The TIPS quality assurance framework for creating open educational resources: validation

URL	http://weko.wou.edu.my/?action=repository_uri& itemid=385

The TIPS Quality Assurance Framework for Creating Open Educational Resources: Validation

Paul Kawachi
Email: kawachi@open-ed.net

Abstract

This Paper presents the validation for the TIPS quality assurance Framework (see this pamphlet at http://www.open-ed.net/oer-quality/tips.pdf) for creating open educational resources (OER). A total of 205 criteria were elicited from more than 60 OER experts around the world and then referred to participants at several international workshops on quality assurance. From these workshops, 65 criteria were identified as essential, and these were then put together to form a four-level framework covering the teaching and learning aspects (T), information and material content (I), presentation product and format (P), and system technical and technological aspects (S) : giving the acronym TIPS. Here this TIPS Framework is validated in a Delphi -style referral back to OER experts to determine its content validity, according to Lawshe (1975). The Content Validity Ratio was calculated for each of the 65 items, and the resulting Content Validity Index was found to be above 0.80, after several items were rejected. The TIPS Framework was also referred to target end - users around the world for further validation. Most OER are authored by university faculty for reuse in universities, and relatively few are authored by school teachers for reuse in pre-tertiary education. The TIPS Framework is designed and intended for school teachers (and teachers not in schools) at the pre - primary, primary, secondary and vocational levels. Teachers were generally unfamiliar with OER initially, but nevertheless returned a high Construct Validity Index, and expressed their personal intention to try out the Framework in their teaching, to introduce e-learning technologies into their traditional practice. A second improved version of the TIPS Framework is published and available online as an OER in itself. While the intention is to help teachers, a final validation is now underway to examine the students to measure the learning improvements achieved from teachers using OER and the TIPS Framework.

Keywords: Quality assurance, Guidelines, Validation, Delphi, Teachers as OER authors

Introduction

Open educational resources (OER) offer an unprecedented opportunity for teachers in the developing world to develop learning materials for themselves and other teachers. OER have been defined variously since 2002 (for a review see Kawachi, 2013a), and here are defined (see **Box 1**) as free-of-cost, with an open-licence attached, allowing adapting or adding into other resources, and derivatives to be created, and at some time in digital format - following the UNESCO-CoL (2011) guidelines to facilitate reuse throughout the context continuum from highly-mediated face-to-face classrooms to independent learning at a distance.

An open educational resource (OER) is defined as a digital self-contained unit of self-assessable teaching with an explicit measurable learning objective, having an open-licence clearly attached to allow adapting, and generally being free-of-cost to reuse.

Box 1 Definition of OER

The present study explores the content validity of the original *TIPS Framework* through referring this back to OER experts for quality assurance. "Determining criteria for assessing quality in higher education requires an understanding of [the potentially] different conceptions of quality that inform the preferences of stakeholders" according to Harvey & Green (1993, p9), where there are five conceptions that can be distinguished *"as exception, as perfection, as fitness for purpose, as value for money and as transformative"*. The *TIPS Framework* offers criteria within three of these five dimensions: achieving fitness for purpose in the eyes of the reusers, achieving efficiency and effectiveness as free-of-cost resources to support education-for-all, and achieving transformation through imbuing 21st-century skills in their current and future students. Accordingly the perspectives held by stakeholder teachers in developing regions around the world as the target end-users are also surveyed to explore the content validity drawing from their 'potentially different' conceptions of quality.

A comprehensive instrument of all known criteria would have the best content and construct validity, and best reliability. However an unwieldy massive instrument would have low utility. Accordingly we reduce the overall number of items and merge items similar in purpose to improve the utility, at some cost to and with some loss in validity and reliability. The objective is to produce an instrument of high utility for practical use in the field, and with validity and reliability within tolerable levels.

