The relationship between longitudinal performance and competence: a pilot study

J.Dickie[1], V. Bissell[1], L. Dawson[2], M. McEwan[1]
[1]University of Glasgow, [2]University of Liverpool

AIMS

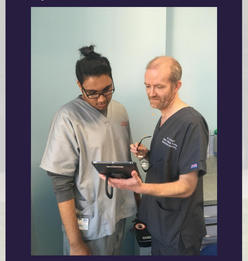
LIFTUPP© is an e-system for providing longitudinal feedback to students on their clinical performance. Accumulated data can also be used as a measure of progress. The utility of the data for determining competence has not been formally demonstrated. The aims of this pilot study were:

- Compare longitudinal data on clinical performance with a simulated one-off competence test and the subjective opinion of teaching faculty.
- Generate lines of enquiry for further research on assessment of dental student competence.

METHODS

13 volunteer BDS3&4 students

Longitudinal clinical performance data

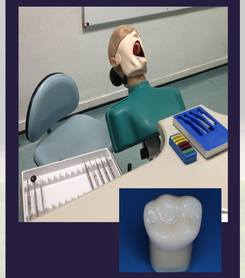


> Students received developmental indicators on a 1-6 scale for clinical procedures in a variety of real-life clinical settings.

> Data on *caries removal and direct restoration* was retrieved from the databased. A n indicator of 4 was set as the threshold for competence.

> Consistency score = number of occasions on which the overall mark for the procedure was ≥ 4 ÷ total number of procedures completed.

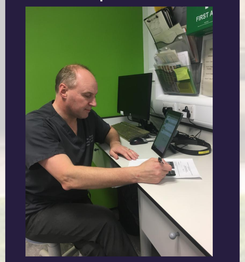
Simulated competence test



> Student participants removed and restored simulated caries from laminated plastic teeth under test conditions.

> Each tooth assessed by 10 Restorative teaching faculty using a criterion-based marking scheme.

Faculty subjective opinion



> Faculty shown photographic profiles of student participants. Two faculty per student.

> Faculty asked to complete a 7-point Likert scale questionnaire gauging confidence in students' ability to restore both occlusal and interproximal caries.

> Faculty asked binary question on whether they think the student would be "*competent*" or "*not yet competent*" in restoring 1) occlusal caries and 2) interproximal caries.

RESULTS

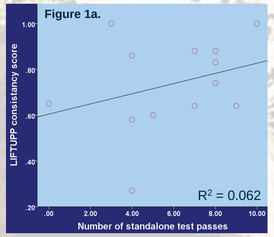
LIFTUPP© Data vs. Simulated Competence Test vs. Faculty Subjective Opinions

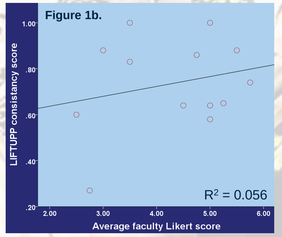
Student	LIFTUPP consistency score	Number of simulated competence test passes (out of 10)	Faculty subjective opinion	
			Average Likert score	Number of competence awards (out of 4)
1	0.64	7	5	4
2	1.00	3	3.5	2
3	0.65	0	5.25	3
4	0.64	9	4.5	2
5	0.88	7	3	0
6	0.60	5	2.5	1
7	0.88	8	5.5	4
8	0.74	8	5.75	4
9	0.27	4	2.75	0
10	0.83	8	3.5	2
11	0.86	4	4.75	3
12	0.58	4	5	2
13	1.00	10	5	3

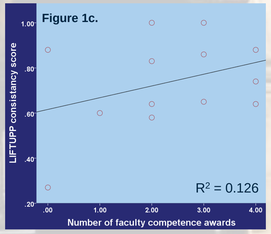
Table 1. Comparison of outcomes from all three data sets

Comparison between all three of the study's datasets (**Table 1**) appears to show no general correlation between student LIFTUPP© data and more traditional means of competence assessment (standalone competence tests and faculty opinion). Both traditional assessments demonstrated a degree of inconsistency when attempts to gain a consensus of student competence were made.

