

VALIDATION OF THE FUNCTIONAL LEVELS OUTCOME MEASURE (FLOM)

Daleen Casteleijn
Sam Crawford
Lauren Graham-Parker
Lerika Karadimos
Shannon Michell

INTRODUCTION

- Measuring the effect of service delivery is becoming an essential step in the occupational therapy process. Routine outcome measurement is done for this purpose.
- In England becoming essential to measure outcomes as the NHS has adopted the “Payment by result” approach to funding and this means that routine clinical outcome measurement has to be done to show evidence of effect of service delivery.
- Although many barriers to routine outcome measurement has been reported, it has been implemented successfully in many institutions.
- Clinicians need to select the most appropriate and valid outcome measure when assessing effect of service delivery.

THE PROBLEM

- Appropriate measures must first of all be relevant for the population served, secondly it should assess the components of the intervention programme and finally, it must have sound psychometric properties, that is reliability and validity.
- A mental healthcare institution for longterm mental healthcare users has been using an instrument called the Functional Levels Outcome Measure (FLOM) but it has never been subjected to psychometric investigations.

- This presentation aims to introduce the Functional Levels Outcome Measurement (FLOM) as the ideal outcome measure for settings who apply the Vona du Toit Model of Creative Ability to large groups of adult in-patients with psychiatric diagnoses.
- One example of how the FLOM is used to direct treatment programmes for large numbers of patients and how to track changes in patients and generate evidence of the effect of service delivery will be presented.
- The emphasis of the presentation will be directed at the construct validity and reliability of the FLOM on a South African population.

Introduction to FLOM

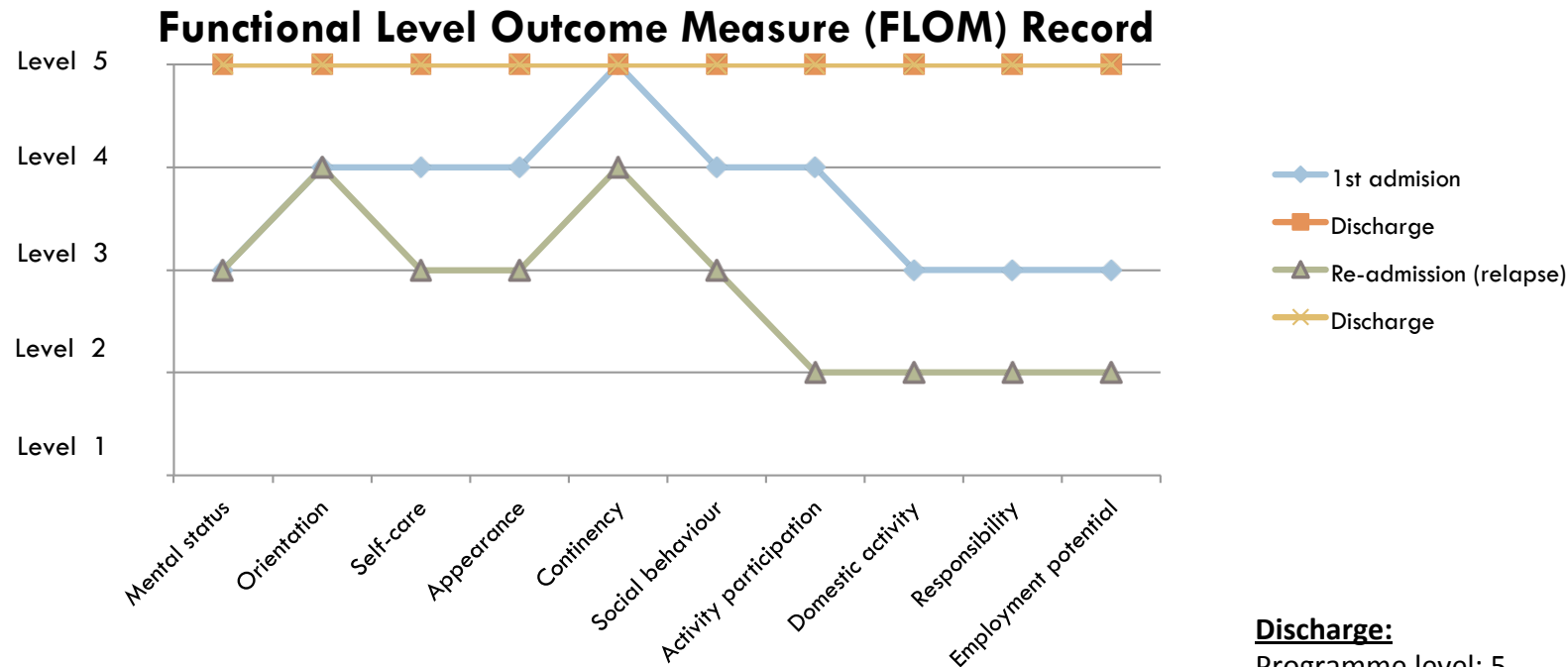
- The FLOM, previously called the Therapeutic Functional Levels Assessment (TFLA) has been used in long term institutions with large numbers of patients (500+).
- It is based on the levels of Creative Ability (CA) as described in the Vona du Toit Model of Creative Ability (VdTMoCA).
- There are 10 items in the FLOM namely mental illness, orientation, self-care, appearance, continence, social behaviour, activity participation, domestic activity, responsibility and employment potential.
- Each item consist of a set of questions for the first five levels of CA, Level 1 - Tone, Level 2 – Self differentiation, Level 3 – Intentional Explorative action, Level 4 – Norm directed action and Level 5 – norm compliance action.
- A clinician answers the question with a yes/no and the level with the most number of yes answers will be the level of the item.
- A score of 1 to 5 is allocated for each item, 1 being the lowest level of Tone and 5 the highest level of Imitative participation.