'Quality assurance' can be described as a cross-sectional evaluation. 'Quality assurance' for OER is thus a checklist of aspects like the *TIPS Framework*. Beyond 'quality assurance', there is a need for 'quality improvement' (Kawachi, 2013b) that aims to improve standards. This can be achieved by adding a rubric alongside the checklist items in the *TIPS Framework*, for a user to tick off on a scale of five boxes to indicate how much the item was adopted. Then future ticking off can show mechanically the changes over time longitudinally to facilitate self-reflection on the quality aspects. Such a rubric is added to the revised *TIPS Framework*.

'Content validity' is a term with an imprecise meaning : according to Fitzpatrick (1983) 'content validity' can refer to (i) how well the items cover the whole field, (ii) how well the user's interpretations or responses to the items cover the whole field, (iii) the overall relevance of all the items, (iv) the overall relevance of the user's interpretations, (v) the clarity of the content domain definitions, and/ or (vi) the technical quality of each and all the items. The first two concern the adequacies of the sampling, and come under 'construct validity'.

Notwithstanding that 'content validity' is an imprecise term, it can be measured quantitatively by asking content experts to rank each item as (i) Essential, (ii) Not-essential but useful, or (iii) Not necessary. Those items ranked as not necessary are likely to be discarded. Among a large number N_E of experts, the number who rank the item as essential N_E is used to calculate the *Content Validity Ratio* for each item as shown in **Figure 1** below. This formula gives a *Ratio* of zero if only half the experts rank the item as essential, and if more than half the experts rank the item as essential then a positive *Ratio* between zero and one.

$$CVR = \frac{N_E - \frac{N}{2}}{\frac{N}{2}}$$

Figure 1 The Content Validity Ratio CVR (from Lawshe, 1975)

For relatively small groups of experts, the average *Ratio* for each item retained in the instrument should be close to one to decide the specific item has content validity with a probability of p<0.05. For larger groups of experts, the likelihood decreases that co-agreement as essential occurred by chance, and the *Ratio* value can be lower while still reaching a probability of p<0.05, with these values (corrected and extended from Lawshe, 1975) shown in **Table 1** below for various group sizes. Items obtaining minimum value, or above, are retained in the instrument. Then the average *Content Validity Ratio* over all items is termed the *Content Validity Index*. Generally the instrument should have an *Index* of 0.80 or above to be judged as having 'content validity'. Some outliers can be discarded on the basis of a low ranking by the experts, while others can be retained despite a low ranking provided there is some other procedure supporting their inclusion.

N of experts	5	6	7	8	9	10	11	12	13	14
Minimum CVR	.99	.99	.99	.75	.68	.62	.59	.56	.54	.51
N of experts	15	20	25	30	35	40	45	50	55	60
Minimum CVR	.49	.42	.37	.33	.31	.29	.27	.26	.26	.25

Table1 The Minimum Averaged Value CVR for an Item to be Retained at p<0.05

Methods

A list of 205 criteria have been collated covering all the five domains of learning, available at http://www.open-ed.net/oer-quality/criteria.pdf, and this is the most complete set of criteria to date available anywhere (see http://www.open-ed.net/oer-quality/others.pdf for all the other 18 known lists). Those 205 criteria were actively discussed in depth at regional and at international workshops, and reduced to 65 criteria. These 65 criteria (Kawachi, 2013) are here referred back to OER experts for content quality validation, according to Lawshe (1975) employing three options against each criterion asking;

Is this item 'Essential', 'Useful but not essential', or 'Not necessary' to the performance of creating a highest quality OER. The resulting analysis produces a minimum set of endorsed essential criteria.

