

## CONSTRUCTION OF CHECKLISTS AND RATING SCALES TO MEASURE SKILLS

### INTRODUCTION

Checklists and rating scales are useful and usable instruments to measure learners *Process or Product Skills*, in many cases unless automatic testing and recording devices are to be used they are the only way.

However, there are two major difficulties to overcome before they can be used with any degree of confidence:

- They must be constructed with considerable care or a great deal of error in the instrument will occur.
- The biases and preferences of the observers or judges must be reduced - a good instrument will go some way to remedy this.

These notes are concerned with:

- Constructing checklists;
- Constructing the major types of rating scales;
- Checking the constructed checklists and rating scales.

### Constructing checklists

There are six steps to follow. If they are followed, a usable, valid and reliable checklist should result.

- Step 1: Specify an appropriate process or product.
- Step 2: List the important behaviours or characteristics or criteria.
- Step 3: Add any common errors.
- Step 4: Arrange the list of characteristics or criteria.
- Step 5: Provide a way to use the checklist.
- Step 6: Pilot it.

## **Step 1: Specify an appropriate process or product**

Checklists are not too difficult to construct and to use, but that should not be the sole reason for using them as internal validation instruments. It is naturally important to make certain that the information needed about the process and/or product that is being trained for and validated is included on the checklist. Hopefully, the analysis of the task and/or job will have provided this; if it has not, it is necessary to look again.

**Process** checklists provide information of the ‘It’s there - it’s not there’ type.

The following question should be asked about the process:

*‘Are there some performance characteristics of this kind of process which are so important that it is valuable simply to know whether or not they occur?’*

If the answer is *yes*, then a checklist is an appropriate measuring instrument to use.

A similar question should be asked about any **product** for which one wishes to construct an internal validation instrument to check learners and their learning.

‘Are there some characteristics typical of this kind of product which are so important that it is valuable to find out about whether or not they occur?’

Again, if the answer is *yes*, then a checklist is appropriate.

Below are a few examples of processes which can be internally validated. For each process *some* of the characteristics or criteria that might and could be included in a checklist have been listed:

### **Processes**

- **Playing a clarinet:**
  - Holds instrument properly
  - Has a satisfactory arm position
  - Uses correct alternate fingering where necessary
  - Checks the key signature and time
  - Counts correctly.
- **Using an electric sewing machine:**
  - Plugs in the machine
  - Threads the machine
  - Checks needle
  - Threads needle
  - Adjusts stitch length.
- **Doing a laboratory experiment:**
  - Writes down what he is trying to find out
  - Gets equipment needed
  - Checks over procedure
  - Gets pad to write down results.

**Step 2: List the important characteristics or criteria**

Examine the task closer and extend the list of criteria. For example, 'Using an electric sewing machine' might look like this:

- Plug in the machine
- Thread the machine
- Remove the bobbin
- Wind the bobbin
- Insert the bobbin
- Remove the needle
- Thread the needle
- Insert the needle
- Adjust the tension
- Oil the machine
- Adjust the stitch length
- Use the foot control
- Properly position the material to be sewn
- Start the machine (insert needle into fabric first)
- Stop the machine correctly (needle out of the fabric).

**Step 3: Add any common errors**

It is important to find out not only if desirable procedures are being followed, but also if undesirable ones are. For example, it may be important in observing someone learning to give a speech in public to find out whether the learner uses gestures - they may be desirable or undesirable ones!

It may also be important to find out if he uses any **distracting** words or mannerisms - for example, saying 'Ah', 'Er', or 'Um', or keeping his hands in his pockets. It is important not to add too many undesirable traits - add only those that are serious and very common.

**Step 4:            Arrange the list of characteristics or criteria**

A **process** checklist will be much easier to use if the list of criteria to be observed is listed in the order in which the criteria are likely to occur.

In a **product** checklist the criteria should be arranged so that the judge can begin at one place and examine it systematically, part by part.

**Step 5:            Provide a way to use a checklist**

There must be a convenient place to check each criteria as it occurs (or as it is seen if it is a product). The best way is to place blanks on the right hand side of the page next to each criterion being checked. A tick (✓) or cross (x) can be put as appropriate – if sequence is important, a space should be left to number it.

The electric sewing machine could now look like this:

<b>Criteria</b>	<b>Yes</b>	<b>No</b>	<b>Sequence</b>
Plugs in machine			
Threads machine			
Removes bobbin			
Winds bobbin			
Replaces bobbin			
Removes needle			
Threads needle			
Replaces needle			
Adjusts tension			
Adjusts stitch length			
Uses foot control			
Positions material to be sewn properly			
Starts machine (inserting needle into fabric first)			
Stops machine (needle out of fabric)			

**Step 6: Pilot it**

Try it before use – see section later, **checking for checklists and rating scales**

Other examples of checklists are added as Appendices:

- Appendix 1:** Saddling a horse
- Appendix 2:** Heavy Goods Driving –  
Performance Monitoring Test – Side 1
- Appendix 3:** Heavy Goods Driving –  
Performance Monitoring Test – Side 2
- Appendix 4:** Heavy Goods Driving –  
Performance Monitoring Test –  
Record card for 4 tests to be taken  
during a 10 day course – Side 1
- Appendix 5:** Heavy Goods Driving –  
Performance Monitoring Test –  
Record card for 4 tests to be taken  
during a 10 day course – Side 2

**Constructing rating scales**

In some cases rating scales have a definite advantage over checklists because they allow the ‘judge’ to make a systematic judgement about **the degree to which** a characteristic or criterion is present or observed.

