out at any time and repeated as desired; hence its timeframe was not limited.

6.3.2 Skills equating

Because of the large classes, it was not possible to assess all students in each skill. To circumvent the problem of assessing students in skills of different demands, skills were equated as shown in Figure 6.1 (below). Using Task 1 as an example, skills 1 and 3 were of equivalent highest demand. Some students could be assessed in skill 1 while others could be assessed in skill 3. Skill 5 was of average demand. Similarly, skills 2 and 4 were of equivalent lowest demand. Skills equating for Task 2 and 3 are presented in Appendix 6.2. Skills equating is a negotiated subjective task which is made more objective by employing more subject matter specialist to individually judge the content and comparing their outcomes to reconcile the discrepancies. The reconciliation process has no hard rules but is premised on mutual discussion and agreement.

Level of demand	Skill				
	1. Preparing a plot (12)	2. Using tools (7)	3. Planting (9)	4. Returning tools & materials to s/room (4)	5. Recording transactions (10)
1					
2					
3					

Figure 6.1: Skills equating for task 1

6.3.3 Task Development

The development of the first version of the prototype was undertaken by the researcher guided by findings in Chapter 5, together with the guidelines discussed in Section 6.2 The two main assessment objectives (Objective 2, Assessing the handling and application of information and problem solving skills; and objective 3, Assessing practical and investigative skills (discussed in Section 2.8) guided decisions related to tasks content and design activities.

The task of designing and developing quality intervention began with scrutinizing terminal objectives of the assessment syllabus (Ministry of Education, 2000b). Terminal objectives described exactly what the students should be able to do in a created learning context, not the real world (Dick et al., 2009). Once terminal objectives were comprehended, subordinate skills were derived, these being building blocks to be mastered by students towards achieving a terminal objective (Dick et al., 2009). Detailed tasks were then developed based on the subordinate skills identified.

The discussion of task development was based on Task 2: *Applying fertiliser as basal dressing*. The task development was divided into two parts, namely: (i) The task, and (ii) Assessment Instrument. The task comprised three aspects: The overall task (Figure 6.2); Pictorial presentation of skills (Figure 6.3); and Skills with their performance criteria (Table 6.1). The assessment instrument also comprised three aspects: Scoring instrument - Checklist and Scale (Tables 6.2 and 6.3); Summary marksheets (Table 6.4); and a detailed description of the assessment criteria (Table 6.5). The objective of conducting the practical was delineated, which was *to apply fertiliser as basal dressing.* The task was then stated generally, as presented in Figure 6.2 (below).

Figure 6.2 (below) states only the task, but does not give methodical details of the steps involved in executing the task. This presents implementation problems and is apt to be interpreted differently, resulting in various schools executing tasks of non-equivalent demands. Detailing subordinate skills pictorially in the form of hierarchical analysis shown in Figure 6.3 was a necessary intermediate step in refining the task.

Given a plot, use appropriate tools from the storeroom to apply a basal dressing fertiliser to your crops.

The task will be complete when you have:

i. applied the correct fertiliser;

ii. applied the right quantity of fertiliser;

iii. used the right tools to measure the quantity of fertiliser;

iv. applied/observed safety to self, crops and others during fertiliser application to crops;

v. returned the tools and other materials to the storeroom; and

vi. recorded all activities carried out.

Your performance on each step will be judged using the following general criteria:

i. performing each step;

ii. executing each step using the appropriate tools in the proper manner;

iii. observing safety to self, crops and others all the time;

iv. keeping detailed records of the activities carried out.

Figure 6.2: The overall task showing each step and general criteria

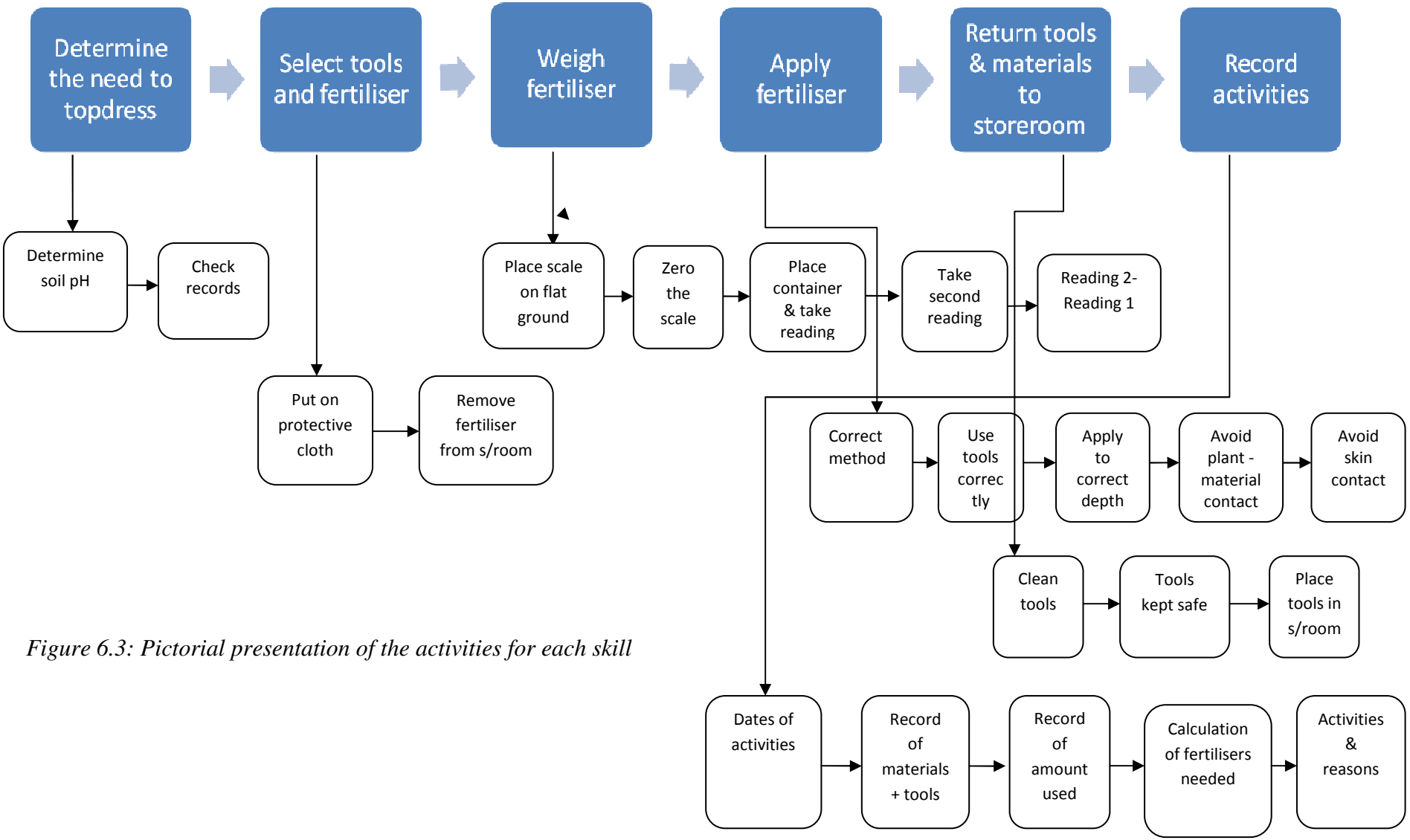

Figure 6.3: Pictorial presentation of the activities for each skill

The subordinate skills and their analysis however were brief for practical purposes. There was a need for increased specificity in the conditions and criteria for performance, as well as prescription of special circumstances (Table 6.1, below), which resulted in performance objectives that guided students as to the precise behaviour expected of them. These included the condition for performance identified using letters CN, the behaviours expected of students identified using letter B, and the performance criteria indicated by letters CR. The condition in this context was the description of the environment, tools and resources that would be available to the learner when performing the skill. The behaviour was the description of the skill that would include actions, content, and concepts, while the criteria were descriptions of acceptable performance of the skill (Dick et al., 2009).