In the event, three sets of survey were performed in parallel; (i) Set-1 of OER experts who were invited individually by personal email, (ii) Set-2 of OER online communities invited by posting a message to the respective group discussion forum, and (iii) Set-3 of school teachers at-the-chalkface who were each invited by personal email. A copy of the survey instrument is available at http://www.open-ed.net/oer-quality/survey.pdf. Those authors who presented a paper on OER quality to

the 2013 *Seventh Pan-Commonwealth Forum on Open Learning* PCF7 http://pcfpapers.colfinder.org were added to the OER experts list and also invited to respond. The OER experts and others were sent personal email over several weeks and were covered in groups at that time. There were 38 survey responses e1 ~ e38 in Set-1 recovered from the online survey website (by 9th May) and analysed offline one-by-one. Of these, 32 were usable. Reasons for discarding a response include incomplete return (the survey was cancelled after only a few items were ticked), and ticking the same column only both options which would otherwise distort the statistics and findings.

There is a need to be reassured about the sample population size. Therefore the 32 responses in Set-1 were re-ordered using standard statistical tables of random numbers, and three groups of ten each were examined using wave analysis (Leslie, 972) to increase the confidence in these being sufficient in quantity that they can be assumed to represent a wider population. Then the first ten (e1 ~ e10) were assigned as wave-1, the next ten (e11 ~ e20) as wave-2, and a third ten (e21 ~ e30) were assigned as wave-3. Two responses were thus unassigned, and indeed the survey is still open: however, wave analysis does not need more than the three waves. In a similar way, there were 17 survey responses in Set-2 recovered from the online survey website and analysed offline one-by-one. Of these, 13 were usable, and after re-sequencing in random order, the first ten (g1 ~ g10) were examined as one wave-4 and compared to the three waves of Set-1. Additionally there were 22 survey responses in Set-3 recovered, and 19 of these were usable. One was randomly chosen to be duplicated to make up two groups each of ten responses with (n1 ~ n10) as wave-5 and (n11 ~ n20) as wave-6 for wave analysis to explore whether a sufficient sample size has been collected.

Results

A pilot study was completed to test out the survey, and minor improvements were made. A planned fourth survey of online teachers communities was abandoned with no responses. However the three surveys on Set-1, Set-2, and Set-3 respectively were carried out, and all response data as of 9 May 2014 were analysed. The response rate was about 50% of those invited individually, and online conversations about comments were enjoyed with about 70% *eg* 27 of the 38 OER experts responding. The resulting population characteristics show acceptable gender equality m/f=37/26, a full age-range, and good target-relevant geographic coverage (Africa-9, Americas-5, Australasia-3, Europe-12, East-Asia-18, and South-Asia-15).