Example of questions for item of “mental illness”

Mental Illness		Yes	No
Level 1-Tone	There is life and that is all i.e. defenseless, dependant, incapable		
Level 2 Self-differentiation	Obviously mentally ill/Intellectually impaired		
	Cannot give account of self/cannot explain own feelings & actions		
	Range of emotions is very limited (Needs to fullfill basic needs like hunger)		
	Behaviour is considerably influenced by delusions or hallucinations/intellectual impairment/mental illness		
Level 3-Intentional explorative action	Start to become aware of the mental illness		
	Gives poor self account (cannot explain own feelings and actions)		
	It is difficult to understand the difference between the real world and delusions and hallucinations		
	Full range of emotions evident		
	High level of anxiety present in unknown situations		
	Has low self esteem		
Level 4 Norm directed action	Disorder not always evident		
	Able to give self account		
	Some mild symptoms are visible eg. Some depressive mood		
	Emotional response is appropriate for the situation		
	Anxiety control is poor		
Level 5 - Norm compliance action	User seems normal, illness not evident		
	Compliant with medication		
	Gives good account of self		
	Has intellectual and emotional insight into mental illness		

- A total score is calculated by summing the score for each item. A maximum total score of 50 and minimum score of 10 can be obtained.
- When all items have been assessed, the level is indicated on a line graph.
- In this institution the users are being assessed every month and the reports are being used for the Review Board as stipulated by the Mental Health Care Act.
- An example of a patient who has been treated, discharged and re-admitted shows how the patient was successfully tracked.

Line graph with 4 assessments



1st Admission:

Programme level: 4
 Score 37/50
 Date: 08/05/2010
 Signature: (HCP)
 Comments: Baseline audit

Discharge:

Programme level: 5
 Score 50/50
 Date: 27/09/2011
 Signature: (HCP)
 Comments: Moved to Independent Living Unit

Re-admission after relapse:

Programme level: 3
 Score 28/50
 Date: 04/11/2011
 Signature: (HCP)
 Comments: Did not adapt to NGO, not sufficient support for taking medication, relapse, admitted with severe depression

Discharge:

Programme level: 5
 Score 50/50
 Date: 31/05/2012
 Signature: (HCP)
 Comments: Gradually recovered from depression, started to take up painting of portraits again (professional artist), successful. Moved to another NGO with sufficient support.

- 
- The name has recently changed from TFLA to FLOM as clinicians realised that this measure is actually an outcome measure and not an assessment.
 - The FLOM is available from Kobie Zietsman who gives training to OT and nursing staff in the application of the FLOM.

