Refinement of the TELT survey instrument
Stage 3, iteration 2
M. Pachman, S. Quinton, A. Logunov
Introduction

Sub-layer 1 of Stage 3 of the TELT Evaluation Framework is aimed at assessing the effectiveness of a given technology for pedagogical use, by way of exploring its usability, usefulness and educational value (see A multi-definitional approach to TELT evaluation at UNSW for a comprehensive description of the sub-layers). The implementation of a staff and student survey instrument in this sub-layer allows for making certain recommendations in relation to trends in overall technology adoption and various tailored staff and student needs.

The evaluative process is designed to include an iterative component and the first iteration of the survey instrument started this process. The 2nd iteration is aimed at further refining the existing subscale and creating linkages between sub-layer 1 and sub-layers 2 and 3 by introducing relational elements (subscales, in our case).

As in the primary documentation, the instrument allows for a synthesis of contextual factors such as students’ technological background and their understanding of their own teaching and learning processes, with the affordances of a given technology. In turn, this allows for an insight into the students’ perceptions of educational technologies and their overall ideas about learning and teaching with the usability, flexibility and emotional affordances of a chosen technology in relation to the actual student population characteristics.

On the other hand, two of the limitations that we have worked with on the way to creating a valid and reliable sub-layer 1 survey instrument included the size of the survey – limited to 30 – 35 items total, and the voluntary participation criterion, which limits the available respondent sample size.

After the first survey revision in October 2010, we planned for a second revision to take place before S1, 2011, namely in March 2011. However, due to logistical problems and an insufficient number of survey responses in S2, 2010, we were prevented from revising the TELT survey instrument according to the original plan. Accordingly, we have used the data from S1, 2011 for the second revision.

Procedures

The participants involved staff and students from four different faculties (COFA, FASS, Science and Built Environment). Completion of the survey was voluntary. A total of 116 students and 7 staff members participated in the study. During the semester (Semester 1, 2011), participants completed an anonymous online survey during the final four weeks of classes. The survey remained available three weeks after the end of the semester.

After exploratory factor analysis (EFA) in the original scaling procedure (1st iteration), several scale items were re-written in order to improve their factor loadings and the reliability of the different subscales. These corrections could potentially
change the structure of the data and the interrelationships between the variables. That is why we had to use EFA again to ensure we have a correct factor structure.

Prior to conducting the EFA, descriptive statistics were generated to get an overview of the data, with acceptable means and standard deviations. Means were in the range of 2.95 to 5.48 on a 7-point Likert scale, and standard deviations were generally larger than 1.2, indicating that we had achieved acceptable variability in the way that participants responded. The correlation matrix was also closely examined for consistency of the survey items and to identify items that were too highly correlated \( r \geq .80 \) (Pett et al., 2003). Several of the redundancies found (between items 7 and 8, and items 12 and 13) were all with \( .80 < r < .82 \). There values are not considered a strong redundancy between the items.

EFA with an oblique rotation - Promax with Kaiser Normalisation \( (\kappa = 4) \) was conducted with the survey scale. Evaluation of the correlation matrix indicated that it was factorable. The Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy value was .872, which closely corresponds to “marvellous” \( (\geq .90) \) according to Kaiser’s criteria (Pett et al., 2003). The majority of MSAs values on the diagonal of the anti-image correlation matrix were in 80s \( (\geq .80) \), which corresponds to “meritorious” according to Kaiser’s (1974) criteria with some values, however, falling into “mediocre” \( (\geq .60) \) or “middling” \( (\geq .70) \) categories. The Bartlett’s Test of Sphericity \( (\chi^2 = 1839, df = 253, p < .000) \) was significant. This test showed that the data used in this study did not produce an identity matrix and thus were multivariate normal and acceptable for applying factor analysis.

The number of factors to extract was determined based on Eigenvalues \( (\text{Eigenvalue} > 1) \). We did not use the factor structure that was extracted and used in the 1st iteration because fine-tuning the existing items and adding some items could have changed the existing factor structure, and we intended to have a full picture of these changes. As a result, five factors were extracted that accounted for 69% of the total variance in the items which was an improvement over the original 60%.