The four waves are presented in **Figure 2** below, and the observed close matches in pattern increase confidence in the data, allowing all responses to be analysed for *Content Validity Ratio* CVR_E for those criteria items indicated as being 'Essential'. All statistics and numerical data are available at http://www.open-ed.net/oer-quality/validation.pdf in the Full Report.
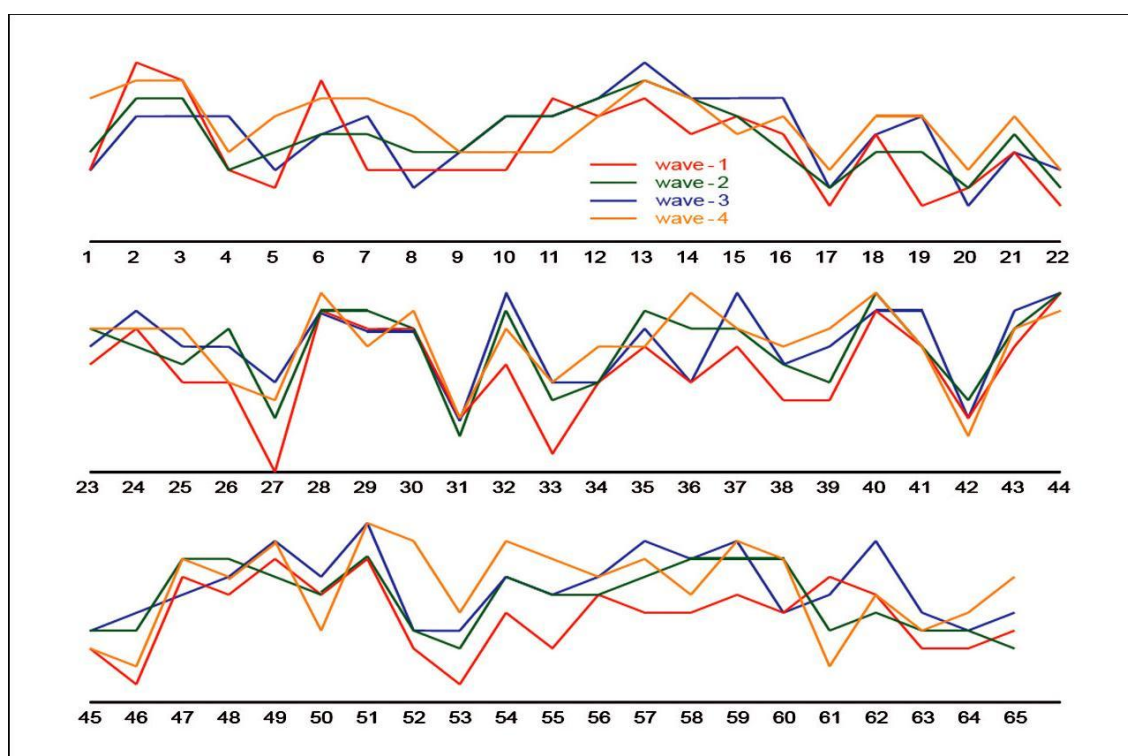
Figure 2 Wave analysis compared anonymous Set-2 responses as wave-4, with the three waves by individual OER experts confirming sufficient size of population sample

There were six criteria items that were indicated as 'Essential' by 27 or more individual OER experts among the 32 OER experts of Set-1 : these six together reach the *Content Validity Index* ≥ 0.80 threshold as valid at the probability level of $p < 0.05$, and these are given in **Table 2** below. Accepting that most respondents do not know that items scored as 'Useful' according to Lawshe (1975) are discarded, the analysis is re-performed using all the items scored as either 'Essential' or 'Useful' to give CVR_{E+U} . The CVR_{E+U} is high for each criterion, and the average CVR_{E+U} , over all the criteria items C-1 to C-65 without discarding any of the lower scoring items, gives the overall *Content Validity Index* CVI_{E+U} for the original instrument to be 0.86 which is > 0.80 and indicates the full original *TIPS Framework* is valid at $p < 0.05$.

From the 13 responses usable in Set-2, only the 18 highest CVR_E sustained the average at around 0.80 and are positioned into **Table 2**. These 18 criteria notably covered all those 6 criteria identified by Set-1. Moreover the average CVR_{E+U}, which is the *Content Validity Index* CVI_{E+U} for the instrument over all the criteria items C-1 to C-65 without discarding any lower scoring items, is 0.88 which is > 0.80 and indicates the *TIPS Framework* is valid at $p < 0.05$.

There were 19 responses usable in Set-3. One was randomly chosen to be duplicated to make up two waves for separately comparing, shown as wave-5 and wave-6 in **Figure 3**. The perspectives of the teachers are different from those of OER experts, and the waves here do not exactly match those in **Figure 2**. After wave analysis, the duplicate was removed, and the 19 responses analysed for content validity, giving an average CVR_E of 0.76 for all those 18 criteria so far in **Table 2**, and after C-36 and C-51 are removed this average increases to 0.79 comparable to that by Set-2 and by Set-1