Objectives

- Investigate the validity of the FLOM in terms of:
 - ▣ Threshold ordering – to see if the categories/levels of the scale are sequentially ordered.
 - ▣ Individual item functioning – to see which items were more difficult to assess than others.
 - ▣ Local independence of the items – each item contributing independently of the other 9 items to the overall construct.
 - ▣ Unidimensionality – whether all items contribute to one overall construct.

Objectives

- Determine the reliability in terms of:
 - ▣ Inter-rater reliability – similarity of scores among several raters (OT staff)
 - ▣ Internal consistency – does the tool consistently give the same scores

METHODOLOGY

□ **Research design**

- ▣ This research consisted of a non-experimental, quantitative study.
- ▣ It focused on numerical data and the analysis thereof
- ▣ This research was also descriptive because it described the psychometric properties of the FLOM
- ▣ The objective of validity was a secondary analysis as completed FLOMs (routinely done in the institution) were provided by the OT manager.

Population and sampling

- The study objectives had different samples. There were two sets of samples.
- The objective of validity and internal consistency used a sample of completed FLOM records of MHCU = sample 1
- The inter-rater reliability objective used a sample of 8 OT technicians/assistants who all assessed the same 3 clients. They had different levels of experience with the FLOM = sample 2

Data collection: validity

- We selected the first 309 completed FLOM forms from the 510 provided. Some forms were not complete e.g. gender not filled in or all items not completed.
- The sample size of 309 was guided by the typical sample sizes for construct validity investigations which is ten or more records per item.

Data collection: Inter-rater reliability

- The 8 OTT/OTAs sat in the same room while one member of the MDT carried out a presentation of each of the 3 clients.
- This presentation included a full medical and social background as well as functional problems that the client presented with.
- Each OTA/OTT then moved through section by section of the FLOM and scored each client without talking to each other.
- The completed FLOM records from the participating members of staff were collected and prepared for data analysis.

Data analysis - validity

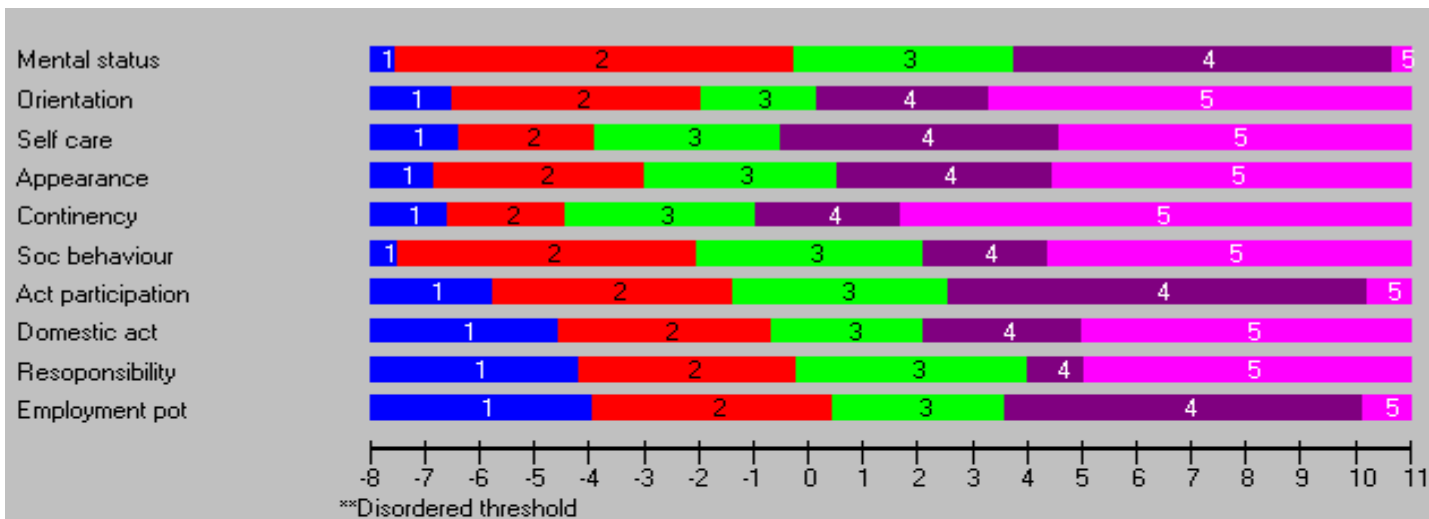
- ▣ The Rasch Measurement Model was used to determine the validity:
 - Threshold ordering – each number on the scale had to appear on a line in a sequential order.
 - Summary statistics – to see if there is an insignificant difference between observed and expected scores.
 - Local independence of the items – each item contributing independently of the other 9 items to the overall construct.
 - Unidimensionality – whether all items contribute to one overall construct.