Table 6.1: *Performance skills and matching performance objectives*

Skill	Performance Criteria	Marks
1. Determining the fertiliser requirements (2)	Given plot (CN), determine the need to basal dress (B).	
	a. Determine the soil pH,	1
	b. Find out what and when fertilisers need to be applied (CR).	1

Skill	**Performance Criteria**	**Marks**
2. Selecting tools and fertiliser(s) (3)	Given tools and fertilisers stacked in a storeroom (CN), select tools needed and fertiliser for application (B).	
	a. Identify tools and fertilisers needed for application	1
	b. Put on protective clothing,	1
	c. Remove the fertiliser from the storeroom to place of weighing (CR).	1

Skill	Performance Criteria	Marks
3. Weighing the fertiliser (5)	With the fertiliser to be applied and the tools needed ready (CN), weigh the fertiliser (B).	
	a. Zero the scale,	1
	b. Place an empty container on the scale and take reading 1,	1
	c. Place the required amount of fertilisers in the container and take reading 2,	1
	d. Subtracting reading 1 from reading 2,	1
	e. Work cooperatively with others (CR).	1

Skill	Performance Criteria	Marks
4. Applying fertiliser (5)	Given the crops growing in a plot (CN), apply the correct amount of fertiliser as basal dressing (B).	
	a. Use the correct method of fertiliser application,	1
	b. Use tools correctly,	1
	c. Apply fertiliser to the correct depth,	1
	d. Avoid fertiliser-planting material contact,	1
	e. Avoid skin contact (CR).	1

Skill	Performance Criteria	Marks
5. Returning tools and materials to storeroom (4)	After applying the fertiliser (CN), return tools and materials to the storeroom (B).	
	a. Clean all tools,	1
	b. Carry tools and materials safely to the storeroom,	1
	c. Place tools and materials in the storeroom neatly and in their correct place,	1
	d. Work diligently with minimal supervision (CR).	1

Skill	Performance Criteria	Marks
6. Recording transactions (10)	As you carry out the activities leading to application of fertilisers (CN), record all the transactions carried out and keep a tidy record (B).	
	a. Dates of activities,	1
	b. Materials and tools used,	3
	c. Calculations of amount of fertilisers needed.	3
	d. Activities carried out and their reasons (CR).	3

Numbers in brackets () under the heading Skill are marks for that particular skill

NB: CN = Condition; B = Behaviour; CR = Criteria

Up to this point, the task was considered fully developed. Efforts were now redirected towards developing the accompanying criteria for assessing the task. In developing the most appropriate instrument to evaluate the students learning, a number of factors were taken into account, including (1) the nature and complexity of the elements observed, (2) the time available for: observation, making judgement, and recording judgment, (3) the accuracy or consistency with which the evaluator can make the judgment, and (4) the quality of feedback to be provided to the learners (Airasian & Russell, 2008).

Given conditions prevailing in the assessment arena of Agriculture in Botswana schools (See Section5.4), one would be obliged to develop an instrument that would be easy to use yet tapping relevant information about the students learning, and resulting in valid and reliable inferences made from assessment scores. A checklist and a rating scale were thus developed (See Tables 6.2 and 6.3). A checklist was used for holistic assessment while a rating scale was used for analytic evaluation of subcomponents of a performance or product (Airasian & Russell, 2008). However, the checklist did not provide enough information for feedback to the students (Dick et al., 2009), and rating scales yielded less reliable scores than checklists (Colton & Covert, 2009). Up to four levels for the rating scales were included for ease of differentiating students' performances. Because of ease of scoring with four levels this helps to improve the reliability of scores.

Table 6.2: *Scoring instrument (Checklist) for the task*

Instructions to the teacher

Score the students using this checklist for skills 1-5 listed in Table 6.1. The scoring is based on 'yes' which represents criteria achieved or 'no' which represents criteria not achieved. The total mark for each criterion is one. Comment on why the student did not achieve the criteria.

Student Name: _____ Score: _____

Skill	**Criteria**	**Scoring Rubric**		**Comment**
		Yes	No	
1. Determining the fertiliser requirements (2)	a. Determine the soil pH			
	b. Find out what and when fertilisers need to be applied			
2. Selecting tools and fertilisers (3)	a. Identifying tools needed for application			
	b. Putting on protective clothing			
	c. Removing the fertiliser from storeroom to place of weighing			
3. Weighing the fertiliser (6)	a. Zeroing the scale			
	b. Placing the container on scale and taking reading 1			
	c. Placing fertilizer on container and take reading 2			

	d. Subtracting reading 1 from reading 2			
	e. Working cooperatively with others			
	a. Using correct method of fertiliser application			
4. Applying Fertiliser (5)	b. Using tools correctly			
	c. Applying fertiliser to the correct depth			
	d. Avoiding fertiliser-planting material contact			
	e. Avoiding skin contact			
	f. Cleaning (if need be) all tools			
5. Returning tools and materials to storeroom (4)	a. Carrying tools and materials safely to the storeroom.			
	b. Placing tools and materials properly in the storeroom.			
	c. Working diligently with minimal supervision			

Teacher's Name_____ Teacher's Signature _____ Date _____

Snr Teacher's Name _____ Snr Teacher's Signature _____ Date _____

Table 6.3: *Scoring instrument (Scale) for the task*

Instructions to the teacher

Score the students using this scale for Skill 6 listed in Table 6.1. The scores range from 0 to 3. Put the student score in the column 'Mark'. Comment on why the student did not achieve the criteria.

Student Name: _____ **Score:** _____

Skill	Criteria	Scoring Rubric				Mark	Comment
6. Recording transactions (10)		0 mark	1 mark	2 mark	3 mark		
	a. Record the date of activities	0-90% recorded	>90% recorded				
	b. Record materials and tools used in each activity	0-50% recorded	50-70% recorded	70-90% recorded	>90% recorded		
	c. Record of calculations of amount of fertilisers needed	All calculations wrong	30-70 % calculations correct	70 -90% calculations correct	>90% correct calculations correct		
	d. Record of activities and reasons	0-50% recorded	50-70% recorded	70-90% recorded	>90% recorded		

NB: A black box means no mark is allocated

Teacher's Name_____ Teacher's Signature _____ Date _____

Snr Teacher's Name _____ Snr Teacher's Signature _____ Date _____

However, the above criteria would be very cumbersome for teachers to carry to the field (practical site), because they would have to carry 35 copies (average class size) of each of the two instruments (Tables 6.2 and 6.3 for checklist and rating scale respectively), and frequently paging through to locate the student under observation to award marks. To avoid this problem and simultaneously seeking to conduct quality scoring, a summary marksheet is used instead. Using the summary marksheet, the teacher would need to carry:

(i) one copy of the class list populated for each student, with letters corresponding to the criteria for each subordinate skill (See Table 6.4).

(ii) one copy of each of the marking criteria for reference (Tables 6.2 and 6.3).

(iii) a copy of the detailed description of marking criteria (Table 6.5).