of OER experts. However the teachers give highest CVR_E at 0.89 to another 6 items beyond those identified by the OER experts. These resulting 24 criteria are indicated as the new list of validated quality assurance criteria (**Table 2**).
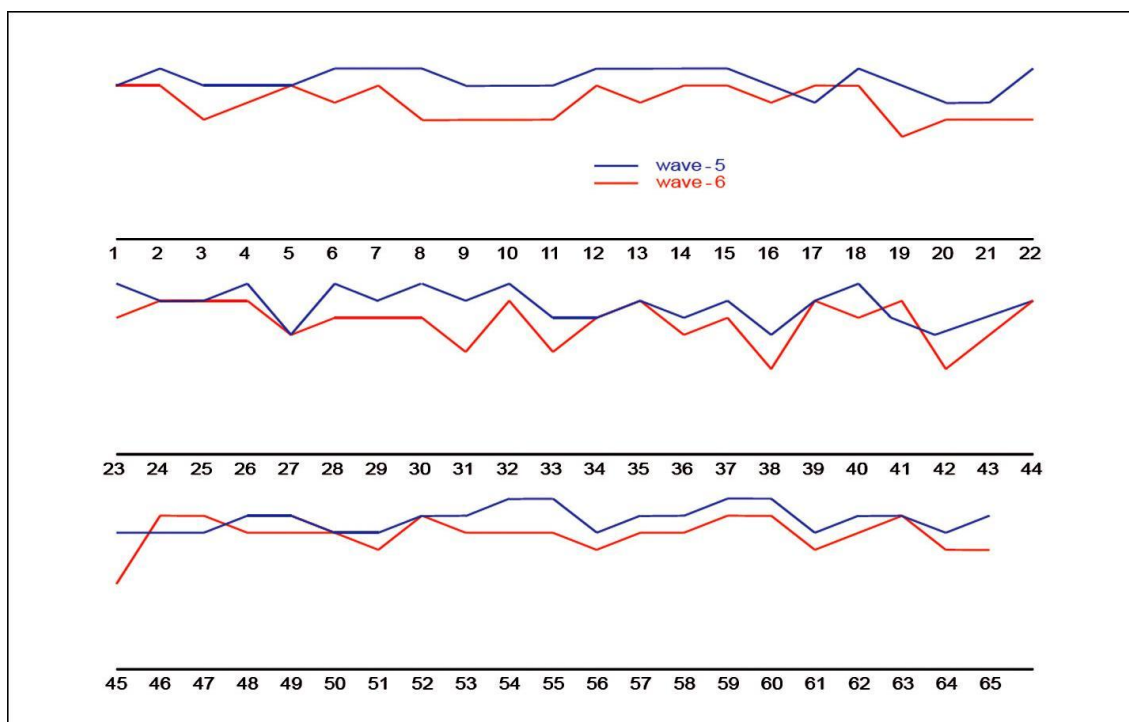
Figure 3 Wave analysis of wave-5 and wave-6 for Set-3 of individual teachers showing similar but not exact match to the wave patterns of OER experts

	Key criterion to be retained	Set-1 N=32	Set-2 N=13	Set-3 N-19
1	Consider giving a study guide for how to use your OER, with an advance organiser, and navigational aids		.69	.79
2	Use a learner-centred approach		.69	.89
3	Use up-to-date appropriate and authentic pedagogy		.69	.68
7	It should be aligned to local wants and needs, and anticipate the current and future needs of the student		.69	.89
12	You should adopt a gender-free and user-friendly conversational style in the active-voice			89
13	Don't use difficult or complex language, and do check the readability to ensure it is appropriate to age/level	.75	.85	.79
14	Include learning activities, which recycle new information and foster the skills of learning to learn			.89

15 Say why any task-work is needed, with real-world relevance to the student, keeping in mind the work needed to achieve the intended benefit			.89
18 Stimulate the intrinsic motivation to learn, eg through arousing curiosity with surprising anecdotes			.89
24 Provide a way for the student and other teachers to give you feedback and suggestions on how to improve		.69	.79
26 Try to offer learning support			.89
28 Make sure that the knowledge and skills you want the student to learn are up-to-date, accurate and reliable. Consider asking a subject-matter expert for advice	.81	1.0	.79
30 All your content should be relevant and appropriate to purpose. Avoid superfluous material and distractions		.85	.79
32 Your content should be authentic, internally consistent and appropriately localised		.69	.89
36 Add links to other materials to enrich your content		.85	.58
37 Be sure the open licence is clearly visible	.69	.69	.68
40 Ensure your OER is easy to access and engage	.88	1.0	.79
44 Present your material in a clear, concise, and coherent way, taking care with sound quality	1.0	.85	.79
51 Use open formats for delivery of OER to enable maximum reuse and re-mix	.69	1.0	.47
52 Consider suggesting which OER could come before your OER, and which OER could come afterwards in a learning pathway		.69	.79
54 Consider adding metadata tags about the content to help you and others later on to find your OER		.85	.79
55 Give metadata tags for expected study duration, for expected level of difficulty, format, and size		.69	.68
59 Your OER should be easily portable and transmissible, and you should be able to keep an off-line copy		.69	.89
60 Your OER and the student's work should be easily transmitted to the student's own e-portfolio			.89
Content Validity Index CVI = average CVR_E =	.80	.79	.79

Table 2 List of 24 Criteria according to each Set for CVI ≥ 0.80

Conclusions

A few OER experts voiced concern that some suggested criteria referred to non-OER and that a narrowest list of quality assurance criteria for OER could be produced. Reflecting on this point, the criteria specific to OER are highlighted in the revised *TIPS Framework*.

The teachers understandably hold a practical perspective that accommodates their want to preserve professional classroom autonomy. Their views as target stakeholders should be taken into account in formulating the *TIPS Framework*. Indeed Lawshe (1975) clearly explains that marginal items can be discarded or included within reason to serve the intended practical purposes. Accordingly those additional six criteria rated highest by the teachers are retained, and included into **Table 2**. These resulting 24 criteria in this list are recommended for including into a revised *TIPS Framework*.

Online discussions often in much detail with OER experts and teachers included other ways to improve the *TIPS Framework* to become more effective guidelines for school teachers. These ways included adding some examples on how to remix open licences, how to re-phrase some of the English used, and indeed examples or demonstration how to create OER and upload them to local repositories. Suggestions universally were supportive of this Project to develop guidelines for practising school teachers to create and author their own OER, and thereby to increase the OER author-base amongst teachers in developing regions around the world.

Acknowledgements

References

CoL (2011). *UNESCO-CoL Guidelines for Open Educational Resources (OER) in Higher Education: OER Workshop*, 25 May, Dar-es-Salaam. Retrieved April 25, 2014, from http://oerworkshop.weebly. com/uploads/4/1/3/4/4134458/2011.04.22.oer_guidelines_f or_higher_education.v2.pdf

Fitzpatrick, A.R. (1983). The meaning of content validity. *Applied Psychological Measurement*, 7 (1), 3 – 13. Retrieved February 21, 2014, from http://conservancy.umn.edu/bitstream/11299/101621/1/ v07n1p003.pdf

Kawachi, P. (2013a). *Quality assurance guidelines for open educational resources: TIPS framework*, version 1.0. New Delhi, India : Commonwealth Educational Media Centre for Asia (CEMCA). Retrieved March 7, 2014, from http://cemca.org.in/ckfinder/userfiles/files/OERQ_TIPS_978-81-88770-07-6.pdf

Kawachi, P. (2013b). Quality not static. *OER Quality*, blog posting 30th January. Retrieved March 7, 2014, from http://oerquality.wordpress.com/2013/01/30/quality-not-static/

Lawshe, C.H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28 (4), 563 – 575. Retrieved February 21, 2014, from http://www.bwgriffin.com/gsu/courses/edur9131/ content/Lawshe_content_valdity.pdf

Leslie, L.L. (1972). Are high response rates essential to valid surveys ? *Social Science Research*, 1, 323 – 334.