AARE Conference Presentation 2005

Applying Item Response Modelling to confirm the underlying construct of a new process instrument in Gymnastics

Refereed Paper Presented at

AUSTRALIAN ASSOCIATION FOR RESEARCH IN EDUCATION CONFERENCE


by

John Haynes, Judith Miller, Rosemary Callingham and John Pegg

University of New England

Abstract

Three cohorts were observed performing a gymnastic locomotor skill – the forward roll. These groups included: children (n=48) aged between 6 years and 15 years; young adults (n= 24) aged between 18 years and 22 years; and, older adults (n = 45) aged between 28 years and 42 years. Qualitative analysis of visually recorded data resulted in the forward roll being categorised into three sequences, the beginning, middle and end. For each sequence a number of indicators were identified. Each indicator was further sub-divided into a number of descriptors. These descriptors were hierarchically ranked and using the Quest (ACER) statistical package a Rasch analysis was employed to determine whether the descriptors confirmed the underlying construct. The application of the Rasch model was confirmatory for identifying the structure of a range of quality movement performances.
Introduction

This study is part of a larger investigation to determine whether a movement skill, namely, the forward roll, could be analysed using Biggs and Collis’ (1980) Structure of Observed Learning Outcomes (SOLO) model. Specifically, whether the cycles of learning found in the other modes of the model are present in the sensorimotor or motor learning mode.

This paper reports on the processes that come before analysis of data from a SOLO perspective. These processes involve, firstly, an examination of existing models used ‘measure’ the forward roll are adequate across diverse skill levels for a range of subjects. Secondly, a new model, which demonstrates dissent from previous models, is proposed. In the first instance several existing models of skill assessment were applied to the sample of subjects to examine performances.

Movement analysis was conducted of the performances for nine purposively-selected subjects from three cohorts, i.e., children (C), aged between 6 and 15 years, young adults, aged between 18 and 22 years (YA) and older adults (OA) aged between 28 and 42 years, from a total of one hundred and seventeen \( (n = 117) \). The criterion for selection was the quality of movement whilst performing the forward roll for each subject, based upon George’s (1980) descriptions of the ideal gymnastic form. Three different levels of movement quality for each cohort are represented in the study.

The instruments used to initially assess the subjects are:

(i) George (1980) provided a Gymnastics perspective, which is based on biomechanically and aesthetically ‘correct’ form.

(ii) Roberton and Halverson’s (1977) Developmental Phases perspective, was modified by Williams (1980), to include hierarchically arranged steps ordered from primitive to advanced for what is termed an initial and completion phase of the roll.

(iii) Gallahue and Ozmun’s (1998) Developmental Stages of body rolling approach. These stages are termed initial, elementary and mature.


Some shortcomings were noted when these models were applied across the different cohorts. Ensuing analysis of all subjects led to the emergence of three hypothesised sequences within the rolling action; the beginning, bridging and end. Further examination of the data led to the development of observation cues that were found to be useful for determining quality of performance. These cues are termed indicators and descriptors. The indicators are the position of various body parts within each sequence and the descriptors are subdivisions within each indicator used to portray and differentiate the range of movement quality.

Emerging from the analysis of the data is a new Model that can be used to gauge the quality of movement for the forward roll. Data that led to the construction of
this framework was subsequently analysed using the Rasch (1980) partial credit model.

**The Conceptual Framework**

Knudson and Morrison (1997:25) have stated:

observational and comprehensive models of qualitative analysis in the discipline of kinesiology (physical education) have commonalities and could be used by almost any profession in the qualitative analysis of human movement.

These researchers also indicated that “there are four important tasks in a comprehensive view of qualitative analysis: preparation; observation; evaluation/diagnosis; and intervention” (p.26). This scaffold provides a comprehensive, integrated model, which includes an outline of important issues associated with each process and is illustrated in Figure 1.

![Figure 1: A Comprehensive, Integrated Model of Qualitative Analysis](Image)

Slight modifications were made to this model to include terms applicable to this study, for example ‘forward roll’. The four processes involved have been retained including the ‘Intervention’ category, even though this aspect of the model was not part of this study.

Figure 1 illustrates the flow of processes involved in qualitative analysis. Brief details relating to each step, commencing with the Preparation stage, followed by the Observation stage and finally the Assessment stages are outlined.
Preparation Step

This step involves issues including, having knowledge of the elements and the critical features of the forward roll. In addition, knowledge of the performers’ capabilities and other personal aspects of the subjects, such as motivation, the gymnastic culture and interpersonal relationships are advantageous.