Detailing the marking criteria helps teachers interpret the criteria in a similar way, for example, if one does not have a clear understanding of what is meant by *Use correct method of applying fertiliser (4(a)), one* can find detailed examples of methods of application of fertiliser in the detailed description (Table 6.5).

.

Table 6.4*: Example of summary marksheet for use by teacher*s

Instructions to the teacher

Use this marksheet in the field to assess the students. This should be used in conjunction with Table 6.5 which gives a detailed description of each criterion. The letters (e.g. a, b, etc) represent criteria corresponding to each skill as detailed in Table 6.4. Circle the letter when the student meets the criterion corresponding to that criterion. Table 6.6 shows how this marksheet should be completed. **Total marks = 29**

Student Name	**Skill**						**Marks**	**Comment**
	1. Determining the fertiliser requirements (2)	**2. Selecting tools and fertiliser (3)**	**3. Weighing the fertiliser (5)**	**4. Applying fertiliser (5)**	**5. Returning tools & materials to s/room (4)**	**6. Recording transactions (10)**		
Lorato Menwana	a b	a b c	a b c d e	a b c d e	a b c d	a 0 1 b 0 1 2 3 c 0 1 2 3 d 0 1 2 3		
Lingani Mesho	a b	a b c	a b c d e	a b c d e	a b c d	a 0 1 b 0 1 2 3 c 0 1 2 3 d 0 1 2 3		
Tate Mgadla	a b	a b c	a b c d e	a b c d e	a b c d	a 0 1 b 0 1 2 3 c 0 1 2 3 d 0 1 2 3		
Ernest Forbes	a b	a b c	a b c d e	a b c d e	a b c d	a 0 1 b 0 1 2 3 c 0 1 2 3 d 0 1 2 3		
Nonny Meshack	a b	a b c	a b c d e	a b c d e	a b c d	a 0 1 b 0 1 2 3 c 0 1 2 3 d 0 1 2 3		
Bonang Ketshabile	a b	a b c	a b c d e	a b c d e	a b c d	a 0 1 b 0 1 2 3 c 0 1 2 3 d 0 1 2 3		

Teacher's Name_____ Teacher's Signature _____ Date _____

Snr Teacher's Name _____ Snr Teacher's Signature _____ Date _____

Table 6.5*: Detailed description of marking criteria for use during field evaluation*

Skill	Criteria	Mark
1. Determining the fertiliser requirements (2)	• Determine soil pH: different crops thrive best in different soil pH. Find out the soil requirements for the crop you are growing, particularly its pH. The teacher should not tell students the type of soil and pH needed. Students should take a leading role in their learning. This criterion is marked from the students' records.	1
	• Find out what and when fertilisers need to be applied: The student should check from literature which fertilisers and how much is used as basal dressing for the variety of the crop planted. Note that fertiliser needs for the same variety of crop may differ from one region to the other depending on influence of the climate. This criterion is marked from students' records.	1
2. Selecting tools and fertilisers (3)	a. Identifying tools and fertiliser needed for application: tools that would be needed are such as scale, hand trowel, while fertilisers are such as N:P:K 2:3:2 (34), superphosphate, kraal manure, potassium sulphate, wood ash (K), basic slag (P), and so on. There could be as many or as few as possible, depending on your location. Prescription of how many tools and/or materials that warrant a point is also dependent on your situation. It is left to the professional judgement of the teacher to determine how many marks the student deserves for the tools and materials enumerated.	1
	b. Putting on protective clothing: Schools should make concerted effort to acquire protective clothing for students and teachers, such as overall, boots, mask, goggles, and respirator. It is a legal requirement that chemicals should be applied putting on protective clothing. This is scored on all or nothing basis.	1
	c. Removing the fertiliser from storeroom to place of weighing: this is a collaborative activity. The teacher should look at what students do when loading and off-loading; whether they are working together as a team. Team building is an important aspect for productivity. Aspects of attitudes are also encompassed.	1
3. Weighing the fertiliser (5) Group work is expected	a. Zeroing the scale: whether digital or .analogue scales should be zeroed to get precise reading. The scale should be at 0.00 before putting anything on it. Putting the scale on a flat area will facilitate achievement of such.	1
	b. Placing an empty container on the scale and taking its reading 1: (see worked example)	1
	c. Placing the desired amount of fertiliser in the container and taking reading 2 (see worked example)	1
	d. Subtracting reading 1 from reading 2. (see worked example in teacher's guide)	1
	e. Working cooperatively with others: cooperation is highly encouraged in Agriculture. Almost all agricultural activities require group work. It should be inculcated into students that helping one another is an important attribute for success. Competition should be discouraged by all means.	1
4. Applying Fertiliser (5)	a. Using correct method of applying fertiliser: methods such as broadcasting, drill, banding, foliage, etc. Students should justify the choice of their method.	1
	b. Using tools correctly: tools last longer if used for the right purpose. Students should use tools for their rightful purposes. However, this does not relegate improvisation whenever necessary.	1
	c. Applying fertiliser to the correct depth: Fertilisers should be applied at the correct depth to be used by plants. If applied too deep, it will leach resulting in stunted growth by plants. If shallowly applied, it will volatilise and escape into the air.	1

Skill	Criteria	Mark
	d. Avoiding fertiliser-planting contact: Chemicals and fertilisers burn crops. The fertiliser should be placed deeper than the planning material or away from the row of planting material.	1
	e. Avoiding skin contact: fertilisers are chemicals and have residual effect, as such, contact with the skin should be avoided.	1
5. Returning tools and materials to storeroom (4)	a. Clean all tools: Tools should be cleaned before they are taken to the storeroom. Cleaning does not imply only using water. Greasing, polishing or removing soil may constitute cleaning. Cleaning is any action that prevents tools/equipments from rusting.	1
	b. Carrying tools and materials safely to the storeroom: tools should be carried safely to the storeroom to prevent injuries. The proper way to carry tools is pointing the sharp end downwards. Tools left lying around are exposed to harsh weather conditions and are likely to wear quickly.	1
	c. Placing tools and materials correctly in the storeroom: tools should be placed on tool racks. It is advisable that such be made in case they are not there, otherwise it would be difficult to assess this skill.	1
	d. Working diligently with minimal supervision: Students should not work only when the teacher is around. They should take a leading role in their learning. Once students have been cultured into doing right things all the time, they work on their own with minimal supervision. Tools won't be left lying around, they will always submit their record books at the right time for scoring, and willing to share tools, and so on.	1
6. Recording transactions (10)	a. Recording the date of activities: This is marked out of 1 because it does not require any skill to do yet it is important.	1
	b. Recording materials and tools used in each activity: The materials and tools used in each activity will vary in terms of number, type and extent of usage from school to school, activity to activity, and student to student. The professional judgement of the teacher is called for in this particular case.	3
	c. Recording calculations of amount of fertilisers needed: Before going to weigh the fertiliser to apply to crops, the student has to calculate how much fertiliser is needed. All the calculations have to be shown for the student to score maximum points.	3
	d. Record activities and reasons for carrying out the activities: This will depend on individual schools, classes, and students. It is difficult to state categorically the number of reasons that warrant 1 or 2 marks. The teacher is better placed to know. Professional judgement should again be exercised.	3
TOTAL		**29**

NB: The total marks are indicated beside each skill in brackets () in the first column and marks for each criterion in the last column.

