Integrating technologies into “authentic” assessment design: an affordances approach
Richard Osborne, Elisabeth Dunne and Paul Farrand
University of Exeter, UK

Current pressures in higher education around student employability are driving new initiatives for change. Assessment is also a topic of debate, as it is a key driver of student behaviour, yet often falls behind other metrics in national surveys. In addition, increasing focus on digital literacies is catalysing new appreciations of what emerging digital culture might mean for both students and staff. These three highly topical challenges were jointly explored by the University of Exeter’s Collaborate project, which aimed to create employability-focused assessments enhanced by technology. By combining existing research on assessment with grounded data derived from local stakeholders, the project has developed a model for assessment design which embeds employability directly into the curriculum. Digital technologies have been aligned with this model using a “top trump” metaphor, where key affordances of technologies are highlighted in the context of the model. This paper explores the design-based research approach taken to develop this model and associated “top trumps”, along with results from the first practical iteration. Results suggest that the model is effective in supporting the design of an “authentic” assessment and that a targeted affordances approach can support the alignment of specific technologies with a particular pedagogic design.

Keywords: employability; assessment; authentic; affordance; evaluation

Introduction
This paper documents the preliminary findings from the University of Exeter’s Collaborate project, which aims to bring together staff, students and employers to create employability-focused assessments enhanced by technology.

Employability: a driver for change
Student employability is currently something of a hot topic in Higher Education, and perhaps for good reason given the importance placed on employability by those thinking of attending university. Recent surveys have highlighted it as among the foremost reasons for considering a university education (Pollard et al. 2008), while over 70% of students stated that improving job opportunities was the most important reason to go to university (CBI/NUS 2011). One possible way for universities to help improve student employability is to “close the gap” between higher education...
and business (Wilson 2012). Exploring ways to embed the development of skills within students’ learning experience has been highlighted as having particular potential:

Universities should reflect on the opportunities that are provided for students to develop employability skills through the formal learning methodologies used within the university.

(Wilson 2012, p. 10)

This sentiment is at the heart of the Collaborate project, embedding employability into the curriculum by focusing on assessment as a vehicle for change.

Assessment: a vehicle for change

Assessment is widely acknowledged as a key driver of student behaviour. As Gibbs and Simpson (2004) summarises, what students attend to, how much work they do and how they go about their studying are completely dominated by how they perceive the demands of assessment. Assessment can therefore be an ideal target for educational innovations. Traditional forms of assessment such as exams and essays, however, are not well suited for embedding employability as they tend to be used infrequently in professional contexts. Other forms of assessment exist which may be better suited to this task, in particular authentic assessment.

The concept of an authentic assessment appears to originate in the US K-12 school system (Terwilliger 1997), as a reaction against the use of multiple choice questions as the core or only way of assessing student performance. The use of the word authentic has proved challenging, however, as Terwilliger summarises: “the term inappropriately implies that some assessment approaches are superior to others because they employ tasks that are more ‘genuine’ or ‘real’” (1997, p. 27). Although authentic assessment appears to be increasingly used in further and higher education, the term is still not widely understood (Falchikov 2005), partly perhaps due to multiple perspectives and definitions (e.g. Gulikers, Bastiaens, and Kirschner 2004; Mueller 2005; Wiggins 1993). Indeed, surveys by Whitelock and Cross (2011) at the Open University suggest that only 25% of the academics surveyed had heard of the terms “authentic learning” and “authentic assessment”.

In order to bring more clarity to this area, Whitelock and Cross (2011) analysed literature from key writers in the field to summarise the common features of an authentic assessment, which included:

- A range of assessment tasks rather than just the “traditional” ones
- Collaboration that is similar to that experienced by practitioners or experts in the field
- Problem tasks that are like those encountered by practitioners or experts in the field
- Resources taken specifically from real-world case studies or research
- Simulations of role-play or scenarios.

(Whitelock and Cross 2011, pp. 3–4)

At Exeter, this summary was combined with research from Herrington (2006), Karet and Hubbell (2003) and local stakeholders, in order to develop a new “dimensions” model for the design of authentic assessments.
Technology: a supporter of change

The value of constructive alignment (Biggs 1996) between intended learning outcomes, teaching methods and assessment is widely acknowledged (Gulikers, Bastiaens, and Kirschner 2004), yet this same constructive alignment is rarely seen in the application of technologies to learning scenarios. Technology use is widespread in higher education, yet it is often “treated as ‘natural’ or given” using a common sense approach (Oliver 2013). Oliver contends that this may in part be due to a lack of theory to explain technology in an educational context, and suggests that one possible approach is to use the theory of affordances (Gibson 1979). Although Oliver argues that affordances may in fact be too problematic to be of practical use, others suggest that affordances can be a useful framework for considering how technologies might be used to support learning and teaching (e.g. Bower 2008; Conole and Dyke 2004; Hammond 2010).

Attempts to align technologies directly with pedagogy may be impractical, as individuals tend to find ways of using technology in unplanned ways to suit their needs. The notion of interpretive flexibility, as described by Pinch and Bijker (1984), suggests that technological artefacts are interpreted by the society that uses them, and tend to be used in flexible ways. Indeed Squires and McDougall (1994) note that in educational contexts technologies can be most effective when teachers and students use them in ways that the original designers had not intended. The theory of affordances suggests ways of addressing this issue of flexibility and has two other characteristics that may allow technologies to be aligned with pedagogy using an affordances approach:

- Any one object/space has multiple affordances, and the affordance that is attended to is based on need. “Needs control the perception of affordances (selective attention) and also initiate acts” (Gibson, Reed, and Jones 1982, p. 411) (e.g. a book can be read, thrown, burnt, etc., but only the affordance that matches need will be attended to). This flexibility of affordances should allow alignment with specific pedagogic requirements.
- An affordance is stable, and it “does not change as the need of the observer changes” (Gibson 1979, pp. 138–139) and exists whether or not it is attended to (e.g. for someone who knows how to read, a book affords reading whether or not they have the intention to read it). This stability of affordances should allow multiple technologies to be designed into a pedagogic model.
- Affordances are personal and relational; that is, a specific affordance is relative to a specific individual (e.g. all books may afford reading for a literate adult, but the philosophy of John Dewey may only afford comprehension to far fewer). This relational nature of affordances should allow individuals to select from multiple technologies based on personal needs and experience.

Although the theory of affordances may have potential for thinking about how technologies might be aligned to a specific pedagogic model, it does not help in suggesting which technologies might be useful. Indeed, the issue of which technologies are appropriate in education is something of a moving target. As the digital world matures, traditional models of technology use in education, with a central IT department responsible for the delivery of learning technologies, are being challenged by a “Bring Your Own Device/Service” approach, a cultural shift...
where “off the shelf” technologies are used in increasingly hybrid ways to support studying. There is evidence that students feel central technology provision has fallen behind what they can provide themselves and that “the type of technology used in HE is increasingly outdated” (NUS 2010). For this reason, together with the overall focus on employability, the Collaborate project has focused on how contemporary “off-the-shelf” technologies might be aligned with the new “dimensions” model.

Method
A Design-Based Research (DBR) methodology was adopted to explore how assessments might be adapted to include features associated with authentic assessments, supported by contemporary technologies. DBR is an emerging methodology, with no strict definition, though researchers have identified a number of principles. Reeves (2006), for example, outlines three principles:

... addressing complex problems in real contexts in collaboration with practitioners; integrating known and hypothetical design principles with technological advances to render plausible solutions to these complex problems; and conducting rigorous and reflective inquiry to test and refine innovative learning environments as well as to define new design principles.

(Reeves 2006, p. 58)

Anderson and Shattuck (2012) suggested the following five characteristics were hallmarks of a DBR approach:

- Using mixed methods
- Involving multiple iterations
- Involving a collaborative partnership between researchers and practitioners
- Evolution of design principles
- Practical impact on practice.

The Collaborate project matches well with these principles/characteristics of DBR, in that it is:

- Working in real contexts with practising academics and employers
- Using multiple iterations to test ideas in varying contexts
- Adapting and changing designs based on experiences, evaluations and feedback
- Blending in contemporary technologies as both management tools and design tools.

Creating a model for assessment design
The challenge for the project was to create a model to support the redesign of assessments, embedding the development of skills and competencies into the module. Various methods were explored, beginning with a set of continuums based on a model of “tensions and challenges” relating to assessment (Price et al. 2011). This initial design evolved through testing with collaborative partners, and eventually words were chosen to summarise the key messages from both literature and local research concerning authentic assessment, which then defined new continuums (see Figures 1 & 2). A key feature of these continuums was that they were designed to be used and filled in rather than simply read.
Work-integrated Assessment: Change Continuum Tool

The concept of a work-integrated assessment is an assessment where the tasks involved more closely resemble those which you would be carrying out in employment. This change continuum tool is designed to scaffold thinking by asking the questions:

- At what position does your current assessment fall on these continuum’s?
- Where would you like your assessment to be placed on these continuum’s?
- How could your current assessment be changed to move in this direction?
- How could the introduction of technology help shift this position?

Assessment:

Example:

Data

- Theoretical, ‘perfect’ data
- Now using data from past modules
- Propose using data from http://data.nasa.gov/
- Real world, messy data

Problem

- Internally derived, theoretical idea
- Externally focused, real world task

Timing

- Single, long term
- Multiple, short term

Data

- Theoretical, ‘perfect’ data
- Real world, messy data

Purely academic learning might require a theoretical problem in order to test a theoretical understanding. In employment though problems tend to be very real.

Assessments are often delivered in the form of examinations at the end of learning. In employment however timing is often out of your individual control, and you usually have to juggle multiple tasks at the same time.

The individual pieces of data that you need to work with in employment rarely come in coherent, standard forms. They are usually in ‘messier’ formats that need to be interpreted to be of use.
Discussions with academic colleagues showed that the use of these linear continuums with two extremes tended to lead to positive and negative perspectives, isolating and reinforcing existing beliefs and polarising opinion. Any attempt to cast
current academic practice near one end and authentic practice nearer the other suggested conflict between the two positions. A new model was sought which would reduce this tendency, with a more “value neutral” position that would allow for the exploration of assessments without negative connotations.

The project team decided to adopt a radar chart for the model, which has no right and wrong ends and therefore does not suggest positive and negative positions. The dimensions shown in Figures 1 & 2 were collapsed from seven to six, with the problem and data dimensions combined, and arranged in a hexagon shape. It was also decided to move away from the term authentic assessment, coining instead the phrase “work-integrated assessment”, due to concerns over the meaning of the word authentic. We defined work-integrated assessment as an assessment where the tasks and conditions are more closely aligned to those experienced within employment. The resulting six dimensions of work-integrated assessment are listed and described in Table 1, and their hexagonal arrangement is shown in Figure 3.

The dimensions model was printed on A3 sheets, and used by collaborative design teams of academic staff, together with employers, students and other staff as appropriate, to help change existing assessments. The model is both a thinking tool, using the dimensions to stimulate discussion, and a redesign tool, annotated in order to visualise and agree collective thinking. It is used in three stages, each of which is marked on the chart:

(1) Analysis: Thinking about a current assessment, plotting on the chart where it might lie on each of the dimensions.

(2) Design: Discussing how to move along the dimensions; that is, how changes to the assessment might be introduced and what impact this might have.

(3) Evaluation: Determining how well the designed changes have worked in practice.

The analysis and re-design stages are completed at the same time, the evaluation stage after the module has finished. Figure 4 is a completed model showing the analysis and design stages.

**Aligning technologies to the model**

The purpose of the current project is not only to create a model to help design work-integrated assessments, but also to find a way to integrate contemporary technologies

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Brief guidance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem/data</td>
<td>Set a real world problem as the core assessment task, supported by real world data</td>
</tr>
<tr>
<td>Time</td>
<td>Move to a more distributed pattern of assessment; consider introducing “surprise” points</td>
</tr>
<tr>
<td>Collaboration</td>
<td>Create teams of students who work together to complete the assessment, encourage collaboration</td>
</tr>
<tr>
<td>Review</td>
<td>Include peer and/or self-review explicitly in the assessment process</td>
</tr>
<tr>
<td>Structure</td>
<td>Lightly structure the overall assessment; reward student approaches</td>
</tr>
<tr>
<td>Audience</td>
<td>Aim to set explicit audiences for each assessment point</td>
</tr>
</tbody>
</table>
Designing a Work Integrated Assessment

The concept of a work-integrated assessment is an assessment where the tasks and conditions are more closely aligned to what you would experience within employment.

Module: __________________________
Co-ordinator: ______________________
Year: _____________________________

Peer / Self Review
Include: peer and/or self review explicitly in the assessment process
Typically the review of assessment (e.g. feedback) in formal education is only provided by teaching staff. In employment, however, much of the review process comes in multiple forms, e.g. informal peer feedback from colleagues, formal and informal reviews from clients, and self-reviews of personal performance. Including peer and self review explicitly within an assessment helps students to develop critical thinking skills, and encourages articulation and understanding.

Multiple Assessment Points
Move to a more distributed pattern of assessment, consider introducing ‘surprise’ points
Assessments are often delivered in the form of one summative assessment, e.g. an exam or essay, at the end of a period of formal learning. In employment however, assessment or evaluation points tend to occur frequently. In addition, timing is often out of individual control, and consequently it can be necessary to judge competing tasks at short notice.

Using multiple assessment points helps to develop reflective thinking, whilst ‘surprise’ points support task prioritisation.

Figure 3. A blank version of the dimensions model.
Figure 4. A completed version of the dimensions model.
into this model. A method was needed which would identify the affordances of “off the shelf” technologies in relation to the model’s six dimensions.

The “Top Trumps” card game, popular in many parts of the world, seemed to offer a format ideally suited to our needs. “Top Trumps” show a series of related objects, listing their strengths and weaknesses against a series of fixed measures. By adopting this “top trump” metaphor, and setting the measures as our six dimensions, ratings could then be assigned for each technology summarising what it “provides or furnishes” (Gibson 1979) from an affordances perspective for each dimension.

A large number of potential technologies were identified for Collaborate, using data from previous technology training sessions and local expertise, and these were then narrowed down to a collection of approximately 60 of the most prominent and well-known technologies, such as Blogger, Facebook and Screenr. Small focus groups of staff then allocated up to five “stars” for each of the dimensions as listed in Table 1, by asking questions such as “what does this provide or furnish for …”, with different questions written to prompt thinking for each of the six dimensions. Blogger, for example, scored highly for Time and Review, as it uses a diary-based approach and has extensive commenting and feedback facilities. It scored lower for Collaboration and Problem/Data, however, as it has limited team-working facilities and does not inherently provide any data or real-world problems. By using this technique the affordances of each of the technologies was rated, in the context of the dimensions model, creating a set of technology “top trumps” for work-integrated assessment. Examples of these “top trumps” are shown in Figures 5 & 6.

The overall model or top trump system is designed to recommend not just one potential technology to support a pedagogic design, but multiple technologies that might be appropriate. In accordance with the features of affordances outlined earlier, this design feature should allow individual students and/or academics to select specific technologies from a choice based on personal needs and experience.
Iteration

Fourteen iterations have currently been planned with the model, in such disparate disciplines as Spanish, Engineering, Geography, Law, Education, Mathematics and Politics. This paper focuses on the first completed iteration, with 25 students taking Professional Issues and Development (PSY1301), a first year module of a BSc Applied Psychology (Clinical) programme.

Stage 1: analysis

The original assessment for the module consisted of a 1½-hour written examination (75%) and a 20-minute oral presentation (25%) that was undertaken in pairs with each student contributing 10 minutes. The overall position of these assessments was decided by a collaborative design team and sketched out on the dimensions model, which can be seen at the heart of Figure 4 as the smallest polygon.

Stage 2: design

Pedagogic changes

The next step was to discuss what changes to the assessment might be made in order to move along each of the dimensions. These changes were discussed by the same team, which made the following design decisions:

- **Problem/Data**: A shift from the examination to the creation of a patient information leaflet
- **Time**: The design and creation of the patient information leaflet would take place across the module, depending in part on the content of the lectures
Audience: Two external groups, with links to the Psychology department, would act as new audiences for the leaflet and additional markers. Collaboration: Students would work in assigned teams of six. Review: Team working would encourage and develop informal review. Structure: Detailed guidance for the creation of the patient information leaflet would not be given, instead signposts would be provided to help students find relevant guidance independently.

The anticipated consequences of this design can be seen in the larger polygon in Figure 4.

Technology alignment
Technologies were then identified that might support this assessment. Of the six dimensions being considered, it was felt that those of structure and collaboration were critical to the overall task, and potentially the most challenging for first year students. For this reason, technologies were sought which matched a star rating of five for both the structure and the collaboration dimensions. Using this process, “top trumps” for Do, Trello and TeamMatch were recommended, as shown in Figures 5 & 6.

- Do is a project management tool that supports work on shared or private tasks, notes, activity feeds and comments to track progress.
- Trello is a project management and collaboration tool that organises projects into boards, lists and cards.
- Team-Match is a web application that helps create teams, analyse team strengths and weaknesses and identify problem zones.

Technology demonstrations
Once technologies had been identified that had the potential to support the pedagogic design, the students were given 10-minute demonstrations to show how these technologies might be useful in supporting their assessment.

- TeamMatch was demonstrated first, and the students completed a personality questionnaire which gave them an individual profile based on the Myers-Briggs Type Indicator, with automatic creation of well-balanced teams based on the results of these profiles.
- Do and Trello were then demonstrated together, using a mocked-up series of tasks, showing how each of the project management tools would deal with the various project management challenges that students would encounter.

Stage 3: evaluation
Once the iteration had been completed, evaluations were carried out with the students, academics and employers.

A student session was organised where students completed three evaluation exercises:
An individual activity using eight A1 sized “mark on the line” activity sheets placed around the walls of the evaluation room, with questions about their overall thoughts on the module. This was followed by an evaluator-led discussion of the collective responses, which were analysed by allocating them into quartiles: the results can be seen in Figure 7.

An individual activity where students answered the question “What’s the most important thing you have learnt?” by writing on postcards designed for the session. Results were analysed using a simple coding technique, where codes were derived directly from the data using an exhaustive process. This was feasible because of the limited amount of text allowed for by the postcard format. The results are shown in Figure 8.

A group activity where students had the opportunity to rate how much they had developed their skills during the module, using a simple drag and drop activity on interactive computer surface tables. These snooker table sized computers allow groups of students to interact with digital content simultaneously, much like a very large iPad. This activity was designed so that the ratings could later be analysed and compared to the dimensions model, giving an overall picture of skills development against each of the six dimensions. Results for this evaluation are displayed on a combined dimensions model for the iteration, shown in Figure 9 as the solid line.

The personal perspectives of academics, employers and students involved with the assessment were evaluated using semi-structured interviews.

Figure 7. “Mark on the line” evaluation data.
Discussion

The evaluation data as shown in Figures 7, 8 & 9 suggest that the overall “authentic” assessment approach taken by the project was effective and that the inclusion of employers in the assessment process was beneficial. Although some students commented that the module design was unexpected, data show that they enjoyed the module nevertheless, and they appear to have developed their skills as had been intended through the design. Detailed discussion focuses on the key aspects of the design.

Use of the dimensions model to design authentic assessments

The dimensions model has now been tested by a wide range of academic staff, from senior members with extensive educational experience to relative newcomers to the field, and has been well received by all. It seems to require little guidance in its use, and in the iteration described it served the designed purpose in supporting assessment change, as skills and competencies do appear to have been built into the fabric of the module. It has been particularly effective in prompting thinking in unexpected ways, such as along the dimensions of audience and structure.

The success of the model may in part be due to the fact that it “invites engagement”, that is, it is in part blank and designed to be actively completed, as opposed to being read passively. Those who use the model therefore have ownership of it; it becomes part of their thinking rather than that of the original designers. Users are effectively constructing their own meaning from the model, and the model itself serves only as a scaffold for their own thinking.

Integrating technologies using “top trumps”

The purpose of the technology “top trumps” was to provide a way of aligning “off the shelf” technologies with the dimensions model. By allocating one to five stars
Figure 9. Drag and drop skills evaluation (completed dimensions model).
against each technology to quantify their affordance in relation to the six dimensions, specific technologies could then be recommended that should have a positive impact on an assessment designed using the model. In this iteration, technologies were recommended to support the structure and the collaboration dimensions, specifically to support students in developing skills in collaboration, team working and project management.

Analyses of data gathered from the evaluation sessions with students (as shown in Figures 7, 8 & 9) all showed that the students had developed their collaboration skills. Analyses of the postcards data and the skills drag and drop exercise suggested that the students had developed time management skills, and their ability to work autonomously. Whilst we cannot claim a causal relationship between the technologies used and these results, there is a strong correlation between the designed objectives and the evaluated outcomes in terms of developing skills and competencies. We propose that this is at least in part due to the alignment of appropriate technologies with the pedagogic design.

Interviews with academic staff and students appeared to corroborate this conclusion. The academic lead commented that some students “adapted technologies that they were familiar with to suit this purpose”, for example two groups decided to use Facebook to organise work rather than the suggested tools. This type of innovative use of technology is predicted by the concept of interpretive flexibility, but suggests that the “top trumps” do need refining further. Crowdsourcing, that is, using student input to moderate star values over an extended period of time, is being explored as a possible long-term solution. Facebook was not considered as providing enough Structure to warrant five stars for this dimension, though does score five stars for Collaboration. The Trello software was only adopted by a small number of students, but it was deemed to be “absolutely excellent” in supporting their team working and time management, and has now been adopted by some students for all their studies. The TeamMatch software was useful in highlighting to the students their own strengths and weaknesses, though it was not used in the first iteration to form teams. No students reported using Do.

The use of affordances

The use of affordances as the underlying theory to align technologies with a pedagogic approach may be questionable, and there is certainly much debate over the pros and cons of the term from an educational perspective (Hammond 2010). Nevertheless the project seems to have had some success using the theory of affordances to highlight the strengths of technologies when considered in a specific pedagogic context.

A possible reason for this success is the project’s focus on Gibson’s (1979) original definition of affordance, rather than later evolutions most often ascribed to Norman (1988). In particular we have focused on affordances as functional, and not as “action possibilities”, a term which now seems synonymous with affordances (although there is no mention of this term in Gibson’s original work). By functional affordance we mean what something “provides or furnishes”, in line with Gibson’s original definition, rather than simply how it can be interacted with, which was more the concern of Norman. We have also avoided broad taxonomy approaches, such as that taken by Conole and Dyke (2004), or detailed classification systems, such as that
proposed by Bower (2008), instead focusing on analysing affordances of technologies in relation to a specific context (i.e., the dimensions model), in a more “targeted” approach.

Understanding affordances as “action possibilities” is arguably limiting its usefulness in educational contexts, as it fixates technologies as tools, “things” to be prodded, pushed or pulled. We argue that a functional perspective brings time into the affordance concept, and allows technologies to be explored not simply as tools but also as specific “places” in their own right, with functions that happen in and across time. Blogger, for example, provides not simply a tool to write, but a place capable of providing for self-reflection and peer review; Evernote is not simply a note taking tool, but a place for centralising and contrasting disparate experiences. By conceptualising affordances in this way, we suggest that they may yet be useful in helping to develop pedagogic strategies similar to what we have presented for integrating technology effectively.

Conclusion

This paper has described the creation of a new model for “authentic” assessment design, which integrates off-the-shelf technologies using an affordances approach. Results from the first iteration using this model suggest that it is effective in supporting the design of an “authentic” assessment and that the targeted affordances approach can support the alignment of technologies with a pedagogic design. Further iterations with students studying Spanish, Geography and Law are also showing strong support for the overall approach, and more iterations are planned across the 2013/14 academic year to refine and develop the concepts presented.

Acknowledgements

We thank Stuart Redhead and Charlotte Anderson, Project Officer and Project Assistant respectively for the Collaborate project, for their invaluable assistance, particularly with the design of the model and the associated technology “top trumps”.

References