A rating scale usually consists of a set of characteristics or criteria to be judged with some kind of scale. The judge uses the scale to indicate the quality, quantity, or level of performance observed. The points along each scale represents different degrees of the characteristic or criterion being observed. A set of directions tells the observer how the scale should be used.

**Example: Rating scale for rating discussion leaders during training****Directions**

Rate the discussion leader on *each* of the following characteristics by placing a tick (✓) anywhere along the horizontal line under each item that represents your view of what you have observed.

1. To what extent did the leader encourage discussion?

<hr/>		
Discouraged discussion by negative comments	Neither discouraged nor encouraged discussion	Encouraged discussion by negative comments

2. How well did the leader keep the discussion on the right track?

<hr/>		
Let the discussion wander	Only occasionally brought the discussion back on target	Did not let members wander from the point target

3. How well did the leader ask controversial questions?

<hr/>		
Never asked controversial questions	Occasionally asked controversial questions	Continually asked controversial questions

4. How well did the leader respond to inappropriate comments?

<hr/>		
Ridiculed the person who made the comment	Treated inappropriate comments the same as appropriate ones	Discouraged inappropriate comments

In the excerpts from the rating scale above, the observer is helped by it to focus on specific observable aspects of the traits being observed and judged. In addition, if it is used on many people learning to lead discussions, they will all be judged from a common frame of reference. This provides objectivity and improves the reliability of observation if the scale is well constructed and used.

### **Steps to take when constructing a rating scale**

- Step 1:** Specify an appropriate training or enabling objective
- Step 2:** List the important characteristics
- Step 3:** Define a scale for each characteristic
- Step 4:** Arrange the scales
- Step 5:** Write the instructions
- Step 6:** Pilot it

### **Step 1: Specify an appropriate training or enabling objective**

There are many training and enabling objectives which can be judged with the use of a rating scale. Here are a few from the field of education - add your own in the spaces provided.

- **Processes:**
  - Singing
  - Playing an instrument
  - Leading a discussion
  - Drawing a picture
  - Using a band saw
  - Solving a problem
  - Handwriting

---

---

---

---

---

---

---

---

- **Products:**  
An oil painting  
A woodworking project  
A handwriting example  
A blue print  
A typed letter  
A jar of jam

---

---

---

---

---

---

---

- **Social-personal traits:**  
Friendliness  
Politeness  
Honesty  
Patience  
Generosity  
Extroversion  
Endurance

---

---

---

---

---

---

---



## **Step 2: List the important characteristics**

This step is essentially the same as Step 2 for constructing checklists. Carefully analyse each objective and decide what is most important to the accomplishment of that objective.

### **Deciding what is important to the accomplishment of a given objective**

- If the objective is a process, ask ‘What important behaviours are exhibited by individuals who can successfully perform this process?’
- If the objective is a product, ask ‘What important features are characteristic of high quality products of this type?’
- If the objective is a result of a social-personal development, ask ‘What important behaviours are most commonly associated with this social-personal trait?’

Below are specific examples of these general questions accompanied by some possible answers. *In most cases these characteristics will have to be obtained by an analysis of ‘good’ and ‘bad’ performers.*

Note that each answer lists **only** those characteristics which are considered to be most important to the objective being rated.

#### ***‘What behaviours are exhibited by individuals who can successfully play tennis?’***

They can:

- Hit a forehand drive
- Hit a forehand lob
- Hit a backhand drive
- Hit a backhand lob
- Anticipate the opponent’s next move
- Place a ball accurately, and so forth.

#### ***‘What important features are characteristic of good charcoal sketches of nature?’***

They are:

- Well composed
- Contain bold as well as thin lines
- Free, not stiff
- Shaded so as to create a mood, and so forth

***‘What important behaviours are usually associated with friendliness?’***

They are:

- A ready smile
- A readiness to speak
- An interest in others
- A willingness to meet others
- A memory for names, and so forth

**Step 3: Define a scale for each characteristic**

A scale defines an underlying continuum. A ruler is a scale used for judging the physical characteristics of length. The continuum defining length (which theoretically ranges from zero to infinity) is marked off in some meaningful units (inches, centimetres, etc., in the case of a ruler).

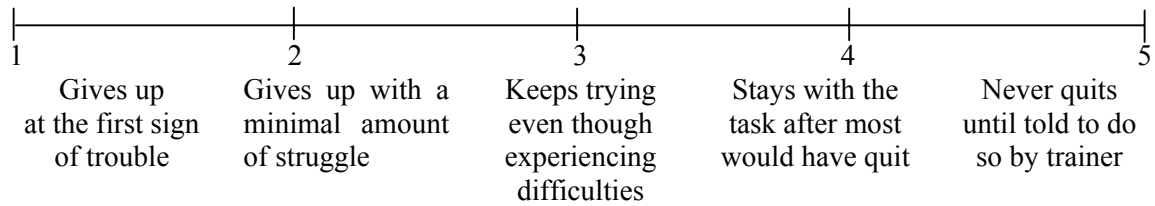
If one thinks of rating scales as rulers designed to measure the quantity or quality of certain characteristics, a number of features emerge.

Whereas rulers are used to measure length, rating scales are used to measure such characteristics as:

- Participation
- Friendliness
- Smoothness
- Appropriateness
- Frequency
- Value
- Balance
- Beauty
- Flexibility
- and so forth

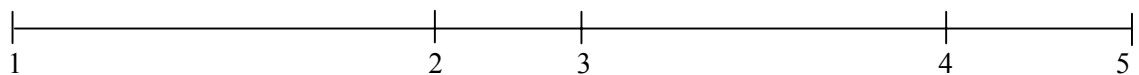
Like rulers, rating scales should also be marked off into meaningful units. However, the units on a ruler are all of equal value, and the distances between the points are equal. This is *not* true of the points along a rating scale. For example, imagine a rating scale had been constructed for rating ‘Ability to stay with a task’. The scale might look like this:

### Ability to stay with a task

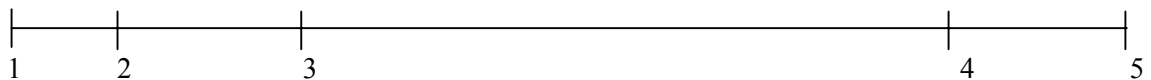


All five points are drawn equidistant from each other, but in reality the significance of the difference between a rating of 1 and a rating of 2 is not known. Nor is it known how that compares with the difference between a rating of 3 and a rating of 4. Is someone who ‘stays with a task after most would have quit’ (a 4 on the scale) twice as good as someone who ‘gives up with a minimal amount of struggle’ (a 2 on the scale)? Obviously it is not possible to answer the questions.

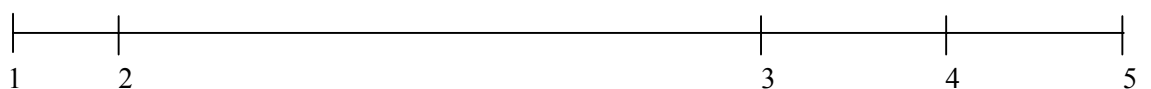
Maybe the scale would represent reality better if it were drawn as follows:



Or like this:



Or like this:



Unfortunately it is not known which drawing best represents reality. Consequently, the scales might as well be drawn *as if* all the points along them represented equal intervals. This is not to say that they should be interpreted in that way. Nor does it mean no concern should be given about this problem when constructing the scales. The problem is obviously a major source of error. Therefore, operating from the premise that it is better to prevent error from occurring than to try to cope with it later, scales should be defined as carefully as possible. The following suggestions may help.

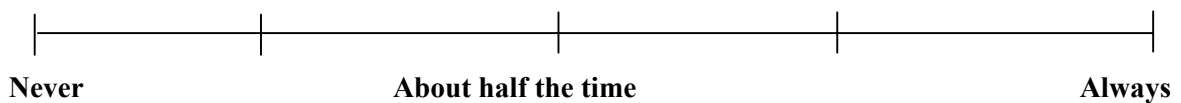
### Establish the extremes of the continuum

When attempting to define a scale, first describe the extreme ends of the continuum. If the characteristic that is being measured is most naturally rated in terms of frequency, then the task is easy. The scales could range from:

- Zero to infinity
- Never to always
- Rarely to very frequently
- 0 to 100 per cent
- 0 to 15, and so forth.

### **Example of a quantitative rating scale**

**How often does the assistant arrive late for work?**



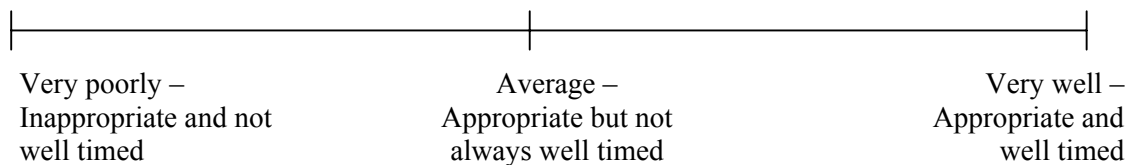
However, when the characteristics being measured is to be rated in terms of:

- ***Degree*** to which it exists
- ***Quality*** of performance
- ***Level*** of development
- ***Quality*** of product