Given the summary marksheet presented in Table 6.4 and the detailed description of marking criteria presented Table 6.5, the teacher can assess a number of students with ease. Let us consider how the summary marksheet is used. Taking Lorato Menwana as an example, under the activity *Determining the fertiliser requirements*, there are two criteria labelled *a* and *b*. From Tables 6.1 and 6.2, it was evident that:

criteria *a* is: determining soil pH

and criteria *b* is: finding out what and when fertilisers need to be applied.

The teacher circles *a* and *b* if the student successfully achieves these criteria. Table 6.6 (below) shows how the scoring is done. Other skills are scored the same. The total mark for *Lorato Menwana* was obtained by aggregating the number of letters and/or numbers circled, totalling 22 out of 29, and captured under the column 'total'. Likewise, the mark for *Lingani Mesho* was computed in a similar way (23).

This kind of assessment infused flexibility and the teacher could assess different students at the same time on the same activity or skill, or assess different activities. Practically, the teacher could target a few students per lesson, perhaps 10, who could be quickly and accurately assessed. The rest of the class can be assessed on other skills of similar demands identified during skills equating. Any criteria not achieved by the student can be reassessed another time or during the conducting of a similar task. After assessing, the teacher transferred the marks from the summary marksheet to the scoring instruments shown in Tables 6.2 and 6.3 (above) for purposes of traceable and retrievable records for accountability. Supervisors and inspectors can peruse these records and use them for reconciliation whenever there is a dispute.

The development of the other two tasks followed a similar procedure, and is presented as appendices 6.3 and 6.4.

Table 6.6: Sample of completed summary marksheet

Total: 29 marks

Student Name	Subordinate skill							Comments
	1. Determining the fertiliser requirements (2)	2. Selecting tools and fertiliser (3)	3. Weighing the fertiliser (5)	4. Applying fertiliser (5)	5. Returning tools & materials to s/room (4)	6. Recording transactions (10)	Total	
Lorato Menwana	a b	a b c	a b c d e	a b c d e	a b c d	a 0 1 b 0 1 2 3 c 0 1 2 3 d 0 1 2 3	22	
Ernest Forbes	a b	a b c	a b c d e	a b c d e	a b c	a 0 1 b 0 1 2 3 c 0 1 2 3 d 0 1 2 3		
Lingani Mesho	a b	a b c	a b c d e	a b d e	a b c	a 0 1 b 0 1 2 3 c 0 1 2 3 d 0 1 2 3	23	
Nonny Meshack	a b	a b c	a b c d e	a b c d e	a b c	a 0 1 b 0 1 2 3 c 0 1 2 3 d 0 1 2 3		

Teacher's Name_____ Teacher's Signature _____ Date _____

Snr Teacher's Name _____ Snr Teacher's Signature _____ Date _____

6.4 FORMATIVE EVALUATION OF THE FIRST PROTOTYPE BY EXPERT GROUP

The preliminary validity of the prototype was to be ascertained through evaluation and feedback could be incorporated into the redesign of the second prototype.

6.4.1 Research Design

The evaluation of the first prototype of the standard task and assessment materials was carried out by experts who were given the three tasks to review against the criteria. The intention of evaluation at this stage of development was to maximise the content validity (De Villis, 2003) as well as the consistency between tasks and assessment criteria (Plomp & Nieveen, 2007). Feedback from evaluation was incorporated into the development process to improve the effectiveness of the intervention (McDavid & Hawthorn, 2006). The review was guided by the evaluation question:

Is there consistency between the tasks and assessment criteria?

6.4.2 Participants

Formative evaluation of the first prototype involved two groups of evaluators, the first of which comprised three Assessment Officers, two Agriculture Education Officers from CD&E and DSE, and a lecturer from the College of Agriculture offering Measurement courses to student-teachers. Officers also had wide experience of the school system as they had been teachers before. The second group of participants comprised five teachers from two senior secondary schools who had vast experience in teaching Agriculture. Demographic information relating to the experts is given in Appendix 6.5.

6.4.3 Data collection strategies

Expert Evaluators completed 4-point Likert scales ranging from 4 representing strong endorsement, to 1 representing weak endorsement, as shown in Appendix 6.6. The first scale sought to find out the sufficiency of the format and the clarity of the language. The next two scales were aimed at determining the adequacy of the task and adequacy of assessment criteria. The scale sought to find out the clarity of instructions and the instrument also provided experts with the opportunity to express their views by answering open-ended questions. Quantitative data was analysed descriptively through determining the reliability of

the instrument. Data was analysed qualitatively using themes and thick descriptions quoted verbatim.

6.5 EXPERTS' VIEWS AND EXPERIENCES WITH THE FIRST PROTOTYPE

Experts were requested to evaluate the *Task* and *Assessment Instrument* for (i) sufficiency of format, (ii) clarity of language, (iii) adequacy, and (iv) clarity of instructions, on scales ranging from 1 to 4. Negatively worded items were reversed during analysis so that the high numbers denoted high endorsement. Table 6.7 presents the number of items and reliability coefficients for each scale. Detailed table is presented in Appendix 6.7.

Table 6.7: *Reliability coefficients for scales of the tasks and assessment instruments*

Scale name	**No of items**	**Reliability coefficient of task**		
		1	2	3
Sufficiency of format of task	3	.82	.64	.91
Sufficiency of format assessment instrument	5	.89	.94	.90
Clarity of language for task	3	.93	.82	.92
Clarity of language for assessment instrument	5	.92	.96	.88
Adequacy of task	10	.83*	.95	.87
Adequacy of Assessment Instrument	7	.77	.87	.90
Clarity of instructions	4	.90	.88	.87

** when 1 item is removed*

Task 1: Preparing a plot and planting

The reliabilities of the sub-scales for task 1 were found to be generally high as shown in Table 6.7. The Generally, there was high degree of agreement among experts concerning the sufficiency of *format of the* task and assessment instrument, adequacy of the task and assessment instrument for Form Four level, and clarity of the language used. However, the

task was found to be physically inaccessible and lacking activities targeting students' of different abilities:

> *Assessment task did not cater for those who are disabled or visually impaired even though they are mixed in the classes with normal students. Their abilities are not assessed because most of the work is done for them by other students or assistant teachers.*

Experts also suggested separating this task into two tasks since *Plot preparation* and *Planting* were adequate enough to stand on their own. Suggestions were made to improve clarity of the instructions. Experts' views are summarised in Table 6.8.

Table 6.8*: Summary of experts views on task 1*

criterion	**Experts views**
Overall quality of task	The task is of good quality and achievable. The assessment instrument is objective and allows the one assessing to spread out the marks across a wide spectrum of performance by the learners. It can be useful for narrowing the gap in mark allocation by different examiners in different places since the instruments clearly elaborate on what must be done in order to award each mark. Land preparation and planting skills be separated. Marking criteria should be differentiated to show that the tasks cater for learners of different abilities.
Content of task	Covers much of the content of the syllabus There is no description of how content of the task relates to the curriculum objectives. In some areas demanding tasks given low marks and level of difficulty should correspond with mark allocated.
Format of task	Well formatted and structured to measure the indented learning outcome. The structure is technically sound and of technical quality. The structure should be such that it shows the relationship between the task and assessment. The task is not differentiated to accommodate the different ability levels. Needs improvement, it is too congested and shorten sentences

Language of task	Simple to be understood by learners at the proposed level. Language is ok but instructions not quite clear. Guidance not quite clear

Task 2: Applying fertiliser as basal dressing

All the sub-scales were found to be internally consistent with generally high reliabilities as shown in Table 6.7 (above). Reliability for *Adequacy of assessment Instrument* was .87 after one item was removed. Experts unanimously agreed that the standard task and assessment materials were well constructed to elicit the desired outcomes.