We have considered factor loadings less than .40 as weak loadings (see Stevens, 2002), which resulted in weak loading for items #22 and #29 and marginally weak loading for #32 (see Table 1). Although we retained all of the above items, because all of them improved the reliability of their corresponding subscales (see Reliability section), the refinement process included re-wording these items or replacing them with similar items from the originally developed question pool. Overall, our survey instrument included the following factors1:

1. Usability (items 7, 8, 9, 10, 11, 12, 13, 14, 15, 16).
2. Feelings toward an application / frustrations toward eLearning applications – items 22, 27, 32.
3. Resistance to implied authorities in relation to one’s teaching and learning (items 20, 21, 23, 24, 25).

---

1 In this document the words factor and subscale are used interchangeably
Table 1: Pattern matrix generated from EFA solution with oblique rotation - Promax with Kaiser Normalisation ($\kappa = 4$)

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
<th>Factor 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. The way the information and resources are organised within the application easily integrates with the way I naturally organise or access my information and resources.</td>
<td></td>
<td>.978</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. The process of searching for, finding and accessing the information I needed was clear and intuitive</td>
<td></td>
<td>.877</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. The application allows me to perform my tasks quickly.</td>
<td></td>
<td>.942</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. In considering the usability and usefulness of the features and functions of this application, how strongly would you agree that you thought there were NO problems that you particularly identified?</td>
<td></td>
<td>.952</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. I think the application has a pleasant, appealing look and feel.</td>
<td></td>
<td>.827</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. The application integrates well into my learning or teaching approach and processes without me having to try look for alternative ways of integrating it and making use of it in my own individual way.</td>
<td></td>
<td>.875</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. I think other students/staff/users would find this application usable and useful.</td>
<td></td>
<td>.843</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Using this application unreasonably complicates my learning and/or teaching in this course(^a)</td>
<td></td>
<td>.527</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. It is somewhat frustrating to access information and download resources using this application(^a)</td>
<td></td>
<td>.639</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Based on your overall experience with the use of other software and websites, how would you rate this application</td>
<td></td>
<td>.904</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. I felt somewhat apprehensive about using the application(^a)</td>
<td></td>
<td></td>
<td></td>
<td>.652</td>
<td></td>
</tr>
<tr>
<td>21. I am not sure how this application, its features and its content are meant to integrate with my learning or teaching(^a)</td>
<td></td>
<td></td>
<td>.474</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. I do not believe that based on my opinions any serious improvements will be applied to the application(^a)</td>
<td></td>
<td></td>
<td></td>
<td>.800</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Reverse items
24. I can conclude that the institution does not understand my unique learning or teaching approach and learning or teaching processes. 

25. I should not be asked to modify my learning and/or teaching approach in order to work effectively with this technology application.

17. I thought about referring to help and documentation during my use of the application.

26. My use of the application was guided by training aids and technical documentation.

28. Some people say that eLearning applications are valuable tools that encourage the active collaboration between students and staff, exploration of different opinions and promote a culture of social learning. Based on your experiences with this application, would you agree?

29. Some people say that eLearning applications enable the effective acquisition of practical knowledge and skills that students can then apply in real life situations. Based on your experiences with this application, would you agree?

30. I believe I understand, and am well informed of the overall way this application, its features and its content should integrate with student learning and teaching.

31. Even if an application does not integrate well into my learning or teaching approach and learning or teaching processes, it is important that it is flexible enough so that I can integrate and use it in my own individual way. 

27. Some people say that eLearning applications are just tools that present existing information, and are of little value in student construction of new knowledge, collaboration and deep learning. Based on your experiences with this application, would you agree?

32. This application is not flexible enough to integrate with my learning and teaching approach.

22. The application makes me complete irrelevant, superfluous or distracting tasks that are not related to my learning (or teaching) processes.

**NB:** items in bold and italics are discussed in Validity section.
1. Flexibility, integration and collaboration in relation to the application (items 28, 29, 30, 31)
2. Use of technical aids and support documents (items 17, 26)

Since certain questions did not clearly and strongly load on one factor, we had to refine these items. This is discussed in the following sections of this document.

Reliability of the scale

Following the original administration of the survey instrument in S1, 2010 and the subsequent analysis we had several considerations regarding the reliability of the survey subscales. While Feelings toward an application and Usability subscales had acceptable (alpha = .76) to good (alpha = .89) reliabilities the other three subscales had low reliability coefficients.

