Customized Learning Analytics: Six Prescriptive Steps

Elspeth McKAY
School of Business Information Technology and Logistics, RMIT University
Melbourne, Victoria 3000, Australia

Allaa BAREFAH
School of Business Information Technology and Logistics, RMIT University
Melbourne, Victoria 3000, Australia

ABSTRACT

For various stakeholders across educational institutions, there is a broad awareness of data analytics. The way learning analytics is defined involves: providing assessment reports for individual learners to know how they rate compared with other learners; highlighting students who may need extra support; assisting teachers to plan supporting interventions for individuals and groups of learners; aids for professional development teams when considering new course design and development; and institutional/ corporate marketing and recruitment management strategies. However, for some people the practice of customized learning analytics may seem a daunting task. Using a prescriptive Learning Analytics Planning model, this paper will show why this perception is wrong. It is vital to understand the importance of validating the measurement tools; these steps describe the key processes that are necessary to carry out customized learning analytics through careful preparation of the testing instruments.

Keywords: customized learning analytics, prescriptive learning analytics model, instructional design, Rasch modelling, item response theory (IRT), human-computer interaction (HCI).

1. INTRODUCTION

The term 'learning analytics' can invoke argumentative discourse. One prior definition evolved from the academic stance that described the process of predicting human knowledge acquisition via intelligent data which was analyzed through social connections [1]. This viewpoint was too narrow for others who felt that learning analytics should encompass a more holistic interpretation [2], and provide support for learners, their teachers and the facilitating institution to understand and predict needs for personal learning performances. Yet there are still other definitions that focus on particular aspects of environmental context such as using a reference model to determine a type of learning analytic that has six dimensions: data; environments; context; stakeholders; objectives; and methods [3] [4]. The one common point of agreement with most of these opinions is that the computational features of learning analytics ought to be linked directly to existing educational research [5].

And so, we can say for the purposes of this paper, that the way learning analytics is defined involves taking a customized approach, to include: assessment reports for individual learners to know how they rate compared with other learners; to highlight students who may need extra support; to assist teachers to plan supporting interventions for individuals and groups of learners; for professional development teams when considering new course design and development; and for institutional/corporate marketing and recruitment management strategies.

This frame of reference may be very rhetorical, as the practicing of such customized learning analytics may seem for some people a daunting task to the extent that progress will be slow, while the computational experts reach agreement with the researchers on best practice.

This paper describes a prescriptive learning analytics planning model in six steps. These key processes are necessary to carry out customized learning analytics practices through careful preparation of the testing instruments. It is therefore vital to understand the importance of validating the measurement tools.

2. LEARNING ANALYTICS PLANNING MODEL

This learning analytics planning model has been coined elsewhere as the 'six easy steps' [6], to initiate a hands-on conference workshop where the participants were given an overview of the data validation phase. The whole prescriptive process involves a well thought-out sequencing of the instructional design tasks that involve (1) preparing the test instrument; (2) setting the scoring regime; (3) validating the testing instrument; (4) modifying the test-items; (5) implementing the test; and (6) analyzing the results.

Step-1. Prepare test instrument

Sequencing the gradual skill-building progression of tasks for achieving any given instructional outcome, usually starts with identifying the easy concepts or declarative knowledge (knowing that), through the mid-range intellectual skills development [7], to achieving the procedural or cognitive strategies (knowing the how), as depicted in the following Gagné matrix [8].

		Instructional Objectives : Programming Knowledge					
		Declarative		Procedural			
		Band-A	Band-B	Band-C	Band-D	Band-E	
		Verbal information skill Concrete Concept Knows basic terms Knows ***"that"***	**Intellectual skill** Basic Rule Discriminates Understands concepts & principles	**Intellectual skill** Higher-Order-Rule Problem solving Applies concepts & principles to new situations	**Cognitive Strategy** Identify sub-tasks Recognizes unstated assumptions	**Cognitive Strategy** Knowing the ***"how"*** Recall simple prerequesite rules & concepts Integrates learning from different areas into a plan for solving a problem	
Task No:	**Learning Domain:**						**Totals:**
9	Solution algorithm						
8	Conditional logic						
7	Until logic characteristic						
6	While logic characteristic						
5	Repetition question						
4	Logic patterns						
3	Basic mathematics						
2	Programming process						
1	Defining diagram						
	Totals:						

Figure-1: Test specification matrix (McKay, 2000 p.175)

Step-2. Set scoring regime

Once the instructional pedagogy trajectory has been decided, using the test specification matrix, the next step is to set about writing appropriate questions and designing activities that test an individual's performance, based on the task-level and knowledge/skill bands as shown in Figure-1.

For example: Task-3 Band-A would require a question that simply asks whether the learner knows basic arithmetic operations such as: addition, subtraction without expecting any further knowledge of how these operations work, while the Task-3 Band-E would require the question to involve the learner in complicated mathematical operations such as manipulation of percentages or applying square roots in a formula.

At this point a 'marking scheme/rubric' is developed. All the answers to each question need to be written out with appropriate scoring to be lodged in a rubric with the ideal answer, involving: dichotomous (yes/no), multiple-choice, and partial credit (where the expected sequential steps to complete the task operation are written out, with appropriate scoring for each accumulated step).

Step-3. Validate testing instrument

Central to the proposed learning analytics planning model is the need to ensure, with respect to the item response theory (IRT) which underpins our data analysis [9], that our test-items fit the Rasch model [10]. To this end, we have used the QUEST interactive test analysis system devised in 1996 by Adams and Khoo, in Melbourne, Australia. QUEST is a psychometric test measurement application that affords the researcher improved analyses of an individual's performance relative to other participants [7], and relative to the test-item difficulty of the instructional content. Central to QUEST is the measurement model developed in 1960 by the Danish statistician George Rasch [11]. QUEST develops a uni-dimensional scale (with equal intervals along each axis), to measure individuals' performance (case) and test-items together. See an example of this in Figure-2 where each participant or 'case' is depicted by an '*x*,' on the left hand of the map, and test-items are shown on the right-side) [7]. The Rasch IRT estimates the probability of an individual making a certain response to a test-item.

When researchers work with this type of Rasch model their primary task is to ensure they produce data that fit the Rasch specification, rather than evaluate the fit in a more "conventional manner (ie., how well the model fits the data)" [9 p. 235]. This approach relies upon the uni-dimensionality of the Rasch model that provides a reliable and steady measurement within the model's unidimensional framework.

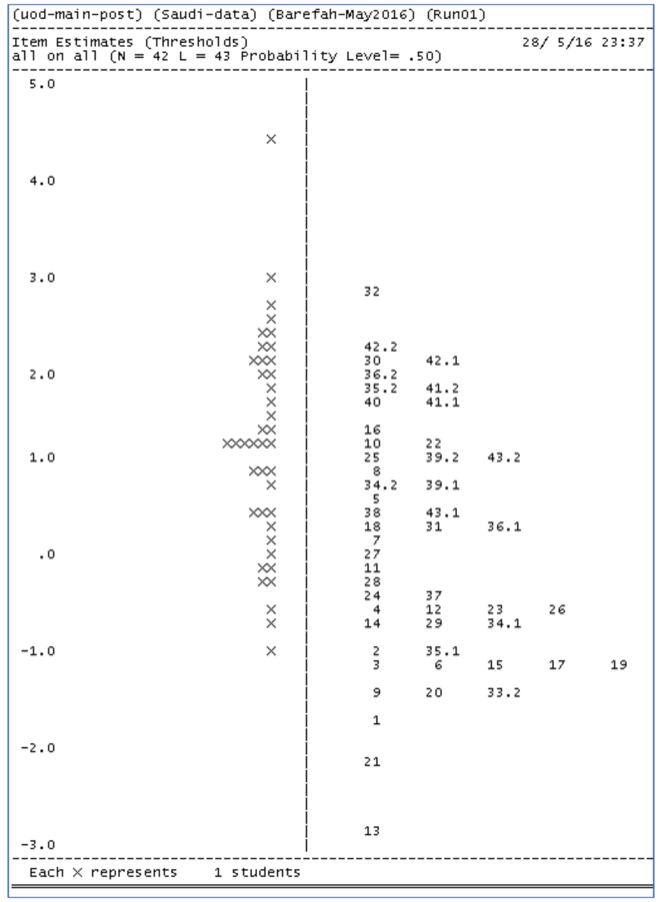

Figure-2: QUEST variable map

And so, we are concerned here with demonstrating that the test-item data points (see Figure-3) should lie within the threshold lines: if not they are either eliminated from the test instrument and/or are rewritten such that subsequent validation can ensure test-items are validated as a fit of the Rasch model, providing a reliable measurement scale and hence are worthy of further investigation. In Figure-3 test-items-39 and 41 are an under-fit; while test-items-2, -6, -27, and -38 would be considered as over-fit of the Rasch model.

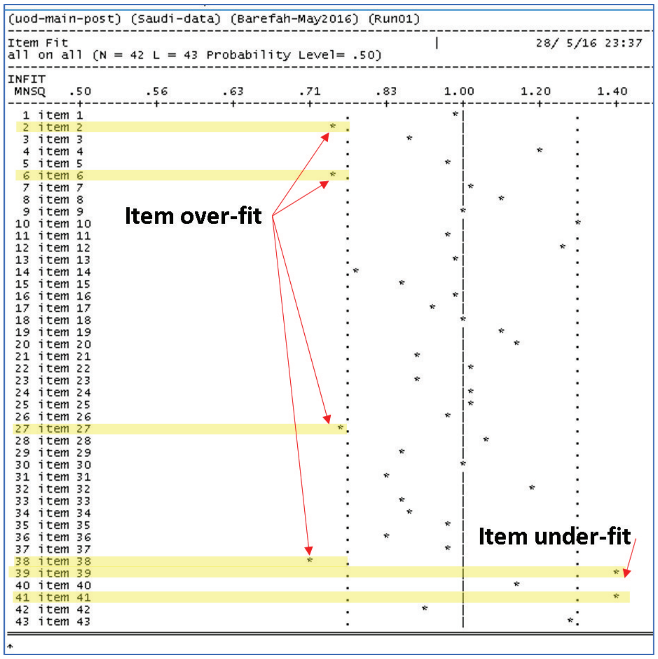

Figure-3: QUEST fit map

Step-4. Modify test-items

The QUEST software package provides the item-fit statistics [12]. The mean-square fit statistics provide a useful way of judging the compatibility of the model and the data [13]. The item thresholds represent the default representation of difficulty used by QUEST: analogous to Thurstonian thresholds, they represent the ability step-level that is required for an individual to have a 50 percent chance of passing that step [12]. The test-items to the right of the right-hand dotted (threshold) line in Figure-3, and mentioned in the previous paragraph, show more than expected variation from the model. Those test-items shown to be plotted on the left of the left-hand dotted (threshold) line in the diagram indicate less variation than anticipated. In their current question form, these non-fit test-items would need to be deleted from the data analysis. Should the researcher need to keep these test-items due to their position in the test specification matrix (Figure-1), they would need to be rewritten and re-tested to ensure they achieved a good Rasch model fit.

Step-5. Implement test-items

Once the instrumentation has been validated through a Rasch model analysis tool (there are a number of applications which perform this type of estimate), and depending upon the research design and methodology, these test-items can be given to participants as pre- and post-tests. For example: a pre-test can be used to identify a participant's prior domain knowledge before an instructional intervention is carried out. This means knowing what their level of skills/knowledge was before being given an instructional/training programme or tutorial class activity. Following the instructional treatment (training/tutorial module), the post-test can be conducted.

Test-items should be distributed throughout the test, spreading the difficulty throughout the test, with the easiest commencing skill/knowledge in the early part of the pre-test, and the most difficult skill/knowledge positioned last. This positioning is to ensure participants are not deterred by difficult items that may cause them not to try answering questions later on in the test. With Rasch modelling it is important for the participant/learners to provide their best answers, because the Rasch analysis does not cope well with missing data.

Step-6. Analyze results

The skill development outcomes can be evaluated in terms of the magnitude of change in learner/trainee proficiency (magnitude of effect size as defined by Cohen's Statistical Power Analysis (13). As the QUEST instigates a Rasch analysis (14) it is also possible to generate a set of hypotheses regarding the interactive dynamics of skill development with and without learning/training intervention. Because the Rasch IRT estimates the probability of an individual making a certain response to a test-item, the pre- and post-test results can therefore be analyzed using a test-item matrix that has each individual's responses for every test-item recorded. Common test-items (identically worded questions) should be 'anchored' so that scale scores on the pre-test are comparable with the scale scores on the post-test. The difference between pre-test and post-test scaled scores indicate whether learning/training has occurred, whether no learning/training occurred, or whether the instructional strategy resulted in reduced achievement

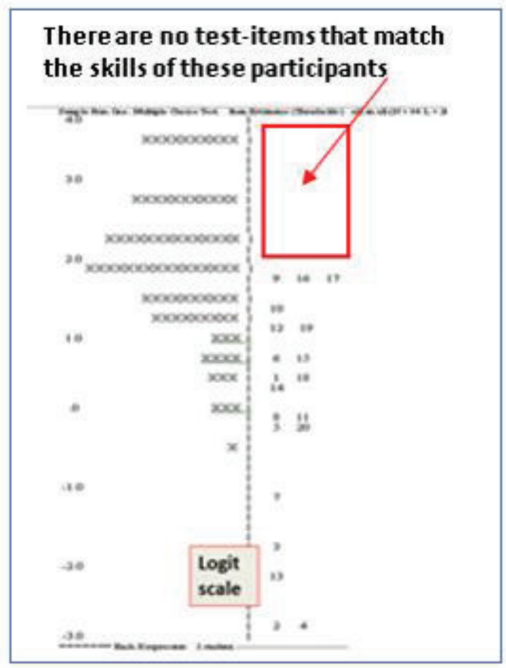

Figure-4: QUEST example of poorly designed instrument

Figure-5 shows a portion of the item analysis table that the QUEST application generates. This is a 'fine-grained' item analysis which involves a comprehensive table affording researchers with the best of both worlds. It shows both the classical test theory (CTT), and the IRT discrimination value. The CTT includes: the count (item achievement number) and achievement percentages; Pt-Biserial (correlation coefficient used when one variable is dichotomous); p-value (small number indicates strong evidence to reject null hypothesis); and the mean ability. A particular item achieving an IRT discrimination value closer to '1' indicates how well the item is able to distinguish between learner/trainees who are knowledgeable and those who are not. The summary section of this table reports on the mean test score, the standard deviation, and the Cronbach alpha index of internal consistency of the whole instrument, producing the correlation between item-score and overall score for each test-item.