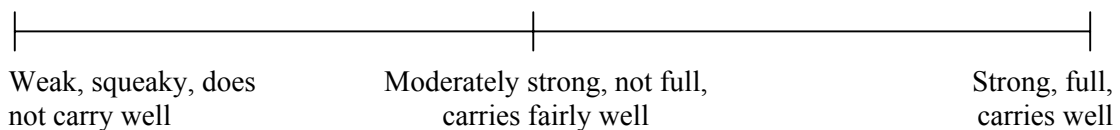
or similar qualitative terms, then the task of defining the extreme ends is far more difficult. Some sample scales are given below:

### **Examples of qualitative rating scales**

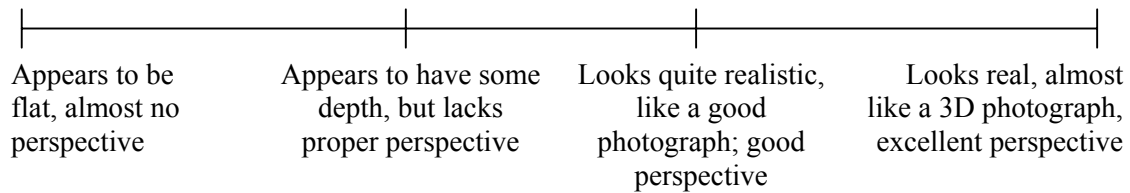
**How well does the discussion leader use gestures?**



**What is the quality of the discussion leader's voice?**



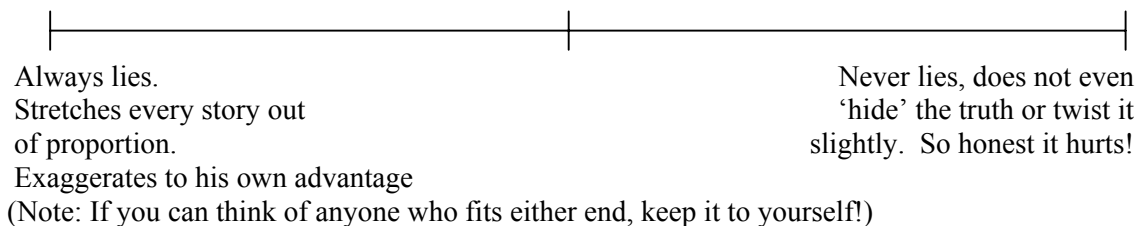
**Indicate the extent to which this drawing conveys a feeling of depth?**



In the above examples, the descriptions at either extreme describe the worst and the best a process or product could possibly be. Whenever defining the extreme ends of a continuum designed to be used in rating a process, analyse good and bad performers and describe the process that exemplifies the negative or low end of the scale. Next describe the process that exemplifies the positive or high end of the scale.

## **Example of a rating scale for a personality trait**

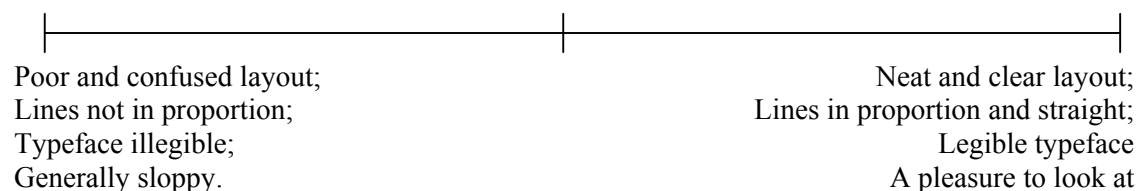
### **Honesty**



This same procedure can be used when defining the ends of a continuum designed to rate products. Analyse bad and good products – 'rejects' and 'ideal'. Describe each one.

## **Example of a scale to rate products**

### **A typed form**



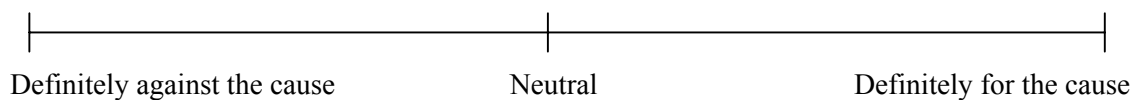
### **Describe the point between the extremes**

It is important to try to describe each point along the scale so that the difference in quantity and quality between any two points is approximately equal. Although it is not always possible to obtain equal intervals along with scale, it is possible to place the points along the scale in the correct rank order. It is important, of course, for example, to make certain that the behaviour descriptive of the fifth point of the scale *does* represent a higher quality performance than the behaviour described at the third point.

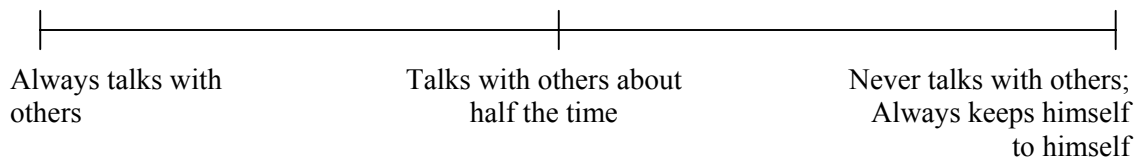
Examine the extremes and then try to describe two or three points representing behaviours (or features) which are not quite so extreme. Usually it is considered better to divide a scale into an odd number of points (3, 5 or 7). This is especially the case when there is a neutral point or a clear-cut mid-point between the two extremes. For this kind of scale, the first point to describe after the extreme ends is the mid-point.

### **Examples of behavioural rating scales**

#### **Commitment to a cause**



#### **To what extent does this officer talk with other officers?**



There are now three points clearly defined:

- Two ends, and a
- Mid-point

This may be, and often is, enough. Sometimes, however, it is helpful to give the judge(s) a little more guidance by describing a few more points.

The procedure then, is to decide two more points on the scale, each of them as a mid-point on a 'new and separate scale' on the scale already created. This is done by looking at the high and low ends of the scale in turn, regarding one end as an end, but the mid-point as the other end – then establishing a new mid-point. The 'talking' scale above would thus look something like this:

## Example of a behavioural rating scale

To what extent does this officer talk with other officers?

Always talks with others	Talks freely with others – sometimes to the exclusion of other things, but occasionally avoids the opportunity to talk	Talks with others about half of the time	Occasionally talks with others, but would rather observe, keep to himself or listen to others, than talk himself	Never talks with others, always keeps himself to himself
--------------------------	--	--	--	--

There are several ways of describing the points on a rating scale. The three most common types of description are:

- Numerical
- Graphic, and
- Descriptive

Sometimes combinations of these are used, for example:

- Descriptive-Graphic

A numerical scale is simply a list of numbers keyed to descriptive labels which remain constant from one characteristic to the next. The judges mark in some way the number which best describes the person, process, or product, being validated (checked).

## Example of numerical rating (with different marking methods)

Circle (O) the number which *best* describes the officers *relationship* with his subordinates. Use the following key:

- 5 = Excellent
- 4 = Above average
- 3 = Average
- 2 = Below average
- 1 = Poor

Tick (✓) the number which *best* describes the officer's *relationship* with his sub-ordinates. Use the following key:

- 5 = Excellent
- 4 = Above average
- 3 = Average
- 2 = Below average
- 1 = Poor

Tick (✓) the box beside the number which best describes the officer's relationship with his subordinates. Use the following key:

5 = Excellent

4 = Above average

3 = Average

2 = Below average

1 = Poor


Underline \_\_\_\_\_ the number which *best* describes the officer's *relationship* with his subordinates. Use the following key:

5 = Excellent

4 = Above average

3 = Average

2 = Below average

1 = Poor

Numerical scales are also used when frequency of occurrence is to be judged.

## **Examples of numerical rating for frequency of occurrence**

**How many times does the speaker say 'ah' during a three-minute speech?**

1 = less than 5

2 = 5 – 10

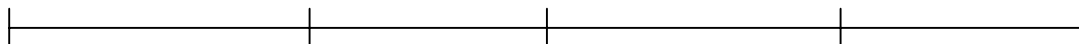
3 = 10 – 15

4 = 15 – 20

5 = more than 20

(The 'barred gate' system (//// ##) may be used to record the 'ahs' as well, before scoring the above).

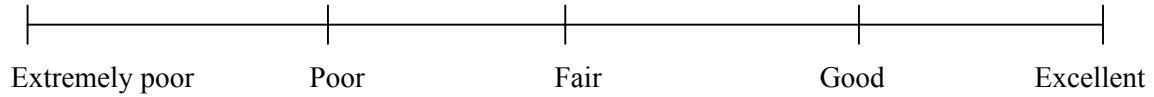
When describing points along a scale graphically, draw a horizontal (or sometimes vertical) line and divide it into equal parts. Each dividing line is accompanied by a label describing that point on the continuum. Observer(s) place their preferred mark (a ✓ or a X) anywhere on the line. This allows marking between two points, closest to the one they feel best describes the individual.



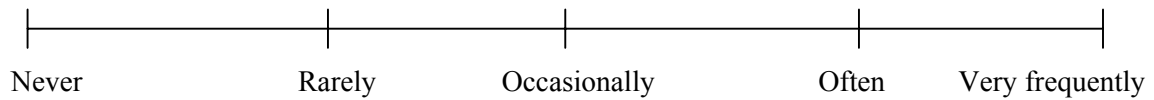


## **Examples of graphic rating scales**

**How good is this officer's relationship with his subordinates?**



**How often does the speaker say 'ah' during a three-minute speech?**

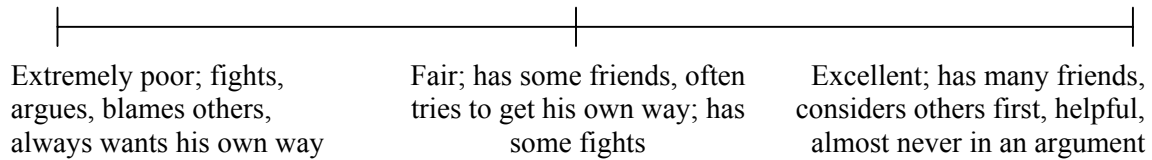


A descriptive scale implies a more complete verbal description of the points along the continuum than the verbal labels used above. It is usual to construct and use a graphic arrangement, producing a description-graphic scale.