DATA ANALYSIS – inter-rater reliability

- Correlations as well as descriptive analysis such as medians and standard deviations were calculated.
- A correlation matrix was done as well as a box and whisker plot to show the similarity among the 8 raters.

RESULTS

The threshold map indicates that the categories on the scale of the FLOM are ordered in a linear fashion.



Local independency

	Item	Mental health	Orientation	Selfcare	Appearance	Continenence	Social behaviour	Activity participation	Domestic activity	Responsibility
Mental health	I0001									
Orientation	I0002	0,188								
Selfcare	I0003	-0,089	-0,03							
Appearance	I0004	-0,118	-0,144	0,253						
Continenence	I0005	-0,109	-0,115	-0,117	-0,085					
Social behaviour	I0006	-0,024	0,008	-0,066	-0,156	-0,143				
Activity participation	I0007	-0,176	-0,073	-0,128	-0,16	-0,213	0,149			
Domestic activity	I0008	-0,139	-0,274	-0,21	-0,174	-0,093	-0,29	-0,087		
Responsibility	I0009	-0,186	-0,264	-0,358	-0,218	-0,188	-0,239	-0,119	0,221	
Employment potential	I0010	-0,112	-0,284	-0,291	-0,238	-0,183	-0,227	-0,153	0,1	0,513

Items with a correlation > 0.01 show dependence

Resolution

	Item	Mental Illness + Orientation	Selfcare + Appearance	Continence	Social + activity participation
Mental Illness + Orientation	ST01				
Selfcare + Appearance	ST02	-0,147			
Continence	ST03	-0,164	-0,114		
Social + activity participation	ST04	-0,114	-0,235	-0,229	
Domestic, Responsibility, Employment potential	ST05	-0,391	-0,42	-0,224	-0,317