Satisfying the first aspect of the preparation step was axiomatic as the principal author is a qualified coach of Women’s (WAG) and Men’s (MAG) Amateur Gymnastics at Level 1 of the Australian National Coaching Accreditation Scheme (NCAS) and has several decades of teaching in the subject of Physical Education as a qualified educator. This background has ensured considerable knowledge of fundamental movement skill analysis, the critical features of movement analysis, and the goal of the forward rolling movement, which is to transfer weight over adjacent body parts including the hands, shoulders (sometimes head) rounded back, buttocks and feet.

The preparation step also requires knowledge of the performers. The minimization of data errors requires an awareness of the predispositions of the target population as well as the possible circumstances surrounding their motivation for participation. To alleviate possible problems associated with researcher/participant conflict, the selected cohorts were introduced to the principal researcher either as a coach or as a lecturer in gymnastics education prior to data gathering. All subjects from each cohort and their carer(s), where applicable, were volunteers. They were introduced to this project through the provision of written information. This process allowed for familiarization with the research and the researcher’s aims.

Relevant systematic observational strategies are an essential part of the preparation stage. The data gathering plan included recording on videotape, from three different aspects, i.e., from a lateral (side) position, an anterior (approaching) position, and from a posterior (back) position. This recording of data from three aspects would allow for post performance scrutiny of video material, which covered the range of possible movements.

Observation Step

The data were gathered from three cohorts performing the forward roll. Cohort 1 included children whose age is less than 18 years. Cohort 2 included young adults, i.e., those persons between the ages of 18 and 22 years. Cohort 3 comprises older adults whose age is between 28 years and 42 years.

Prior to performing the roll each subject was requested to stand at the end of the designated area, and asked perform a single forward roll then return to the start position and perform a second roll. At the conclusion of these two rolls the subjects were required to move to the second designated area where they were instructed to stand at one end of the second set of gymnastic mats and roll once. The subjects could roll either towards the camera or away from it in the first instance and then perform one roll moving in the opposite direction. Subjects
were instructed to perform these rolls on the verbal instruction ‘go.’ This instruction was issued immediately after the video camera was switched to record mode.

During data gathering some deviations to the proposed techniques were required when videotaping took place at the site where the data pertaining to children were gathered. The gymnastics centre manager considered that videotaping might interfere with practice, and in addition, if poorer performers were videotaped, the centre might ‘look bad’. Thus some videotaping took place when angles/locations were not standardized, i.e., the camera was not directly in front of, behind or at right angles to the performance on some occasions. To some extent this perceived limitation was overcome when individuals could be videotaped separately, at the conclusion of their regular practice program. Individuals videotaped separately were selected on the basis of their status as either beginners (or who were attending the centre for the first time) or on their grading across a range of gymnastic levels.

To aid analysis data observations were transcribed into written form, couched largely in anatomical terminology as exemplified by Luttgens, Deutsch and Hamilton (1992: 627 - 633). In addition, the videotaped data were transferred onto compact disks (CD).

**Assessment Step**

This step involves evaluation and diagnosis of the movement performance. The following two sections provide details, firstly, relating to the qualitative analysis of the forward roll and secondly, details of a quantitative analysis using Rasch (1980).

**Qualitative analysis**

The four models, which were previously examined for the assessment the forward roll, did not comprehensively measure the quality of performance of either young adults or older adults. The following description, which incorporates a new model, is offered as an alternative instrument suitable across the lifespan.

Qualitative analysis of data for the roll revealed three sequences, i.e., *beginning, bridging* and *end*. From the investigation of movement patterns and positions of all subjects’ during these sequences it became apparent that some body configurations were more crucial than others in the execution of the roll; these configurations are termed *indicators*, which are representative of the quality of a performance. Following is an outline of the *indicators* for each sequence of the roll.

The four *indicators* for the *beginning* sequence are:

(i) The position and placement of the hands on the surface.

(ii) The position of the arms and the elbows of those of low and medium quality display different positions and configurations than those of high quality.

(iii) The part of the head making contact with the surface indicates the most apparent difference between the subjects.
(iv) The number of body contact points at the commencement of the roll with the surface clearly delineates the different quality levels across all cohorts.

The two indicators for the bridging sequence are:

(i) The position of the hips and knees as the individual rotates.
(ii) The position of the arms during rotation

The three indicators for end sequence are:

(i) The leg movements, which are apparent towards the end of the roll.
(ii) The position of the feet at the end of the rotation.
(iii) The rotational factors that contribute to errors when attempting to attain a standing position.

For each of the indicators a hierarchical set of descriptors emerged from the data. They provide a precise account of the limb and body position of each subject ($n=117$). Both indicators and descriptors are used to assist with deeper analysis of the quality of movements exhibited within and between cohorts. For example, the beginning sequence, as outlined above, has four indicators. For each indicator, four descriptors (common to all cohorts) emerged from the data, these are: hands shoulder width apart; hands close to shoulder width; hands wide; and hands less than shoulder width. These hand positions are used to make comparisons between subjects across all cohorts. This process is replicated for each sequence of the roll for their indicators and descriptors.

Utilising, both the indicators and descriptors a model was constructed that can be applied across all cohorts. Figure 2 depicts each part of the process, which when followed vertically downward involves increasingly detailed information. The descriptor codes are represented in parentheses.
Figure 2: A Model for Assessing Movement Quality of the Forward Roll

This model (Figure 2), which emerged from the data analysis, holds potential for assessing the quality of movement for the forward roll for a wide developmental range of subjects.
Précis of Qualitative Analysis: Figure 2

A schematic overview of the instrument derived from the qualitative analysis of the forward roll is presented in Figure 2. The information accompanying each arrowed link informs the process of analysis and provides a broad template for assessment. Working through the model, commencing with the three sequences of the forward roll, the beginning, bridging or end, by following the arrows to the indicators and then to the descriptors for each sequence one can begin to determine the quality of a performance.

The descriptors, which are presented in abbreviated format and accompanied by their code (within each descriptor box), are arranged in hierarchical order from the most ideal to the least. The first descriptor from each indicator box determines a performance of highest quality. The last descriptor from each indicator box delineates the lowest quality. However, for each individual descriptor the degree of quality of performance level can vary.

For the beginning sequence, for example, the hand position the first descriptor (1) points to a high quality performance, the second to a medium, whereas the third is low quality. Likewise for the arm/elbow the first descriptor relates to high quality, the second and third descriptor medium and the fourth descriptor low quality. The first descriptor of the head position indicates the highest quality, the second descriptor indicates slightly less quality, and the third descriptor indicates low quality.

The descriptors for the number of body contacts with the surface deviate slightly from a pattern of an increasing number of contact points, as three body contact points is not as desirable as four contact points. Three contact points indicate an undesirable asymmetrical configuration of the body.

The first descriptor for the bridging sequence, i.e., hip/knee position indicates a performance of high quality. The legs remaining in a ‘tucked’ position (legs bent) indicates medium quality. The last two descriptors indicate low quality. For shoulder/arm descriptors the first indicates a high quality performance. The second descriptor is a medium quality performance whilst the remainder are arranged hierarchically down to delineate increasingly lower quality performances.

The first descriptors for the end sequence; the three leg positions indicate high, medium and low quality. With the exception of the first descriptor, i.e., ‘legs together’, a range of movements is included for each of the other descriptors. Firstly, for descriptor 2, feet apart with knees together, or knees apart with feet together form a single descriptor indicating a medium quality performance. Secondly, ‘leg movements’ related to the legs crossed, separated laterally or separated longitudinally are grouped, which represent lower quality performance. Whereas the leg positions are expressed through different movements, the three feet descriptor positions indicate high, medium and low quality, one descriptor for each quality. The descriptors for the ‘final movements’ indicator for the end sequence, in the first instance, persons with high quality performances can rise to
a standing position smoothly. Of the remaining movement descriptors, i.e., if the performer loses balance after attaining a standing position (2) or there is a momentary pause in the roll (3) this would indicate medium quality. Finally, if the hands are used to provide forceful thrust on the surface (4) or the roll stops (5) the movements are deemed to be low quality.

Generally, high quality performances display all the first mentioned descriptors. Medium quality performances may show some descriptors from high quality and some from medium quality. Low quality performances will show mainly descriptors at the end of each descriptor list. However, the performers may not always exhibit descriptors that place them in discrete high/medium/low categories. For example, if performers cannot be placed discretely within a category there is the suggestion that a performance continuum exists, along which an individual may be placed.

The following section provides details related to how the descriptors, explained in the section above, are analysed using Rasch.

**Rasch Analysis**

“The Rasch model, provides a useful method for attaining approximate measures that assist with the understanding of the processes underlying the reason why people and specifically chosen items behave in a particular way” (Bond & Fox, 2001:19). Thus Rasch measurements are particularly suited to investigations in the wide range of human sciences (Bond & Fox, 2001:189), which according to these authors is the only technique generally available for constructing such measures.

The software package used to calculate the Rasch scales is termed *Quest*, developed by the Australian Council of Educational Research (ACER). A feature of this model is that “when a variable indicating a single particular construct has been identified, within a targeted population, the measurement of the subject’s ability (in this case movement quality) is independent of the set of items that were administered, and the item difficulty is independent of the set of persons used to calibrate the item” (Snyder & Sheehan, 1992: 88). *Quest* output includes estimates of item difficulty and subject ability for both dichotomous and partial credit observations. The data for the forward roll was scored polytomously and thus the partial credit form of the Rasch model is employed. The partial credit model specifically incorporates the possibility of having differing numbers of steps for different items on the same test (Masters, 1982).

The organization of the data for analysis was based on the descriptors. The hierarchically ordered descriptors were allocated a numeral to represent their ranking. The actions judged to be the lowest quality are allocated a zero, through to highest quality, which is given a sequential numeral depending the number of descriptors. For example, for the beginning sequence the descriptors for hand position are allocated the following numerals: shoulder width ‘0’; close to shoulder width ‘1’; and wide of shoulder width ‘2’. This procedure is repeated for each descriptor for the other indicators. In addition, each sequence was allocated a code. The codes for the beginning sequence are all given the initial, ‘B’. The bridging sequences (or middle) have the tag ‘M’, and the end sequences, ‘E’. Each indicator was allocated additional letters, providing a unique cryptogram. The following Table 1 provides a summary
pertaining to \textit{indicator} codes, the acronym for each indicator and numerical \textit{descriptor} range.

\textbf{Table 1: Indicator Codes and Descriptor Range for the Three Sequences}

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Indicator</th>
<th>Code</th>
<th>Descriptor Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginning (B)</td>
<td>Hand position</td>
<td>BAH</td>
<td>0-2</td>
</tr>
<tr>
<td></td>
<td>Arm/elbow</td>
<td>BAE</td>
<td>0-2</td>
</tr>
<tr>
<td></td>
<td>Head contact</td>
<td>BHT</td>
<td>0-2</td>
</tr>
<tr>
<td></td>
<td>Contact points</td>
<td>BCP</td>
<td>0-4</td>
</tr>
<tr>
<td>Bridging (M)</td>
<td>Hip/knee</td>
<td>MHK</td>
<td>0-3</td>
</tr>
<tr>
<td></td>
<td>Shoulder/arm</td>
<td>MSA</td>
<td>0-4</td>
</tr>
<tr>
<td>End (E)</td>
<td>Final Leg movements</td>
<td>ELM</td>
<td>0-2</td>
</tr>
<tr>
<td></td>
<td>Foot placement</td>
<td>EFP</td>
<td>0-2</td>
</tr>
<tr>
<td></td>
<td>Final rotational movements</td>
<td>ERM</td>
<td>0-4</td>
</tr>
</tbody>
</table>

The first column, in Table 1 indicates the sequences for the forward roll. The second column lists the \textit{indicators} for each sequence. The third column lists codes for all subjects. The final column gives the possible range of numerals allocated to each \textit{descriptor}, which represents the range of the quality.

As quantitative analysis provides additional information relating to the subjects’ movement quality, “meaning is added to the qualitative analysis” (Bond & Fox, 2001:14). To this end an estimate of the difficulty ranking is provided for the items by comparing them with the subjects’ success rates, producing fit statistics that aid in the identification of the discriminatory nature of the items.

The output from \textit{Quest} shown in Table 2 includes the item estimates and reliability statistics. The item reliability index provides an indication of the degree to which the range and distribution of item difficulty levels is sufficient to differentiate between subjects of near equal movement quality. The \textit{Quest} commands used in the analysis may be found in Bond and Fox (2001: 104).
Table 2: Summary of Item Estimates

| Item Estimates (Thresholds) | all on all (N = 117 L = 9 Probability Level= .50) |

---

Summary of item Estimates

<table>
<thead>
<tr>
<th>Mean</th>
<th>0.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD</td>
<td>.91</td>
</tr>
<tr>
<td>SD (adjusted)</td>
<td>.77</td>
</tr>
<tr>
<td>Reliability of estimate</td>
<td>.71</td>
</tr>
</tbody>
</table>

Fit Statistics

<table>
<thead>
<tr>
<th>Infit Mean Square</th>
<th>Outfit Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.01</td>
</tr>
<tr>
<td>SD</td>
<td>.07</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Infit t</th>
<th>Outfit t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>.10</td>
</tr>
<tr>
<td>SD</td>
<td>.61</td>
</tr>
</tbody>
</table>

0 items with zero scores
0 items with perfect scores

The reliability of the item estimates is satisfactory at .71 on a 0 to 1 scale. This indicates confidence in the replication of item placement across other samples. That is, if the descriptors for the indicators for the roll were applied to other samples for which they were suitably similar then the order of item estimates would probably be replicated. Item reliability can be interpreted in the same way as Cronbach’s alpha (Cronbach & Meehl, 1955), which is the “extent to which [a] … test [or score] may be said to measure a theoretical construct or trait” (Overton, 1999 as cited in Bond & Fox, 2001: 192).

The fit statistics are close to 1 (1.01 and 1.00), for the unstandardised fit estimates with both the infit and outfit mean squares showing little spread from the ideal. When fit scores are standardised, the mean square values are transformed so they are distributed like t, with a mean of 0 and a standard deviation of 1. In this case the t values of .10 and .13, indicate that the items are useful for the sample of subjects. The note at the end of the table indicating that ‘no’ items were too easy or too difficult also verifies this output information.

The output shown in Table 3 includes the case estimates and reliability statistics.
Table 3: Summary of Case Estimates

<table>
<thead>
<tr>
<th>Mean</th>
<th>-.12</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD</td>
<td>1.61</td>
</tr>
<tr>
<td>SD (adjusted)</td>
<td>1.50</td>
</tr>
<tr>
<td>Reliability of estimate</td>
<td>.87</td>
</tr>
</tbody>
</table>

Fit Statistics

<table>
<thead>
<tr>
<th>Infit Mean Square</th>
<th>Outfit Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>.99</td>
</tr>
<tr>
<td>SD</td>
<td>.54</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Infit t</th>
<th>Outfit t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-.01</td>
</tr>
<tr>
<td>SD</td>
<td>1.00</td>
</tr>
</tbody>
</table>

0 cases with zero scores
4 cases with perfect scores

The output shows that the reliability of the person movement quality is high at .87. The mean of the infit squares at .99 and the outfit means squares at 1.00 are close to or the same as the Rasch modelled expectations of 1.00. Consequently, they produce standardised fit t values around zero (-.01 and .15). The mean person estimate (i.e., group average) is close to 0 indicating a well-matched item list. The person movement quality estimate mean of -.12 is an indicator that, for the individuals involved, the items are slightly difficult to achieve. The standard deviation of 1.61 for person estimates indicates a greater spread of person measures or variation in those measures than with item measures. The spread in the modelled fit scores for persons (infit t SD = 1.00 and outfit t SD = .85) suggests that the person movement quality estimates have error estimates well inside the conventionally acceptable range of -2 to +2.

There are no cases with zero scores, but 4 persons (subjects) with perfect scores. These subjects have been identified as exceptional (elite) gymnasts, and therefore items in general do discriminate movement quality for the majority of this group. The fact that 4 subjects are outside the scope of the framework means, however, that an additional more difficult indicator may need to be included in the repertoire of items in future research. Graham, Holt/Hale & Parker (1998) refer to the use of apparatus such as rolling on objects, e.g., the balance beam or aerial versions of the forward roll (saltos) as a more advanced skill.

Item Fit
The *Quest* generated item infit mean square map identifying those items with infit mean square values that fall inside and/or outside the interval of .77 and 1.30. This is the interval, suggested by (Wright & Stone, 1979) within which items should lie if they are jointly to represent a single underlying construct, which in this case is movement quality. This is represented in Figure 3.

Figure 3 also incorporates a theoretical idealisation (also called construct or fiction or latent trait) of the data’s interrelations, an unachievable state that is mathematically represented as the ideal straight line (seen here as the dashed line in middle of the map). The concept involves observing whether the data is on (near as possible) to this hypothetical line (Bond & Fox, 2001: 28). Thus answering the question concerning whether the data fits with one’s fiction!

In developing the straight line fit, Rasch addresses the concept of order, which is basic to the idea of measuring the human condition and is central to popular surveying tools such as a Likert scale. Performances are attributed relative importance in proportion to the position they hold on the measurement continuum. Thus, when a data matrix reflects a successful attempt to implement a theoretically guided line of inquiry with a sample for which the inquiry was appropriate, then a number of propositions are supportable. These include, (i) more able persons have a greater likelihood of correctly achieving all items in the observation schedule, and (ii) easier items are more likely to be reached correctly by all persons (Bond & Fox, 2001: 2).

---

**Figure 3: Item Fit for 9 Indicators of the Forward Roll**

Figure 3 illustrates that there are no infit mean square values less than .77 or greater than 1.30, suggesting that the items are all elements of the same construct, i.e., the varying levels of quality performance of the forward roll. It is noteworthy that, if items did lie outside the boundary lines (pathway) they can not be interpreted meaningfully with regard to this pathway (Bond & Fox, 2001: 23). Sometimes the steps/item chosen by the researcher are satisfactory for some people but not others, e.g. the choice of items may be developmentally consistent with the majority of people. However, sometimes people are moving in a way that is outside the normal
pathway of items. Analysis from Quest confirms that this is not the case for the data utilised in this study.

Item difficulty

*Quest* software includes an item-person map in which person movement quality and item difficulty relations can be seen. Estimates of fit and error are tabulated along with movement quality and difficulty estimates. Item difficulty is expressed in terms of logits: zero is average; negative easier; and positive becoming more difficult. Person movement quality is estimated in relation to item difficulty estimates the higher the positive values the better the quality of movement (Bond & Fox, 2001: 33).

Figure 4 is provided as a further confirmation showing the hierarchy of item difficulty levels. In interpreting this map the lower levels of quality are indicated by the presence of the number 1 for each descriptor. For each increase in movement quality the number increases by one increment, i.e., step difficulties are ordered within each item.
Figure 4 Key:
- BAH Beginning Hand Position
- MHK Bridging Hip/knee
- ELM End Final Leg movements
- BHT Beginning Head contact
- MSA Bridging Shoulder/arm
- EFP End Foot placement
- BEA Beginning Arm/elbow
- ERM End Final rotational movements
- BCP Beginning Contact points

Figure 4: Item Person Fit Map

It is essential that the estimates of person movement quality and item difficulty in the data matrix are meaningful. This can only be achieved if each and every test item contributes to the measure of a single attribute (Bond & Fox, 2001: 25). The idea that the recorded performances are a reflection of a single underlying construct, that which the investigator creates to represent the items or observations is necessary for construct validity, i.e., not confusing two or more human attributes into one measure or score (Bond & Fox, 2001: 26).

Note that the measurement unit is common to both person movement quality and item, and is displayed down the far left hand side of the map. Because it is a logit scale, i.e., an interval scale, the equal distances up and down that scale have equal value. Persons and items are located on the map according to their movement quality and difficulty estimates, respectively (Bond & Fox, 2001: 45). The mean of item difficulties is adopted by default as the 0 point. The threshold is the representation of the item difficulty and is the movement quality level that is needed for a person to have a fifty percent chance of achieving success for the item (Adams & Khoo, 1993:86). The Xs down the centre-left of the map represent the distribution of case (subject) estimates on the logit scale. Each X represents the estimate for one subject (person), which signifies a fifty percent probability the subject will be able to achieve that item (descriptor) at the same position on the logit scale.

Conclusion

The model that emerged from the data represents an assessment instrument, which is sufficiently fine-grained to be applied across all cohorts. The output from Rasch analysis provides more than reasonable support for the single underlying construct; movement quality. There are three major implications for these results. Firstly, the new model for the assessment of the forward roll represents an instrument that more adequately applies across the ‘lifespan’ for differentiating the quality of a performance.

Secondly, with the advent of the new model, Rasch provides a level of confidence in the underlying construct, i.e., the varying levels of quality performance of the forward roll.

The Rasch findings reveal that both items and persons were shown to act in a predictable manner and it is pragmatic to postulate that the applicability of the new model is worthy of continued investigation in the future. Thus, based on the acceptable levels of the fit statistics, the case estimates could now be used to explore additional aspects of group performance, such as differences between the qualities of movement for each cohort, using multivariate techniques.
Thirdly, the descriptors can form the basis for coding the movement quality data from a SOLO perspective ... an examination of the SOLO cycles in the motor learning mode.

References