The most evident part of the problem with two out of these three subscales was the small number of items per subscale: each of these subscales included only two items which is a low number of items per subscale according to measurement literature (Clark & Watson, 1995). In order to improve the reliability coefficients we have added an extra item to each of these subscales. We also replaced one of the items in the third low reliability subscale. The replaced item was not clearly worded (according to the participants’ comments), and was lowering the scale’s reliability. All these revisions positively influenced the reliability readings in S1, 2011 survey administration (see Table 2). Only in one case of adding an item to the subscale did not help increase the subscale reliability because the item did not load on the intended subscale. Thus, the reliability of Use of technical aids and support documents subscale turned out to be below the 0.70 level acceptable in social sciences (Netemeyer, et al., 2003; Nunnaly, 1978), and this subscale has only two items. This subscale covers an important aspect of technology evaluation and any technology being used by the novice inexperienced learners - namely, the existence and usefulness of the support documentation.

Table 2: Cronbach’s alphas reliabilities based on retained items (N = 123)

<table>
<thead>
<tr>
<th>Scale</th>
<th># items</th>
<th>Alpha</th>
<th>Means Inter-Item Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Usability evaluation</td>
<td>10</td>
<td>.94</td>
<td>.61</td>
</tr>
<tr>
<td>2. Feelings toward an application / frustrations toward eLearning applications</td>
<td>3</td>
<td>.72</td>
<td>.46</td>
</tr>
<tr>
<td>3. Resistance to implied authorities in relation to one’s teaching and learning</td>
<td>5</td>
<td>.79</td>
<td>.43</td>
</tr>
<tr>
<td>4. Flexibility, integration and collaboration in relation to the application</td>
<td>4</td>
<td>.70</td>
<td>.37</td>
</tr>
<tr>
<td>5. Use of technical aids and support documents</td>
<td>2</td>
<td>.66</td>
<td>.49</td>
</tr>
</tbody>
</table>
One more item was added to this subscale, and the subscale was retained as a part of the measurement instrument.

The question added to the *Use of technical aids and support documents* subscale expanded on the previous question covering the necessity of using the training aids. It read: “This application requires an extensive use of help functions, training aids and documentation”. The reason for including this question was to help distinguish the limits of support particular for each application.

The other measures used for reliability analysis:

1. We inspected the inter-item correlation matrix for possible highly correlated items. Correlations larger than .8 between two items suggest redundancy, and therefore the researcher should consider deleting one of the items. The only case of such a high correlation was present between items 7 and 8 (*Usability* subscale).
2. We inspected the mean and standard deviation of the inter-item correlations (Table 1). Ideally the average inter-item correlations should be in the range of .4 - .7 (Pett et. al, 2003). The only subscale having a mean inter-items correlation outside of this range is discussed in the Validity section of this document, as well as changes made to this subscale.

The proposed changes, i.e. adding a question to the *Use of technical aids and support documents* subscale and changing one item in the *Flexibility* subscale were intended to fine-tune the existing subscales. Otherwise, at this stage of the instrument development we could consider four out of five subscales fully and totally acceptable from the point of view of reliability with no further changes necessary to improve the coefficients.

**Validity considerations**

After the 1st iteration of the survey instrument, we had some new items being added to the survey subscales in order to improve their validity / reliability. This goal was mostly reached for both - the subscales’ reliability and validity. However, there are still some validity considerations in regard to 2 or 3 survey items. These considerations are outlined below.

Question # 22 existed in the original version of the survey and was loading on *Feelings toward an application* subscale (factor 2). Its loadings were not strong but were in an acceptable range, i.e. loadings of .40 to .51. With the introduction of the new items into *Feelings toward an application* subscale after the 1st iteration of the survey the loading of this particular item dropped to rather low .357. To strengthen the loading of this item we have proposed a refined version. The refined item reads: “This application often requires me to perform irrelevant steps that are not related to my course activities”.
Question # 29 was originally written after the first iteration of the survey to improve the reliability of the flexibility subscale (factor 4). While this goal is reached and flexibility subscale has a relatively high reliability (see Reliability of the scale subsection) the question item itself does not have a strong loading on a subscale. Besides, mean inter-items correlation coefficient for this subscale is relatively low and could be improved. All these reasons served as a basis for an item revision. The refined item reads: “This application aligns with my natural learning or teaching processes and enables the effective transfer of knowledge”.