Experts agreed that the task and assessment instrument were sufficiently formatted and the language clear enough to be understood by the intended users: "*The task is well framed and phrased for one to easily follow and understand and I will suggest you retain them in this format"*. Likewise, experts agreed that the task and the assessment instrument were adequate to be used for Form Four level to measure students' capabilities more efficiently than before, with the exception of not being physically accessible to all students. Experts felt that the task was not providing opportunities for all students to interact and cooperate within a group. . The only concern raised was that it seemed inadequate in assessing students' cognitive skills. Table 6.8 summarises experts' views for task 2.

Table 6.9: *Summary of experts views on task 2*

criterion	Experts views
Overall quality of task	Clearly, well defined, explicit, good, elaborated and clearly outlines the actual activities that occur during fertiliser application for easy understanding by both teachers and students. Comprehensive and user-friendly to execute. I haven't been able to come up with a document/planned assessment instrument to use, now light has been shed.… just wish the instrument could be used on a small sample of students to see if it is user friendly. Instructions to the tasks not clear particularly for the pictorial presentation.
Content of task	Good, self-explanatory, well understood, fig 6.4 skill 6 criteria (b) and (c) will require more details, covers all content areas for the level of intended users. Content does not show any differentiation among the students' of different abilities. Activities need to be reduced.
Format of task	Well structured, detailed, clear and easy, too congested & a lot of wording. The structure does not present itself to allow interaction among learners.
Language of task	Comprehensive, and clear but shorten writing, easy to understand and suitable to both students and teachers.

Task 3: Controlling weeds using chemicals

The reliabilities of the sub-scales for task 3 were found to be high as presented in Table 6.7 (above). Experts expressed satisfaction with the format, language and adequacy of task 3. However, suggestions were made to improve the structure so that it would differentiate between students of different abilities, as well on its feasibility given the nature of resources involved in this particular task. *The task on weed control using chemicals is well described but I am in doubt of its feasibility in a school set-up particularly taking into account the residual effect of herbicides in a garden with multiple users.* Table 6.10 summarises experts' views on task 3.

Table 6.10: *Summary of experts views on task 3*

criterion	Experts views
Overall quality of task	Hope this exercise will be given the seriousness it deserves because at the moment or for a very long time agric practicals haven't been well assessed. Opinions of different bosses have been used as assessment tools. Well described, but its feasibility in a school is doubtful taking into account the residual effect of chemicals. Task composition is mainly suitable to science based students with better knowledge of mathematics. Need to have a set standard of tasks for each syllabus topic and common for all schools
Content of task	Covers content of syllabus and relevant. Covers all activities in the task. If there are no weeds, shouldn't the learner choose an appropriate method s/he would like to control the weeds?
Format of task	Well structured and easy to follow. Measure what is intended to measure. I have learnt the detailed step by step approach for assessing tasks and the benefit of using a well prepared scoring instrument to assess and grade students. Improve the format
Language of task	Clear and simple

6.6 CONCLUSION

There was high endorsement on the need to develop the assessment intervention to improve the assessment of practicals and consequently enhance its contribution for certification. Expert welcomed the development of the standard task and assessment materials as the solution to some of the problems that had bedevilled practical assessment for a long time.

Almost all experts expressed lack of understanding on the use of the marksheets, calling for a thorough training of teachers before they could be given the assessment instruments to use. The tasks were found to lack i) physical accessibility to all students; ii) demonstration by students of understanding in a variety of ways; iii) activities for groups of differing abilities;

iv) opportunities for all students to interact and cooperate within a group set up, and (v) assessment of affective skills. Suggestions were made on how to include these.

Other suggestions made to be factored in the design of the second prototype included improving clarity of instructions, and inclusion of date of task execution as authentication, as well as teacher's and senior teacher's signatures as a quality assurance step, justification for high marks for record keeping as they lacked face validity.

6.7 IMPLICATIONS FOR FURTHER DEVELOPMENT

Findings of the experts' evaluations resulted in tasks being revised to make them open and more interactive, both between students and between students and teachers. Interaction helps students to learn from multiple sources in different ways, as well as use information to evaluate themselves and others.

Since experts indicated that tasks did not cater for special needs students, expertise was sought from the Botswana Examinations Council's Special Education Officers on how to craft tasks and assessment instruments so as to cater for special needs students. Tasks were refined to cater for different abilities and differentials in intellectual development. In addition, tasks were made to cover in-depth challenging content of knowledge and skills (Lane & Stone, 2006) to provide students with the opportunity to use multiple skills and abilities (Diez, 2002; Rennert-Ariev, 2005; Ryan, 2006).

Instructions were revised so as to be easily understood by all users, a crucial move since assessment materials would be given to students prior to teacher assessment for familiarisation and self-evaluation. A number of affective objectives would be introduced as they are an integral part of the curriculum, and are no longer considered part of the 'hidden curriculum' as was the case during the 1980s (Jarolimek, 1981). However, they would be infused in skills assessment (Gronlund, 2003) for ease of assessment.

6.8 DESIGN OF THE SECOND PROTOTYPE - PILOT

The intention of design and evaluation at this stage was to explore the validity and practicality of the standard task and assessment materials in the context of Botswana Agriculture practical assessment with Form Four students. That is, to find out if teachers were able to use the standard task and assessment materials with their students as intended by the designer. The design and review were guided by the evaluation question:

What is the practicality of the intervention that aims at supporting performance assessment in agriculture?

The review of the first prototype was the first cycle of formative evaluation and highlighted a number of issues to be included in the design of the second prototype, as indicated below:

1. Incorporation of collaborative activities: initially, collaborative activities were implicitly infused in the task. These had to be made explicit to guide teachers and students precisely to the behaviour expected to improve the validity and reliability of the assessment.

2. Clarity of instructions: in a number of cases, instructions were improved for ease of understanding. An *introduction* and *Directions to teachers* were included to guide teachers and students to the interpretation of the intervention. The *Detailed description of the marking criteria* section of the standard task and assessment materials was elaborated to clarify how each objective could be achieved and scored. The *Implementation plan* was also introduced, detailing how each skill should be implemented.

3. Inclusion of tasks catering for different abilities: the design of the prototype included a number of critical thinking, abstract, problem-solving, and reasoning skills to cater for students with different abilities. These enabled teachers to differentiate between students who needed help and those who needed more challenging work.

4. Inclusion of affective skills: a number of affective skills in the development of the second prototype were included, for instance: (a) Working diligently with minimal

supervision; (b) Working cooperatively with others; and (c) Observing safety to self and others.

5. Marks allocation: experts had raised questions on the allocation of marks, for example why more marks were allocated to record keeping. Efforts were made to explain in the *Detailed description of the marking criteria* why this was the case.

6. Accommodation for special needs students: tasks were made flexible so that they could be easily modified for special needs students.

7. Resources needed: for ease of preparing for the lesson, materials, tools and other resources needed for each task were outlined. This would enable schools to acquire them in advance for effective implementation of tasks.

8. Induction of teachers: well-thought out notes were prepared for use during workshops with teachers on how to implement the intervention. This was particularly important to facilitate standard implementation of the intervention throughout the country.

6.9 FORMATIVE EVALUATION OF THE SECOND PROTOTYPE

This section discusses the research design employed in evaluating the intervention, the participants involved, and data collection strategies.

6.9.1 Research design

To ascertain consistency or logical design of the standard tasks and assessment materials, three teachers and their students from one school were involved in piloting the standardised materials. Purposive sampling technique was employed for both the teachers and students, having agreed to voluntary participate through a signed consent form. Each teacher piloted one of the three tasks and assessment instruments.

6.9.2 Participants

Three teachers and their Form Four students from one Government-Aided school participated in the study. There were four teachers in total, three of whom volunteered to participate. Demographic information about the participants is presented in Table 6.11 (below). All teachers were male, with adequate experience of teaching and held Senior Teacher's Grade II position. They possessed at least a bachelor's degree qualification, and their class sizes were large, ranging from 42-45.

Table 6.11: *Demographic information of participants*

Variable	Teacher 2	Teacher 1	Teacher 3
Age	38	34	40
Sex	Male	Male	Male
Academic qualification	Degree	Degree	Degree
Professional qualification	BSc. Agric Ed	BSc. Agric Ed	BSc.Agric Ed
No of year teaching	15	7	14
Post	Ag. ST Grade I	ST Grade II	ST Grade II
Class size	45	45	42
Sampled students	10	5	5

A total of 20 Form Four students participated in the study, ranging in age from 16-18 years. As indicated above, Form Four students were used because the final students (Form Five) had already completed all their practicals. During the time of conducting the study, the Form Fives were preparing a report for the final project (see Sections 1.3 and 2.8) to be ready for scoring and moderation in September. Form Fours were considered to have enough experience in conducting practicals since they were did them at junior level where Agriculture was offered as a core subject (See Section 2.5.3).

6.9.3 Data collection strategies

(i) *Procedure*

Teachers were taken through the standard task and assessment materials step-by-step prior to implementation, during the workshop conducted in the afternoon when they had no lessons. However, they had difficulty in conceptualising the implementation strategy, particularly the

Summary marksheet, so further efforts were made to explain, reinforced by hands-on facilitation for better conceptualisation. Teachers were subsequently requested to explain to their students how the materials were used. Teacher 1 selected task 3: *Controlling weeds using chemicals*, teacher 2 selected task 1: *Preparing a plot and planting*, while teacher 3 chose task 2: *Applying fertilisers as basal dressing,* as shown in Table 6.12.

Table 6.12: *The tasks selected by teachers*

Teacher	**Task**	**Task Name**
1	3	Controlling weeds using chemicals
2	1	Preparing a plot and planting
3	2	Applying fertilisers as top dressing

Materials were delivered to the teachers four days before the planned day of implementation for distribution to students before the start of the lesson. The conduct of the practicals was made during the last week of the term after end-of-term examinations. Though teachers were busy marking examinations, they managed to accommodate the piloting because they understood its likely implications for improving students' learning. During the day of piloting I went early to the school to assist with logistical issues to ensure proper implementation. All the three practical lessons were conducted the same day during normal teaching time.

The researcher accompanied teachers to the garden, where the practicals took place and the teachers introduced me to the students so that I did not appear as a stranger. They gave a brief explanation of the objective of the practical before the students could work on their own. Teacher 1 selected ten students, while teacher 2 and 3 selected five students each for assessment. Task 1 was relatively easy to assess, hence more students were selected. Teachers interacted with the students as they were assessing their processes. Meanwhile, the researcher completed an observation schedule, and at the end of the lesson, teachers completed a self-administered questionnaire to reflect on how they perceived the practicality of the standard task and assessment materials. The researcher requested six students in all from the three classes to be interviewed in the afternoon as a focus group. After the students'

interviews, one teacher was interviewed and the other two were interviewed the following day. All the interviews were audio-taped.

(ii) Data collecting instruments

Teacher evaluation questionnaire and interview

At the end of the lesson, each of the three teachers completed a questionnaire which sought their views on the implementation of the standard task and assessment materials, with both closed-ended and open-ended questions. The former targeted teachers' views on instructional behaviour, knowledge of assessment, standardising assessment and class management, while the latter sought the views of teachers on quality, content, format and language used on the standard task and assessment materials.

A structured interview was administered at the end of the lesson, the aim of which was to capture respondents' views about the impression of the intervention in their own words. Issues discussed ranged from the usefulness of the standard task and assessment materials, its feasibility, things they did not like, and how the tasks could be improved. The teacher evaluation questionnaire and interview schedule are presented as appendices 6.8 and 6.9 respectively.

Lesson observation

All lessons were observed and the researcher posed as a silent observer. An observation schedule was used to collect data of the activities of the lesson. The observation focussed on instructional behaviour of teachers, knowledge of assessment by the teacher, and resources availability. The instrument was developed in the form of a rubric to fully describe the activities of the lesson. The *Instructional behaviour* scale had five-level descriptors which holistically described the behaviour of the teacher. The teacher's knowledge and resources availability were evaluated through four-point and five-point analytic scales respectively. The instrument had provision for field notes to capture what transpired during the lesson. The observation schedule is presented as appendix 4.5.

Student interview

A focus group of six students was interviewed at the end of the practical lesson, using a semi-structured interview schedule, with probing and follow-up questions to get insight into

students' views about the standard tasks and assessment materials they had been implementing. The interview lasted for about one hour and transactions were audio-taped for later transcription. The interview schedule is presented as appendix 4.6. Students completed a self administered questionnaire comprising a Likert scale with 11 items and five open-ended questions. The questionnaire is presented as appendix 4.8.

6.10 RESULTS OF THE EVALUATIONOF THE SECOND PROTOTYPE

Results presented are based on lesson observation, standardising marking, students' understanding of assessment practices, completion of the assessment instrument, and record keeping. Lesson observation is divided into instructional behaviour, knowledge of assessment and resources availability.

6.10.1 Lesson Observation

Lesson observations were conducted with the view to understand what the teacher was doing during performance assessment. Lesson observations constituted instructional approach adopted by the teacher, teachers' knowledge of assessment, and resources availability.

Instructional approach

Table 6.13 shows the extent to which teachers' conducted activities to facilitate effective learning. Generally, teachers' instructional practice in assessment was average to above average indicating that their instructional approaches were student-centred. Figure 6.4 (below) shows the frequencies of instructional approaches diagrammatically.

At the beginning of the lesson, students had been given copies of the standard task and assessment materials just before the lesson. Teachers' introduction of the lesson's objective was not clear to the students as they did not understand what to be achieved at the end of the lesson. Teachers asked a number of questions to gauge the students' knowledge on the topic, but that was not related to everyday life. Only teacher T_3 made some effort.. Teachers interacted with the students as they were working and their assessment was individualised. Teachers generally agreed that the intervention was very useful for instructional effectiveness.