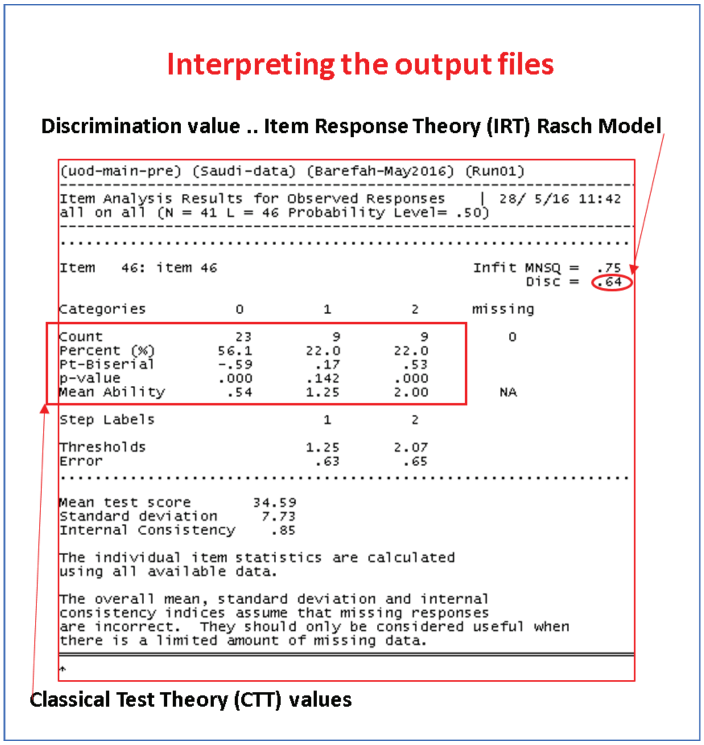

Figure-5: QUEST item analysis table

3. CONCLUSIONS

'Learning Analytics' means many things to many people, but it is not something that the average researcher should be afraid of. By systematically following the proposed Learning Analytics Planning Model, most research projects can achieve measurable and valid analysis results. Firstly, one must define the knowledge levels within the study domain, and then design a test instrument to capture participant skills. Then a structured scoring rubric must be developed to scale participant achievement on the test. The instrument must be validated, with poor test items (too easy, too hard, or poorly worded such that participants cannot understand what is expected) either removed or modified and re-validated. Once the instrument is fully validated then it must be implemented. This methodology recommends a pre-test to measure participants' prior domain knowledge, followed by the instructional treatment, and then a post-test to measure what they know after the treatment. The participant performance is then measured as the difference in QUEST scores between the pre-test and the post-test (after anchoring common test items). The internal consistency of the Rasch IRT model and the statistical variables generated by packages such as the QUEST Interactive Test Analysis System, allow the researcher to produce quantitative research results from limited data and fewer participants. The Cohen statistical power analysis can be used to determine the magnitude of effect of different treatments or different participant characteristics on performance outcomes (13).

4. REFERENCES

[1] G. Siemens, "What are Learning Analytics?" **Elearnspace**.2010 http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/.

[2] W. Greller & H. Drachsler, "Translating Learning into Numbers: Toward a Generic Framework for Learning Analytics", **Educational Technology and Society**, Vol. 15, No. 3, 2012, pp. 42–57.

[3] M.A. Chatti, A.L. Dyckhoff, U. Schroeder, & T.H. Hendrik, "A reference model for learning analytic". **International Journal of Technology Enhanced Learning (IJTEL),** Vol. 4, No. 5/6, 2012, pp. 318-331.

[4] M.A. Chatti, V. Lukarov, H. Thüs, A. Muslim, A.M.F. Yousef, U. Wahid, C. Greven, A. Chakrabarti, & U. Schroeder, "Learning Analytics: Challenges and Future Research Directions." **eleed**, 2014, Vol. 10.

[5] D. Gašević, S. Dawson, & G. Siemens, "Let's not forget: Learning analytics are about learning". **TechTrends**, Vol. 59, No. 1, 2015, pp. 64-71.

[6] E. McKay, "Learning Analytics - Six Easy Steps", in proceedings of the **E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2016**, V. Chesapeake, Editor. 2016, Association for the Advancement of Computing in Education (AACE). Alexandria, Washington DC, USA. pp. 1520.

[7] R.M. Gagné, **The Conditions of Learning: And the theory of instruction.** 4th ed. 1985, NY: Holt/Rinehart/Winston.

[8] E. McKay, "Instructional Strategies Integrating the Cognitive Style Construct: A Meta-Knowledge Processing Model (Contextual components that facilitate spatial/logical task performance)", in **Doctoral Dissertation (Computer Science and Information Systems)**. *Total fulfilment: Available online from* http://tux.lib.deakin.edu.au/adt-VDU/public/adt-VDU20061011.122556/.2000, Deakin Univ: Australia. p. 3 Volumes.

[9] T.G. Bond, & C.M. Fox, **Applying the Rasch Model: Fundamental Measurement in the Human Services** (3rd Edition). 2015, New York and London: Routledge, Taylor & Francis.

[10] G. Rasch, **Probabilistic Models for Some Intelligence and Attainment Tests. 1960**, Copenhagen: Nielsen & Lydiche.

[11] Abb ED:230.931 and (microfiche), *Reading, Comprehension, and Memory Processes*. 1983.

[12] R.J. Adams, & S-T. Khoo, **QUEST:The Interactive Test Analysis System**. Vol. 1. 1996, Melbourne: Australian Council for Educational Research.

[13] E. McKay, & J. Izard, "Planning effective HCI courseware design to enhance online education and training". in **The Proceedings of the 18th International Conference on Universal Access in Human-Computer Interaction.** Toronto, Canada, 2016, pp. 183-195: http://www.hci-international.org/.