Quantitative correlation testing (**Figures 1a-c** and **Table 2**), based on regression analysis and calculation of Spearman's rank correlation coefficient, also found a lack of association between LIFTUPP© data and the traditional competence assessment methods. The lack of correlation is signified by the low R^2 values and widespread data points from the scatter charts, as well as the low values of the Spearman's rank coefficient. Low significance was anticipated because of the study's small sample size.







Figures 1a-c. Regression analyses using scatter plot diagrams for comparisons between a) LIFTUPP© data and the simulated competence test, b) LIFTUPP© data and faculty Likert scores, and c) LIFTUPP© data and the number of competence awards given by faculty.

	LIFTUPP© consistency scores *vs* simulated standalone competence test results	LIFTUPP© consistency scores *vs* average faculty Likert score	LIFTUPP© consistency scores *vs* number of faculty binary competence awards
Spearman's rank correlation coefficient	.137 (very weak)	.199 (very weak)	.279 (weak)
Significance (p-value)	0.655	0.515	0.356

Table 2. Spearman's rank correlation coefficients for comparisons between LIFTUPP© data and i) the results from the simulated competence test and ii) faculty subjective option. NOTES: all p-values >0.05. Spearman's rank strengths: .00 - .19 = "very weak"; .20 - .39 = "weak"; .40 - .59 = "moderate"; .60 - .79 = "strong"; .80 – 1.0 = "very strong".

CONCLUSIONS

Competence is now widely seen to be neither permanent nor independent of context. The lack of correlation seen in this study is therefore not surprising and, indeed, accepting the small sample size, adds further to concerns about one-off competence assessments and subjective opinion. Much recent scholarship on the subject suggests that longitudinal data reflecting performance situated in the workplace is required for trustworthy summative decisions. **Figures 2a-d.** represent examples of patterns of student performance identified through production of barcode graphics, derived from the actual LIFTUPP © data of student participants in this study. A dark line indicates an incidence where the student did not achieve the "competence threshold" for provision of a direct restoration. Lighter areas signify occasions where that threshold was attained.

Figure 2a. Barcode of a student who was inconsistent to begin with, but appears to have improved with further experience and become consistent.

Figure 2b. Barcode of a student who has been inconsistent throughout their clinical practice, who may require further development.

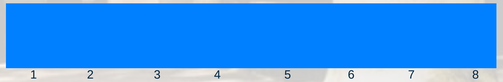


Figure 2c. Barcode of a student who has been consistent from the outset (though at this point, only 9 restorations have been completed).

Figure 2d. Barcode of a slowly developing student, who may require both intervention, and significantly more time to develop

LIFTUPP © enables patterns of performance over time to be identified, which may provide a rich contribution to decision making that is based more on qualitative than psychometric approaches, and is the province of the expert panel or "*interpretive community*". Future research should aim to understand the patterns of performance associated with effective and safe practice, with the aim of improving both learner development and the quality of patient care.

References

1. Albino, J.E., Young, S.K., Neumann, L.M., Kramer, G.A., Andrieu, S.C., Henson, L., Horn, B, and Hendricson, W.D., (2008). Assessing dental students' competence: best practice recommendations in the performance assessment literature and investigation of current practices in predoctoral dental education. *Journal of Dental Education*, 72(12), pp.1405-1435.
2. Dawson, L.J., Mason, B.G., Bissell, V.. and Youngson, C., (2016). Calling for a re-evaluation of the data required to credibly demonstrate a dental student is safe and ready to practise. *European Journal of Dental Education,* March 2016 (Epub ahead of print).
3. Chambers, D.W., (1998). Competency-based dental education in context. *European Journal of Dental Education*, 2(1), pp.8-13.
4.Govaerts M., van der Vleuten C., (2013). Validity in work-based assessment: expanding our horizons. *Medical Education*, 47, pp.1164-1174.