UNIDIMENSIONALITY

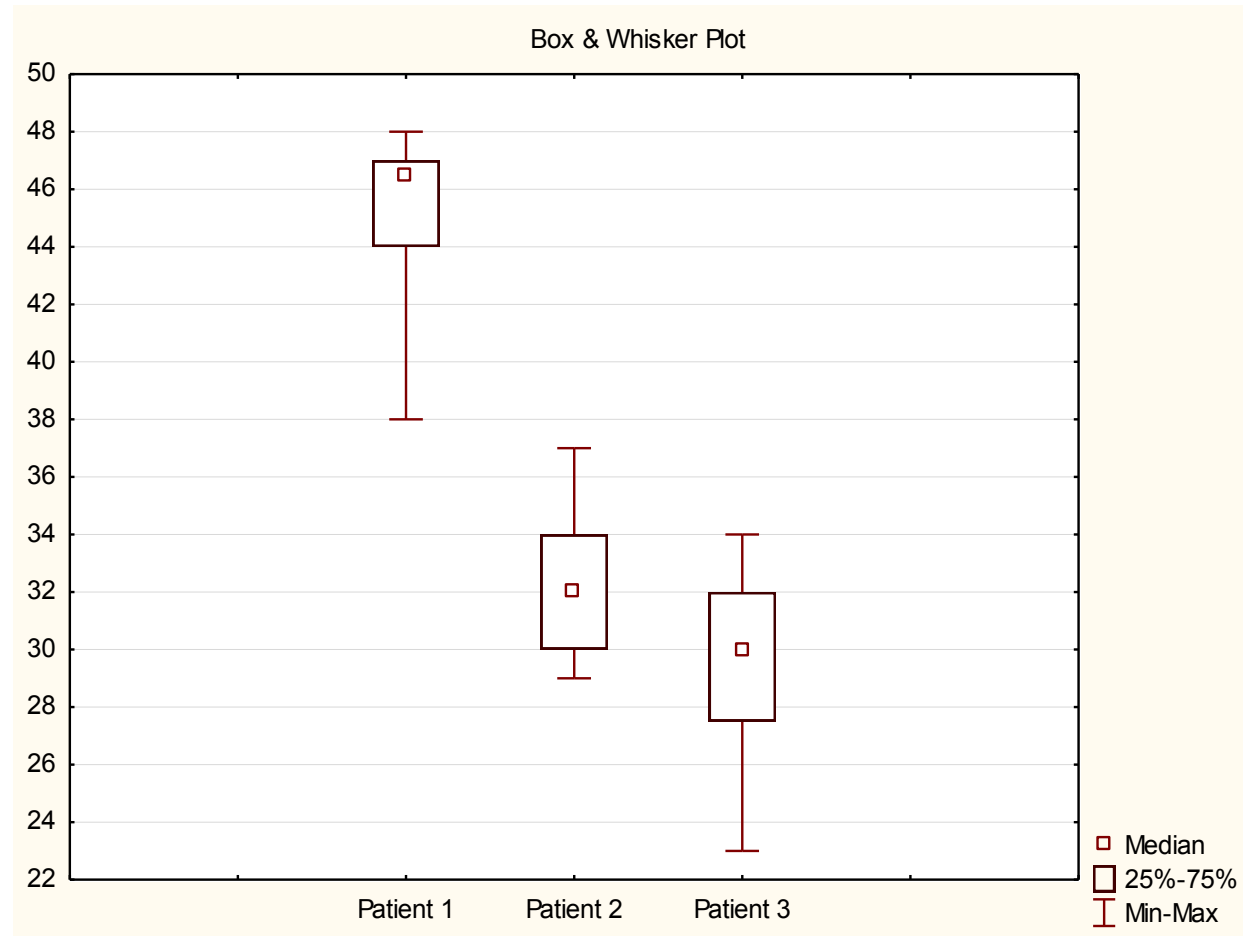
- The FLOM with the 10 items did not conform to unidimensionality criteria.
- More than 5% fell outside the range
- The item continence is theoretically a client factor and not strictly speaking on a functional level although it impacts on independent functioning.
- When the item Continence was deleted, unidimensionality improved considerably with only 3.7% outside the range and a p-value of 0.036.

Inter-rater reliability

	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Rater 7
Rater 1							
Rater 2	0,90						
Rater 3	0,98	0,97					
Rater 4	1,00	0,93	0,99				
Rater 5	0,99	0,84	0,95	0,98			
Rater 6	0,99	0,95	1,00	1,00	0,96		
Rater 7	0,87	0,57	0,76	0,83	0,92	0,79	
Rater 8	0,97	0,98	1,00	0,99	0,94	1,00	0,74

Average correlation = 0.92

INTER-RATER RELIABILITY.



INTERNAL CONSISTENCY

- The internal consistency determines whether the items in a test will provide consistent scores at one point in time.
- In the Rasch measurement model the internal consistency is measured by the person separation index, this should have a minimum value of 0.85 to support internal consistency of the scale.
- The person separation index found in the FLOM was 0.939 which indicates that the items of the FLOM provides consistent scores

Discussion

- The internal construct validity of the FLOM yielded positive results. The Rasch analysis showed that all items are ordered on a 5-category scale, each category representing a level of CA.
- Summary fit statistics and individual item fit indicated that the data fitted the Rasch model and therefore the ordinal scale was converted to an interval scale
- local independence were achieved after 5 sub-tests were formed
- Unidimensionality was achieved when the item continence were dropped.

Discussions

- Inter-rater reliability was good with 8 OTTs scoring 3 patients. Although some variation among the item scores, all indicated the same level when scores were summed.
- Average correlation was high 0.92

Implications

- The FLOM has a scale with ordered thresholds. The summary fit statistics fitted the Rasch model which means that the scores of each item may be summed to get to the total score.
- However, the unidimensionality was not good with the item continence included. When this item was removed, unidimensionality was achieved.
- Implications of this is that the sub-tests 1, 2, 4 and 5 can be summed and the item continence should be interpreted on its own.

Recommendations

- The score of the item continence should be separated from the rest of the items and should not be summed together to get the level of CA.
- It could be moved as the last item in the scale and presented in the line graph as the last item.