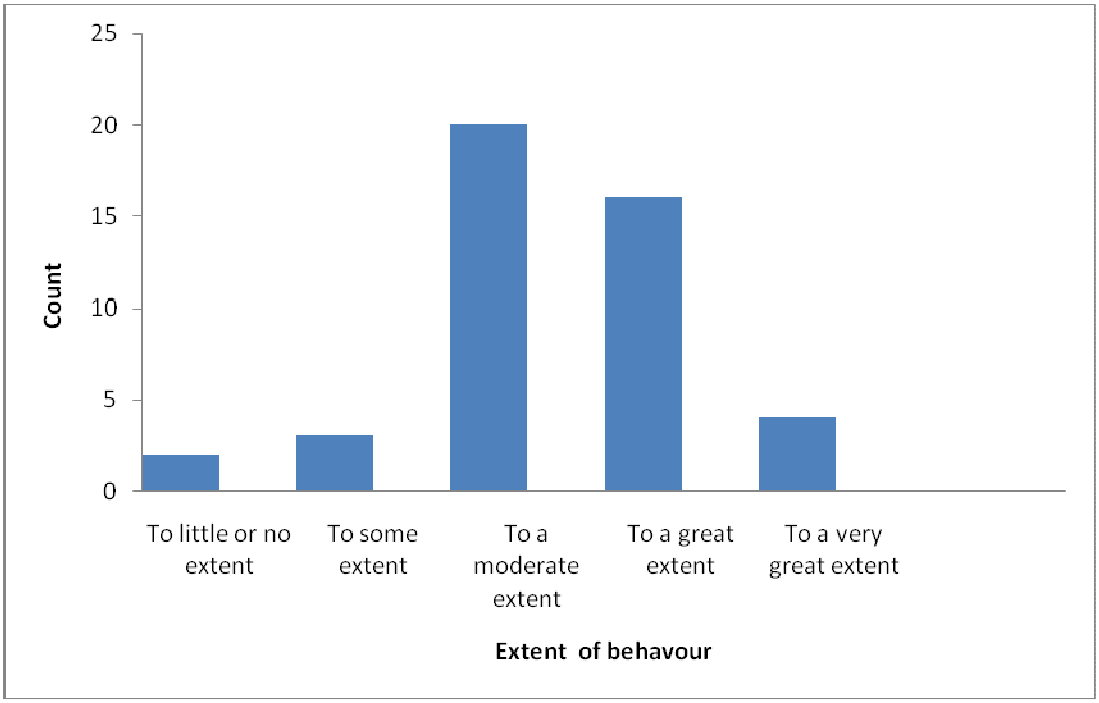

Figure 6.4: The *occurrence of the extent of instructional behaviour*

Table 6.13*: The extent of conducting activities by different teachers*

	Activity	To little or no extent	To some extent	To a moderate extent	To a great extent	To a very great extent
1.	Stating the performance assessment objective before the start of the practicals			T_3 T_1	T_2	
2.	Asking questions to gauge the students level of knowledge of the activity		T_2 T_3		T_1	
3.	Linking the relevance of the practical to everyday lives			T_2 T_1 T_3		
4.	Clarifying what resources are to be used and how they are to be used			T_1 T_3	T_2	
5.	Clarifying what will be assessed and how			T_1 T_3	T_2	
6.	Stressing observation of safety			T_2 T_1 T_3		
7.	Spelling out observable aspects of performance that should be judged			T_2	T_1 T_3	
8.	Distributing the task to the class before the start of the practicals			T_3		T_2 T_1
9.	Managing time well			T_2 T_3	T_1	
10.	Organising materials for the practical			T_3	T_2 T_1	
11.	Stating the students' role				T_2 T_1 T_3	
12.	Stating the students' task				T_2 T_3 T_1	
13.	Emphasising reasoning as opposed to rote learning	T_2 T_3	T_1			
14.	Providing an appropriate setting to elicit and judge the performance or product				T_3	T_2 T_1
15.	Providing a judgement or score to describe performance			T_2 T_1 T_3		

Teachers' knowledge of assessment

Teachers' knowledge of assessment practices is presented in Table 6.14. The least knowledge is represented by 1 while 4 represents adequate knowledge. Teachers' knowledge of assessment was modest. Although teachers discussed the roles of each party in assessment, modest action was taken to gauge students' readiness for assessment. Assessment was conducted without students' consent for readiness.

Table 6.14*: Knowledge of assessment displayed by individual teacher*

Activity	**1**	**2**	**3**	**4**
Providing opportunity for students to be assessed when ready		T_2 T_1	T_3	
Assessing processes				T_3 T_2 T_1
Forming groups during practicals			T_2	
Assessing individuals in group work			T_2	
Assessing all students in a class on the same skill in one day	T_3 T_2 T_1			
Perusing students' records		T_3 T_2 T_1		

1 = little knowledge, 5 = adequate knowledge

All teachers had insufficient knowledge on assessing specific skills instead of all, to an extent that teacher T_1 respondent that the intervention helped very little in assisting them to assess all students in a class at the same time. Knowledge on what exactly to look for when perusing students' records was inadequate. Students did not have proper books for record keeping, but made records on the standard task and assessment materials that they had been given. All teachers assessed students as they were working. As indicated above, only teacher 2 formed groups as more students were involved, but was however not grounded in individual assessment of students in group work.

Teachers appreciated the idea of students assessing themselves and assessment conducted by at least two teachers though they did not practice that. In addition to motivating students, it helped teachers to manage their classes better, particularly in taking care of tools and

implements. Teachers felt that tasks had content that was excessive for both the students and the teachers.

Resources availability

Resource availability is paramount to the successful implementation of the assessment programme. The provision of resources to schools to enable effective conduct of practicals was a challenge. Due to large class sizes, two to three students shared tools and in some cases up to five shared the same equipment, a situation that hindered effective connection between materials taught and students' experience in the field setting (Finn et al., 2003; Jones, 2006). The situation was serious when more than one class were involved in practicals simultaneously. There were no protective clothing provided for either students or teachers, despite the demand by curriculum for students to conduct practicals, some of which involved the use of chemicals. However, the provision of space in the garden was enough for students to have individual plots. Table 6.15 presents the extent of availability of resources in schools.

Table 6.15*: Availability of physical resources in schools to facilitate performance assessment*

Resources	**Availability**				
	1	**2**	**3**	**4**	**5**
Tools			✓		
Equipment	✓				
Other materials	✓				
Garden space					✓
workload		✓			
1= least available			5= most available		

6.10.2 Standardising marking

Interviews with teachers revealed high appreciation of the intervention to standardise performance assessment, because the guide they had been using was very subjective resulting in varied interpretations across schools. The assessment syllabus states the number of performance assessments to be made but it did not categorically dictate the level of difficulty. This culminated in schools administering tasks of different demands, format and frequency. The format of performance tasks administered ranged from products assessment, to

interviewing students about their conduct of practicals, to administering written practical tests.

6.10.3 Students' understanding of standardised assessment materials

The outcome of the students' questionnaire is presented in Table 6.16 (below). Three items which were negatively worded were reversed before analysis was done, while two items were removed which correlated very little with the item total. The resulting scale had high internal consistency of 0.80.

Generally, students understood and appreciated the standardised assessment materials as reflected by their high endorsement on most of the statements. However, concern was the 31.6% and 21.1% of students who did not understand the importance of *marking the processes*, and *marking of the products* respectively. This could imply that some students aimlessly conducted practicals, as a result of teachers failing to state the objectives explicitly as discussed above under *instructional approach*. Understanding an individual's role in group work was a problem to 33.3 % of the students.

Also of concern is about a quarter of students (26.3%) who had some difficulties linking the practicals to the theory. One student said: *"In practical, I can understand much rather than theory because one can see how it is done"*. Students liked the idea of being given the assessment materials before the commencement of assessment because it helped them know expectations in advance (Black & Wiliam, 1998; Harlen, 2006), practice first before teacher assessment, and helped them to reconcile the marks with the teacher after scoring: *the teacher marked you according to what you do ...you are given marks you deserve*".

Students like the idea of being scored by two or more than one teacher:

> *"one teacher is not good because s/he can make a mistake during the assessment unlike when they are two ... their scores are going to be different but at some point they would agree that ok these students deserve the mark they have given"*

Table 6.16: *Students' understanding of assessment practices*

The way I have been *doing practicals for the past week* makes me:	**SA**	**A**	**D**	**SD**
Understand the importance of practicals	26.3%	68.4%	5.3%	0.0%
Understand the link between practicals and theory	36.8%	36.8%	26.3%	0.0%
Enjoy doing practicals (reversed)	36.8%	52.6%	10.5%	0.0%
Like learning more about the topic	36.8%	52.6%	10.5%	0.0%
Feel encouraged to do practicals (reversed)	47.7%	47.7%	5.3%	0.0%
Understands my role in group work (reversed)	33.3%	33.3%	33.3%	0.0%
Realise the importance of working cooperatively with others	47.4%	42.1%	10.5%	0.0%
Be responsible in caring for tools	47.4%	47.4 %	0.0%	5.3 %
Understand the importance of safety in practicals	31.6%	63.2%	0. %	5.2%
Understand the importance of marking practicals while we are doing them (processes)	15.8%	52.6%	31.6%	0.0%
Understand the importance of marking practicals when we have finished doing them (products).	26.3%	52.6%	21.1%	0.0%

SA = Strongly Agree; A = Agree; D = Disagree; SD = Strongly Disagree

Students also suggested some improvements to be made, such as thorough explanation of assessment materials, and being given a wide choice of what they wanted to do rather than being required to do one thing.

6.10.4 Completion of the assessment instrument (Checklist)

The completion of the assessment instrument was a problem, with teachers not following the example given in the assessment guide despite prior training. The completion of the form was insufficient and unsystematic. Teachers selected skills to score, for example the form in Table 6.17 (below) shows that the teacher selected skills 1, 3 and 5 but the mark allocated to a certain skill was a total rather than for individual criterion, hence it was extremely difficult to know which criterion was not achieved. Furthermore, not all students were assessed on skill 3 – *calibrating the sprayer*, and no comments were made as to why the students did not achieve a certain criterion. The completion of assessment form revealed their lack of understanding of analytic scoring, resulting in insufficient and unsystematic scoring (Airasian & Russell,

2008). They were used to holistic scoring, as a consequence, the outcome could not be used as the basis for formulating remedial or enrichment strategies (Salvia & Ysseldyke, 1998).

Table 6.17: *An example of scoring by teachers*

Student Name	**Skill**						**Comments**
	Identifying weeds (4)	Organise materials (3)	Calibrating the sprayer (10)	Preparing and spraying chemical (7)	Returning tools & materials to storeroom (4)	Recording transactions (7)	
Lorato Menwana	a (i) (ii) 0 b (i) (ii) 0	a b c 1	a b c d 2 e 0 1 2 3 2 f 0 1 2 3	a b 0 1 2 3 c 0 1 2 3	a b c d 4	a 0 1 b 0 1 2 3 c 0 1 2 3	
Lingani Mesho	a (i) (ii) 0 b (i) (ii)	a b c 2	a b c d 1 e 0 1 2 3 3 f 0 1 2 3 3	a b 0 1 2 3 c 0 1 2 3	a b c d 4	a 0 1 b 0 1 2 3 c 0 1 2 3	
Tate Mgadla	a (i) (ii) 0 b (i) (ii) 1	a b c 0	a b c d 2 e 0 1 2 3 3 f 0 1 2 3 1	a b 0 1 2 3 c 0 1 2 3	a b c d 4	a 0 1 b 0 1 2 3 c 0 1 2 3	
Ernest Forbes	a (i) (ii) 1 b (i) (ii) 0	a b c 0	a b c d 2 e 0 1 2 0 f 0 1 2 3	a b 0 1 2 3 c 0 1 2 3	a b c d 4	a 0 1 b 0 1 2 3 c 0 1 2 3	
Xenicxi Xhabagkh	a (i) (ii) 1 b (i) (ii) 1	a b c 0	a b c d 2 e 0 1 2 3 0 f 0 1 2 3	a b 0 1 2 3 c 0 1 2 3	a b c d 4	a 0 1 b 0 1 2 3 c 0 1 2 3	

6.10.5 Record keeping

Record keeping is an important activity during performance assessment, as they are kept to helpdevise methods that can improve the learner's development (Le Grange & Reddy, 1998). Students did not keep proper records of their practical transactions. They wrote their record on the assessment document they had been given during the practicals due to the conspicuous absence of record books, which is an important material for practicals. Some presented their record in a tabular format, with number of columns and headings differing from one student to the other. Samples of the records kept by students are presented below.

Sample 1

Date	**Activity**	**Tools/materials**	**Reasons**
12/08/2010	weeding	Digging fork	To prevent competition for nutrients and water
12/08/2010	watering	Watering can	To encourage the process of photosynthesis To cool the plants

Sample 2

Date: 12/08/10

Activity	**Reasons**	**Tools**
weeding	To avoid competition of food	Spade
watering	For easy absorption of nutrients in the soil by plants	Watering can
cultivation	To prevent water logging. To encourage good aeration	Digging fork

Others presented their records in a vertical format, placing headings below the other and writing continuously.

Sample 3

Date: 12-08-10

Activity: weeding

Tools: spade, rake

Reasons for weeding: to prevent/reduce competition between the weed and the crop for minerals, water, space

To maintain tidiness of the plot

Reasons spade: to cut/remove the weed with the roots in order for it to not grow again

Reasons rake: to obtain a fine tilth after weeding; to obtain tidiness around the plot

Close scrutiny of the records revealed that reasons advanced were more textbook-like than experienced during practicals. For example, the reasons given for watering in sample 1: "*To encourage the process of photosynthesis"* was more theoretical than practical. As a consequence, record keeping did not reflect any critical thinking by the students. This prompted the researcher to develop a standard form for recording transactions of the activities.

6.11 CONCLUSION

Generally, teachers' instructional practices in assessment were student-centred but the processes of assessment in a student-centred learning approach were found to be insufficient. On the other hand, teachers' knowledge of assessment was found to be average, possibly due to inadequate training on assessment as initially discussed in Subsection 5.2.3. Teachers therefore needed further training in assessment practices to improve students' performance.

Resources were found to be insufficient with the exception of the garden space. Equipments and other materials such as protective clothing were found to be in acute shortage. Scoring by teachers was holistic which provided an overview of the students' performance rather than providing a separate score for each criterion. Teachers found analytic scoring to be presenting more work in classes with more students' hence high workload. Students were not provided

with record books to record their daily activities, and recording of their transactions was not standardised. However, both teacher and students appreciated the new approach to assessment.

6.12 IMPLICATIONS FOR THE SUBSEQUENT DESIGN

The findings implied infusing and modifying a number of factors in the design of the subsequent prototype. Instructions were improved for the intervention to be self-explanatory since the document was given to students in advance, so that they understood it with little explanation from their teachers. Training of teachers was strengthened to include mock-assessment before the actual implementation to ensure that teachers grasped fundamental principles of analytic scoring. Training concentrated on how to select skills to assess a given number of students; how to complete the assessment form; and how to keep record of activities taking place and of assessment outcomes. In addition, a training manual was produced detailing procedure for administering the standard tasks and assessment materials.

A record keeping guide was developed to standardise record keeping throughout the country. Students were guided on how to record activities they carried out. A strategy was devised to encourage them to answer open-ended questions of the questionnaire to provide valuable information to improve the validity of the standard task and assessment materials.

Prior to commencement of lesson observation, the researcher worked with teachers to ensure that they thoroughly prepared for the lesson, such as organising all requisite materials, giving the assessment instrument in advance, explaining the instrument to the students, and how to the implement of the intervention. The strategy did not yield positive results and had to be changed.