Whenever constructing a verbal description of points along a scale, be certain to be terse, crisp, and to use action, observable terms as if writing training or enabling objectives.

## Example of a descriptive rating scale

**How good is this assistant's relationship with his fellow assistants?**

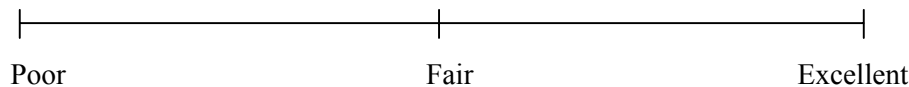


The advantage of a numerical over a graphic scale is that it yields numbers (which are easier to record and which can be manipulated). The graphic scale, on the other hand, has an advantage because the traits being rated are more explicitly stated in observable terms.

### Step 4: Arrange the scales

For any given process, product, or social-personal trait, several scales will be constructed and used (one for each of the important behaviours, characteristics or criteria). If there is any logical arrangement for the scales it is preferable to place them in that order – for example, in the order of expected occurrence.

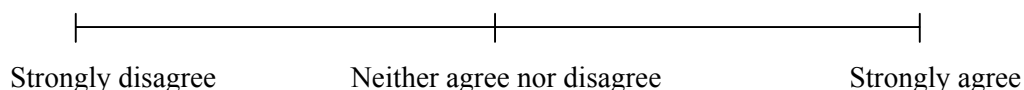
The 'direction' of the scale should be considered – it can go in one of three directions:



Or, it can go from positive to negative:



Or, it can go from strong (not necessarily positive) to neutral and back to strong:



Notice that to disagree is not necessarily negative. It depends with what you are disagreeing!

When judges are marking a number of scales in a row, it is easier if they are always arranged in the same direction, for example, negative to positive. However, with this arrangement it is easy to establish a response pattern – for example, 'This officer is pretty good, I'll mark all 4s' – and just mark all down one side of the page, marking each scale about the same.

If this is thought as likely to be a problem, it can be avoided by occasionally changing the direction of some of the scales – for example, every so often putting in a positive to negative arrangement. The expectation is that observers will have to read and consider each scale before establishing their decision.

**Step 5: Write the instructions**

Whether or not the creator of the scale is going to use it exclusively himself it is vital to write instructions that clearly specify the judges' task. This will help to reduce any error which might result from a misunderstanding of how the scale is to be used.

The instructions to the judges need not be elaborate. However, they *must* contain the following information:

The instructions accompanying any rating scale should contain the following information:

1. A statement naming or describing the overall process, product, or trait, being rated.
2. Directions on how to mark the scales.
3. Any special instructions, for example, add any comments at the end or omit any characteristics you do not feel qualified to judge.

**Examples of instructions for rating scales**

Rate the learner's ability to present a convincing argument. Place a tick (✓) anywhere on the horizontal line below each of the characteristics. In each case place the (✓) at the place which you think best describes this learner's ability.

For each feature below, circle (O) the number which best describes this apple. Make any additional comments about the apple at the bottom of this page.

The above example instructions can be adapted to fit particular situations and are the most popular. They can help produce analytical ratings – for example, for a given

- Process
- Product
- Trait

that is analysed and the most important characteristics are identified and carefully described.

Instead of trying to devise a series of descriptive scales which would define a specific attribute of a product it is better to provide actual examples of products. These sample products can be used as examples illustrating each of the points along some scale. Ratings would be made by observers by comparing the product being rated to the various sample products at the various points along the scale.

## Example

Although the following is taken from a non-industrial/commercial environment, the principles hold good.

A group of art teachers were trying to rate products produced by a group of their learners. The products were creative wax sculptures, and they found it difficult to define the important characteristics. (Creativity would be one, but how would you describe the points along a scale of creativity except as ‘not creative at all’ to ‘very creative’?) Although they could not always verbalise all their reasons, the three art teachers were in almost complete agreement as to which sculptures were the best, which the worst, and which ‘about average’. Knowing that they would be having the next group of learners do the same assignment, they decided to develop a product scale using the learners’ own work as models.

The first thing these teachers did was to select the three best and the three worst sculptures. These would be used to anchor the ends of the scale. There is nothing magical about selecting three examples at each end; often one is enough. However, they felt that by selecting three very different but equally good creative pieces as ‘best’ examples, they could demonstrate the fact that any particular form was not necessarily best. They had thus decided to use the product scale they were developing as concrete ‘objectives’ – a demonstration to future learners. By looking at the products on the scale, learners could get a fairly good idea of what would not meet the teachers’ expectations.

They identified then, three ‘about average’ sculptures as examples of the mid-point on the scale. Next, they identified some which seemed to fit about halfway between the ‘average’ and the ‘best’ and between the ‘average’ and the ‘worst’. Now they had a five-point scale. Each point on the scale was represented by three actual, learner-produced sculptures.

The final touches were placed on the scale by writing brief descriptions of the examples at each point. In the descriptions, they tried to point out the ‘good’ and ‘bad’ features. To use this scale in the future, the teachers would simply take a piece of sculpture and compare it to the examples on the scale. If it looked most like those at position 2, it would receive a score of 2, if it looked most like those at position 5, a score of 5, and so on.

‘Process’ skills scales although more difficult to construct than ‘product’ ones, can be constructed in the same way as the above example.

## Step 6: Pilot it – checking checklists and rating scales

### Checking for obvious flaws

The validity of a checklist or rating scale is determined to a great extent by which behaviours, criteria, or characteristics, were selected to be checked or rated. It is vital to be certain that the most important ones are included in the instrument. An easy first step is to obtain the views of a colleague. He needs to know in general terms what the checklist or scale is trying to measure. He needs to be asked to produce his list of important criteria and compare them with the created one. Differences need to be discussed.

Then ask him to read over the descriptions created to describe the particular behaviours listed on the checklist or placed on the scale. Ask him to ‘translate’ the descriptions into his own words. If some of the created descriptions are ambiguous, rewrite the descriptions.

Finally, have the instructions checked and rewrite the unclear ones.

### **Trying out the instrument**

Pilot the instrument and, where possible, have other people use it too. Notes should be made on any problems. For example:

- Were there times you wanted to put something down but had no place to put it?
- Were there characteristics you did not get an opportunity to observe?
- Were there behaviours as described typical of those you observed?
- Did the performance move too fast for the number of ratings you had to make?

These and many similar questions often arise as a checklist or rating scale is used. They signal problems with the instrument that could probably be corrected with a little rewriting.

Whenever piloting an instrument, a note pad should be kept handy so that any questions which may arise or any problems which may occur can be jotted down. Sometimes there will not be the time to make notes while you are using an instrument of observation. It may not be possible to jot down your impressions until *after* using the instrument. Therefore, it is advisable to plan to have extra time immediately after piloting the instrument of observation. Use this time to write down impressions while they are still fresh in the mind.

### **Validity**

There are a number of ways to check the validity of an information-gathering instrument. The easiest way for a trainer-made test is to simply verify the information which the instruments yield with ‘work done for real’, ie in the work situation, with manager, peer, and if appropriate, quality control comment. If they confirm the information gained, there is validity; if it is not confirmed, it is invalid. But also check the instrument against some other form of source of information.

### **Reliability**

It is highly unlikely that valid results will be obtained if the instrument is unreliable. However, there is a major source of error in checklists and rating scales:

- Judge’s error

The procedure to establish this is relatively simple:

Two or more judges obtain information about the same group of individuals, processes, or products, *at the same time*.

The information obtained by one judge is compared with the information obtained by another judge.

A large amount of agreement means high inter-judge reliability, a small amount of agreement means low reliability.

When correlation is used to make the comparison, the resulting coefficient gives a numerical estimate of the degree of consistency between judges.

Statistical techniques suitable for use are described in other notes.

## **Summary**

- Checklists and rating scales are useful instruments for observing learners' processes and products.
- There are a number of steps to take when constructing a checklist:
  - Specify an appropriate process or product;
  - List the important characteristics of criteria;
  - Add any common errors;
  - Arrange the list of characteristics or criteria;
  - Provide a way to use the checklist;
  - Pilot it.
- Checklists provide one with:
  - 'It's there – it's not there' kind of information.
- Rating scales allow judges to make systematic judgements about the degree to which a process, product, or characteristic, is present.
- There are a number of steps to take when constructing a rating scale:
  - Specify an appropriate training or enabling objective;
  - List the important characteristics;
  - Define a scale of each characteristic;
  - Arrange the scales;
  - Write the instructions;
  - Pilot it.
- Product scales (and sometimes) process scales, can be constructed by using actual examples (or pictures of them) of appropriate or sample products or processes.
- The validity, reliability of each of use of checklists and rating scales should be checked before they are used as well as after they have been used. Any faults should be corrected before they are used again.
- Judges should be carefully selected and trained.

## **Appendix 1**

### **Saddling a horse**

<b>Criteria</b>	<b>Yes</b>	<b>No</b>
Is the saddle properly in place on the horse?		
Is the saddle properly positioned on the saddle pad?		
Are the stirrups at the proper height for the leg length of the rider?		
Are the stirrups firmly secured to the saddle?		
Is the girth strap tightened so that the saddle will not slide?		

and so forth

## Appendix 2: Performance Monitoring Test – Marking Sheet

Name

Authority

Date

Course day/duration Test: 1 2 3 4

Time	Traffic	Weather	Finish	
Start	conditions	conditions	Start	
Finish			Total miles/ kilometres	

Notes:

				1	2	3	4	5
S t a r t	1	Fails to ensure	H/B engaged <sup>1</sup> gear in neutral <sup>2</sup> clutch down <sup>3</sup>					
	2	Moving off	Procedure <sup>1</sup> U; Hill <sup>2</sup> D/Hill <sup>3</sup> Blind spot <sup>4</sup> Stall <sup>5</sup>					
V e h i c l e  C o n t r o l	3	Gear selection	Appropriate <sup>1</sup> two speed axle <sup>2</sup>					
	4	Gear production	Up <sup>1</sup> Down <sup>2</sup> Neutral <sup>3</sup> Exercise <sup>4</sup>					
	5	Clutch	Control <sup>1</sup> Riding <sup>2</sup> Coasting <sup>3</sup> Exercise <sup>4</sup>					
	6	Footbrake	Late <sup>1</sup> Fierce <sup>2</sup> Jerky <sup>3</sup>					
	7	Handbrake	Release <sup>1</sup> Stopping <sup>2</sup> Necessary <sup>3</sup> Unnecessary <sup>4</sup>					
	8	Accelerator	Blipping <sup>1</sup> Jerky <sup>2</sup> Excessive <sup>3</sup> Insufficient <sup>4</sup>					
	9	Steering	Hold <sup>1</sup> Oversteer <sup>2</sup> Understeer <sup>3</sup>					
R o a d  B e h a v i o u r	10	Overtaking	Early <sup>1</sup> Late <sup>2</sup> Too close <sup>3</sup> Dangerous <sup>4</sup>					
	11	Progress	Road conditions <sup>1</sup> Traffic conditions <sup>2</sup>					
	12	Roundabouts	Assess <sup>1</sup> MS <sup>2</sup> Position <sup>3</sup> Speed <sup>4</sup> Observations <sup>5</sup>					
	13	Road junctions	Assess <sup>1</sup> MS <sup>2</sup> Position <sup>3</sup> Speed <sup>4</sup> Observations <sup>5</sup>					
	14	Cross roads	Assess <sup>1</sup> MS <sup>2</sup> Position <sup>3</sup> Speed <sup>4</sup> Observations <sup>5</sup>					
	15	Lane discipline	Crown <sup>1</sup> Kerb <sup>2</sup> Straddle <sup>3</sup> Change <sup>4</sup>					
	16	Following ORU	Separation distance <sup>1</sup>					
	17	Meeting ORU	Incorrect <sup>1</sup> Dangerous <sup>2</sup>					
	18	Crossing ORU	Incorrect <sup>1</sup> Dangerous <sup>2</sup>					
	19	Ped. crossings	Approach <sup>1</sup> Invite <sup>2</sup>					
M i r r o r s	20	Normal stop	Selection <sup>1</sup> Position <sup>2</sup>					
	21	Before	Starting <sup>1</sup> Adjusting speed <sup>2</sup> Stopping <sup>3</sup>					
	22	Before	Changing direction <sup>1</sup> Signals <sup>2</sup> General use <sup>3</sup>					
S i g n a l s	23	Overtaking	Before <sup>1</sup> During <sup>2</sup> After <sup>3</sup>					
	24	Indicators	Wrong <sup>1</sup> Timing <sup>2</sup> None <sup>3</sup> Unnecessary <sup>4</sup> Cancel <sup>5</sup>					
	25	Arm	Wrong <sup>1</sup> Timing <sup>2</sup> None <sup>3</sup> Unnecessary <sup>4</sup>					
S i g n s	26	Horn	Necessary <sup>1</sup> Not necessary <sup>2</sup>					
	27	Automatic T/Sig	Approach speed <sup>1</sup> Fails to comply <sup>2</sup>					
	28	Observation	Traffic signs <sup>1</sup> Road markings <sup>2</sup> Hazards <sup>3</sup>					
A t t e n t i o n	29	Traffic cont.	Fails to comply <sup>1</sup>					
	30	Other road users	Alertness and anticipation <sup>1</sup> Response to ORUs actions <sup>2</sup> Response to ORUs signals <sup>3</sup>					



Appendix 3

M a n o e u v e s	31	Forward steering	Accuracy <sup>1</sup> control <sup>2</sup> Observation <sup>3</sup>	1	2	3
	32	Reversing	Accuracy <sup>1</sup> control <sup>2</sup> Observation <sup>3</sup>			
	33	Braking	Accuracy <sup>1</sup> control <sup>2</sup> Observation <sup>3</sup>			

<div>General Comments</div>
-----------------------------



**PMT Record card**

Checks and tests		<b>Highway code and signs</b>		<b>Vehicle Safety</b>				<b>Vehicle check</b>		<b>Accessories</b>		<b>Uncoupling and recoupling</b>						
	<b>1</b>	Theory test %	Comments	Theory test %		Comments												
	<b>2</b>	%		%														
	<b>3</b>	%		%														
	<b>4</b>	%		%														
Manoeuvres	<b>Forward steering</b>					<b>Reversing</b>					<b>Braking exercise</b>							
	Test No.		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	Test No.		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	Test No.		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
	Accuracy						Accuracy						Accuracy					
	Control						Control						Control					
	Observation						Observation						Observation					
	General comments					General comments					General comments							
PMT Comments	<b>Test 1</b>																	
	<b>Test 2</b>																	
	<b>Test 3</b>																	
	<b>Test 4</b>																	