While the previous two questions clearly had weak loadings on their relevant subscales, loadings of .40 and greater are often considered acceptable in a research literature (Stevens, 2002). Thus, question #32 was left unchanged for one variant of the survey. However, it was modified in the other variant of the survey in order to possibly increase its loading on its relevant subscale.

Question # 32 was originally written after the first iteration of the survey to improve the reliability of the flexibility subscale (factor 4). However, introducing new items into the existing instrument always has a potential danger to change the structure of the factors or subscales (like in our case). Question # 32 has acceptable loading on Feelings toward an application subscale (factor 2). This item could be left unchanged in its current format, or slightly modified to increase its loading from acceptable to good. The refined item reads: “This application is not flexible enough to promote social construction of knowledge”.

Discussion of a new factors structure

The original factors structure was mostly preserved with minor changes discussed below. Originally, Preconceived notions about eLearning subscale was rather complimentary to the demographic information we have already collected. It was not fundamentally improving our understanding of the survey participants or bringing any additional insights in regard to their individual features. The reliability of this subscale was also quite low (alpha = .52) and required some further enhancements. Thus, the new factor structure did not include the Preconceived notions subscale but included the other subscale estimating the Use of technical aids and support documents. The latter subscale allows for approximation of the initial support needs. These data serves as a pre-requisite for development of the further assessments and linkages between the sub-layers 1, 2 and 3 of the TELT evaluative framework (see TELT Evaluation Framework Stage 3 document). Currently, the Use of technical aids and support documents subscale has alpha = .66 reliability and consists of 2 items. The recommendations outlined in Reliability section of this document were related to enhancing this subscale with an additional item which is aimed at improving alpha coefficient to the acceptable ≥ .70 level.

The other four subscales stayed the same in terms of measuring the same psychological constructs, as described in the original refinement doc., such as:
1. Usability subscale measures the extent to which the given technology application is perceived to be usable.
2. Feelings toward an application subscale measures the emotional component of the learners’ perception of a given application and their possible frustrations toward eLearning applications.
3. Resistance to implied authorities in relation to one’s teaching and learning subscale measures the extent to which students differentiate between their own learning style, and the institutional regimen and affordances in relation to a particular technology.
4. And finally, flexibility, integration and collaboration in relation to the application subscale measures how flexible the given application is and how much the application can be integrated with and adjusted for individual preferences.

All the above subscales have acceptable to perfect reliability and can be used in further learner-centred assessments of educational technologies.

**Summary**
This iterative refinement cycle helped to correct the remaining reliability and validity issues of the survey instrument and aided to further finalize the instrument design. One additional iterative cycle might be necessary to confirm the design but the current reliability of the subscales has reached an acceptably high level.

While contextual factors are known to have a critical influence on implementation, adoption, and future use of educational technologies at the higher education level, a vast majority of research studies assess learner satisfaction alone, choosing to overlook who the learners are and overlooking their own technological background. As a result, these accounts are somewhat one-sided, emphasising the “good” or “bad” technology rather than the pedagogical and technological factors. Our survey instrument, taking into consideration learner background and their conceptions of teaching and learning, thus brings an important addition to the educational technologies research field. Furthermore, the current shortage of valid and reliable instruments for evaluating the actual effectiveness of educational technologies, calls for the creation of a coherent set of measurement instruments allowing for assessing and improving already-adopted educational technologies, bringing new technologies on board, and exploring course design issues related to a particular technology and learners with a particular technical background. A development of this survey instrument and the inclusion of contextual factors in this refined scale may contribute greater value for researchers and evaluators in accomplishing these goals.

The practical significance of the 2nd iteration is that it is another step on the way towards creating a valid, reliable and effective evaluative scale that could help institutions to optimise their decision-making processes in selecting a range of educational technologies appropriate to their staff and student needs. It is also a
step on the way of linking the sub-layer 1 of the TELT Evaluation Framework with the subsequent sub-layers and with the needs of the overarching evaluation report.
References:


