






UNIVERSITY OF
CAMBRIDGE
Discover something new

Digging for Validity
Developing new ways of assessing extra-curricular achievement





Carenza Lewis MA ScD FSA
Director, Access Cambridge Archaeology

www.arch.cam.ac.uk/access/ McDonald Institute for Archaeological Research, Downing Site,
Cambridge CB2 3ER



Digging for Validity

- HEFA
 - Course aims
 - Course structure
 - Assessment aims
 - Assessment methods
 - Outcomes
- Refining the HEFA assessment model
 - Articulating the aims of the learning programme
 - Developing the construct base
 - Evaluating the assessment model
- The new assessment model
 - Application and use of the assessment model in the future
 - HEFA & HEFA G&T
 - CAP
 - A freestanding qualification/certificate
 - Concluding thoughts

Digging for Validity

- The assessment procedures for the Higher Education Field Academy, recently refined in collaboration with ARD, has highlighted a number of conceptual challenges which have the potential to resonate more widely within the field of assessment, especially of 'soft' skills or PLTS.
- A principal challenge has been to construct a model for assessing 'soft' skills in a way that is sympathetic to the intentions of the learning programme whilst also providing robust observation-based outcomes and avoiding heavy requirements that might corrupt learning relationships.
- The ideal outcome would be an approach to validly assessing 'soft' skills that could be scaled up in, an activity that requires high staff-learner ratios.

Digging for Validity



Higher Education Field Academy

Aims of HEFA

To **raise the educational aspirations and achievements** of **young people from less privileged backgrounds**, particularly with regard to progression to university, in order to enable them to be a vital part of the highly-qualified workforce that will be needed in the future.

To enable participants to develop a wide range of **practical, cognitive, personal and presentational skills** vital to success in school, university and the workplace, while **contributing to new and current research** at the University of Cambridge.

To allow **young people and local communities** to work together to **advance knowledge and understanding of the rural historic environment**.



Digging for Validity

Higher Education Field Academy

Method

HEFA gives pupils the chance to **do something new** to them which is **challenging, structured, multi-skilled and goal-oriented** and which results in a substantive piece of independently carried-out, individually-completed assessed work which is of **value to themselves and also of benefit to others**.

HEFA pupils achieve this by carrying out their own **two-day archaeological excavations** over two days working independently in mixed-school teams of 3-4 excavating on of around **10 small simultaneous 1m² test pits** in gardens and other open spaces within villages and hamlets.

Digging for Validity



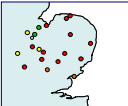
Higher Education Field Academy

Task suite 1

To complete all stages - from initial site preparation through to writing up - of one of several concurrent 1m square archaeological 'test pit' excavations to professional standards to look for archaeological finds

In order to help tell

- 1) How and when humans used each place in the past
- 2) How an entire settlement developed in the past
- 3) How settlements across a region developed and what this can tell us about wider historic events and processes.



Digging for Validity



Higher Education Field Academy

Task suite 2

To acquire, improve, apply and demonstrate skills which will be useful for life and learning.

These skills include

- 1) Data collection and analysis
- 2) Working to set standards
- 3) Verbal communication
- 4) Creative thinking
- 5) Reflective learning
- 6) Team working
- 7) Report writing.





Digging for Validity



Higher Education Field Academy

Task suite 3

To find out more about university and how to get there

To do this by

- 1) Taking part in this university-level excavation led by a university team on Days 1 and 2
- 2) Meeting people who are teaching or learning at university
- 3) Coming to spend a day in a university on Day 3
- 4) Attending a workshop on applying to university.





Digging for Validity



- Success in all three suites of tasks are vital to the success of the HEFA programme
- The use of a robust valid model for assessment is key to the **second** suite of HEFA tasks - **acquiring, improving, and applying skills useful for life and learning**
- Being able to **demonstrate** that these have been achieved is crucial to the long-term success of HEFA and maximise HEFA's benefit to students



Digging for Validity



- Assessment (original)
 - Achievement of HEFA aims
 - Standards of student performance



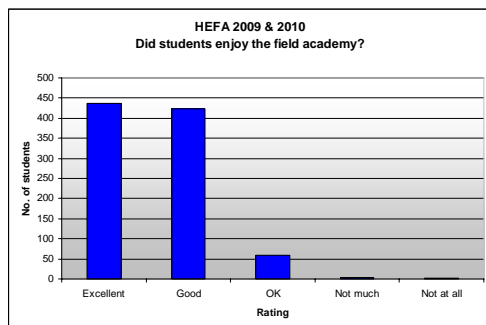
Digging for Validity



- Assessment methods for student performance (original)
 - Student self-assessment and impact assessment – feedback forms completed after day 3 by students
 - Fieldwork assessment form – performance assessed by on-site observer/supervisor assessing 4 students each
 - Written report assessment – performance assessed by remote assessor assessing all reports submitted each year
 - Return of assessment (comments and marks) to students

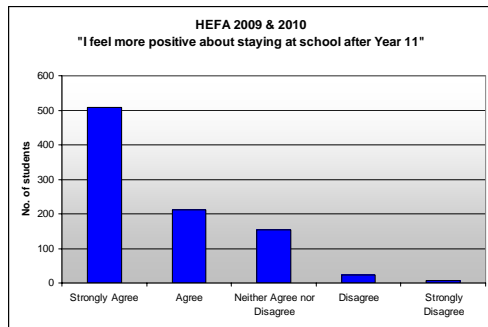


Digging for Validity



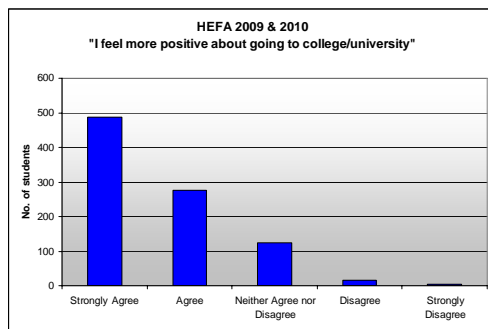


Digging for Validity





Digging for Validity





Digging for Validity



2009 & 2010	Male (no.)	% of total male	Female (no.)	% of total female	Total (no.)	% total
90-100% (A*)	253	69%	238	63%	491	66%
80-89% (A)	50	14%	80	21%	130	17%
70-79% (B)	37	10%	32	8%	69	9%
60-69% (C)	14	4%	18	5%	32	4%
59% and below (D&E)	12	4%	11	3%	23	3%
Grand Total	366		379		745	

HEFA 2009 & 2010 fieldwork grades (combined for effort, attitude and standard of work) assignment marks



Digging for Validity



2009 & 2010	Boys (no.)	% of total boys	Girls (no.)	% of total girls	Total no.	% of total
80% and above (A/A*)	26	19%	80	44%	106	33%
70% - 79% (B)	43	31%	55	31%	98	31%
60% - 69% (C)	58	42%	43	24%	101	32%
59% and below (D-E)	12	9%	2	1%	14	4%
Total	139		180		319	

HEFA 2009 & 2010 written assignment marks



Digging for Validity



Refining the assessment model

- Articulating the aims of the learning programme
- Developing the construct base
- Evaluating the assessment model



Digging for Validity



- Articulating the aims of the learning programme (1)
- The first phase of the project involved explicating the aims of the HEFA learning programme.
- This was important in order that the development process could align the assessment model with these aims without compromising them.
- This phase involved
 - analysing published HEFA documentation (including promotional documents, programme outlines, existing assessment materials etc)
 - focussed discussions with the HEFA team members so that the ARD could come to a shared understanding of the programme aims.



Digging for Validity



- Articulating the aims of the learning programme (2)
- An important part of setting out the aims was to express the values that underpinned them.
- These elicited values included the belief that knowledge and practical work need to be considered as interconnected, and that there is intrinsic value in allowing young people to develop and demonstrate the application of multiple skills in an integrated way.
- An important part of the programme was also its focus on 'inclusion', and how this issue of interest in general education could be dealt with through the specific context of working to professional standards as a member of an archaeological team.



Digging for Validity



- Articulating the aims of the learning programme (3)
- One early concern that arose from this articulation of values was that attempts to move the programme into a formal national qualifications framework context could damage the programme by overburdening the programme with a heavy assessment architecture.
- Bearing this in mind the next stages of development involved making explicit the construct bases of the learning programme so that these could be reflected in the new assessment framework.



Digging for Validity



Developing the construct base (1)

- Learning programmes (and therefore allied assessment procedures) can be described in terms of three related concepts:
 - *tasks* (the activities that afford particular behaviours). These concepts are also useful in structuring evaluations of assessment programmes.
 - *constructs* (the skills and attributes that the programme covers and that the assessment seeks to recognise);
 - *observable behaviours* (the identifiable actions which give inferential evidence about the constructs to be assessed and which might be scalable).



Digging for Validity



Developing the construct base (2)

- The tasks of the HEFA programme are well articulated; being structured around the requirements of professional archaeological activity.
- For the HEFA/ARD collaboration the articulation of constructs was a key part of the development as this informed the generation of assessment criteria that would help assess particular observable behaviours.
- The construct development phase involved two aspects;
 - construct elicitation
 - construct review.



Digging for Validity



Developing the construct base (3)

- **Construct elicitation** was a complex task and involved a variety of methods and participants:
- Experts involved in developing the HEFA programme reviewed other course material, such as curriculum documents and assessment models from cognate areas designed for young people at the same stage of learning.
- Two researchers from ARD attended a HEFA course to collect data about constructs through observations of young people and supervisors at work during a HEFA course.
- Views from stakeholders were taken into account through a survey of teachers who had taken part in past HEFA courses.



Digging for Validity



Developing the construct base (4)

- Construct review was carried out once the constructs had been elicited and codified into a new assessment framework.
- The process of construct review was undertaken through a series of joint HEFA/ARD meetings.
- The ultimate aim of this review phase was to be able to establish that the sampling of constructs was adequate to support any validity claims for the assessment.
- These meetings evaluated the construct base of the new assessment framework through considering:
 - the overall balance of construct representation,
 - whether the constructs were considered to be essential or desirable,
 - whether the constructs were binary or scalable,
 - whether there were any overlaps in constructs..



Digging for Validity



Developing the construct base (4)

- Once the process of construct review was completed the constructs were broken down into separate criteria, based on observable behaviours, during a series of meetings.
- This process involved the HEFA experts using a variation of an Angoff (1971) standard setting method.
 - Experts' conceptualisations of moderate competent performance on a task were used to establish a reference point.
 - This point was used to help articulate performance descriptors that exceed or fall short of this standard.
 - These criteria were then developed into either binary or scaled descriptors.



Digging for Validity



Structured working			
Plan and carry out scheme of work in structured manner and within required time.	This individual has shown minimal interest in planning or carrying out the required tasks in the correct order or within a reasonable time-scale, taking either far too much or far too little time on tasks.	This individual has used the handbook to plan and follow the required work programme correctly. With some help he/she has identified objectives and defined tasks clearly. Task-flow has been planned and managed; progress has been monitored effectively.	This individual has used the handbook without prompting to pro-actively plan and help drive through an effective programme of work. He/she has identified objectives and defined tasks clearly and has been pro-active in monitoring progress and ensuring tasks are completed on time.

Extract from HEFA assessment framework



Digging for Validity



Developing the construct base (5)

- At this stage the assessment framework consisted of four sections, with two of these containing a number of subsections.
- In total the framework included 39 separate assessment criteria, broken down into scaled or binary assessment statements.

Section	Subsection
Record of participation	
Data collection exercise	<ul style="list-style-type: none"> •Setting up and workspace management •Data/sample collection •Measuring and recording •Working to set standards
Personal, learning and thinking skills (PLTS)	<ul style="list-style-type: none"> •Verbal communication •Structured working •Creative thinking •Reflective learning •Effort and persistence •Team working
Written report	<ul style="list-style-type: none"> •Report structuring and data presentation •Research skills •Writing skills •IT skills •Independent working

The HEFA assessment structure

Section	Subsection
Record of participation	
Data collection exercise	<ul style="list-style-type: none"> •Setting up and workspace management •Data/sample collection •Measuring and recording •Working to set standards
Personal, learning and thinking skills (PLTS)	<ul style="list-style-type: none"> •Verbal communication •Structured working •Creative thinking •Reflective learning •Effort and persistence •Team working
Written report	<ul style="list-style-type: none"> •Report structuring and data presentation •Research skills •Writing skills •IT skills •Independent working

The HEFA assessment structure

Section 2 - Data collection				
Assessment Criteria	Range Descriptor (low)	Range Descriptor (middle)	Range Descriptor (high)	
<p>6 Use appropriate equipment effectively to excavate a precise square test pit exactly 1m², creating and maintaining straight vertical sides at all times and extracting spoil in standard 10cm spits (containing horizontal surfaces clean and level on completion of each context).</p>	<p>This individual has rarely used the correct tools (spade, mattock and shovel) in the correct way, and has shown no interest in acquiring or displaying any proficiency in their use. He/she has made hardly any contribution to the task of loosening and removing spoil from the test pit, keeping the test pit the correct shape, size and precision or he/she has repeatedly excavated beyond the required level or required repeated prompting in order to stop at the appropriate level.</p>	<p>This individual has used the correct tools (spade, mattock and shovel) in the correct way more often than not, and has shown interest in acquiring or displaying proficiency in their use. He/she has made a useful contribution to the task of loosening and removing spoil from the test pit and keeping the test pit the correct shape, size and precision. He/she has excavated up to but not beyond the required level most of the time or with some prompting.</p>	<p>This individual has used the correct tools (spade, mattock and shovel) in the correct way at all times, has shown great interest in acquiring or displaying proficiency in their use. He/she has made a very significant contribution to the task of loosening and removing spoil from the test pit and keeping the test pit the correct shape, size and precision. He/she has excavated up to but not beyond the required level at all times without prompting.</p>	
<p>8 Process a series of extracted 0.1m² spoil samples using appropriate methods (sieving and/or hand sorting) in order to identify, recover and retain target data (in items of archaeological interest).</p>	<p>This individual has taken minimal care to search spoil for finds, has only used the sieves correctly with repeated prompting and has not engaged with identifying when sieving is appropriate and when hand sorting would be more useful. This student has made hardly any contribution to the task of processing spoil.</p>	<p>This individual has taken care to search spoil for finds, has used the sieves correctly with only limited prompting and has been able to identify when sieving is appropriate and when hand sorting would be more useful with only limited prompting. The student has made a useful contribution to the task of processing spoil.</p>	<p>This individual has at all times taken the greatest possible care in searching spoil for finds, has used the sieves correctly with no prompting and has been able to identify when sieving is appropriate and when hand sorting would be more useful. The student has made a very significant contribution to the task of processing spoil.</p>	
<p>8 Critically observe extraction site while extraction is in progress in order to ensure that contextual information is not overlooked.</p>	<p>This individual has taken hardly any care to look for and identify features such as walls or cuts into the ground in any way, during soil loosening or travelling, despite repeated prompting.</p>	<p>This individual has mostly taken care, after a little prompting, to look for and identify features such as walls or cuts into the ground by (a) making visual observations while loosening and removing spoil, (b) using a trowel to remove loose spoil from the surface of the pit at least every 10cm and/or at the bottom of each 10cm context and (c) looking for differences in texture and appearance of cleaned surfaces.</p>	<p>This individual has taken the greatest possible care, within a reasonable time-frame, to look for and identify features such as walls or cuts into the ground by (a) making visual observations while loosening and removing spoil and (b) using a trowel to remove loose spoil from the surface of the pit at least every 10cm and/or at the bottom of each 10cm context and (c) looking for differences in texture and appearance of cleaned surfaces.</p>	



Digging for Validity



Evaluating the assessment model (1)

- The third development phase gathered empirical evidence about the functioning of the assessment framework in two stages.

Stage 1:

- The first stage of evaluation involved researchers observing nine team supervisors as they assessed a total of 26 young people taking part in a HEFA course over two days.
- This evaluation gathered data in a variety of areas:



Digging for Validity



Evaluating the assessment model (2)

- One area was to consider the practicalities of the assessment through looking at the pattern of assessment behaviours;
 - when were assessments taken,
 - how the assessors completed the assessment documents,
 - how the assessors organised their data gathering
 - how long these assessments took.
- The evaluation also compared the assessment outcomes of assessors who were located in one test pit for the whole course with those outcomes of assessors who rotated around different test pits.



Digging for Validity



Evaluating the assessment model (4)

- Researchers also used interviews to gather qualitative evidence from team supervisors/assessors to investigate
 - whether the language of the assessment framework was clear,
 - what the effect of the assessment was on their relationship with those being assessed,
 - levels of confidence in making assessment judgements based on the framework
 - whether there were any duplicated or missing constructs.
- Those being assessed were also surveyed to investigate whether they perceived the assessment process was invasive and how it affected their learning.
- Analyses of the assessor evidence suggested that the supervisors who used the assessment scheme felt confident in its application and the validity of its outcomes.



Digging for Validity



Evaluating the assessment model (5)

- The final aspect of the first evaluation stage was a statistical analysis of the assessment outcomes.
- Comparisons between the quantitative mark outcomes for the young people generated by the assessment framework with independently gathered rank orders of their overall performance were used to validate the assessment framework.



Digging for Validity



Evaluating the assessment model (6)

- Following this first evaluation stage a number of amendments were made to the wording of the assessment framework.

Stage 2:

- This led to a second evaluation stage which considered issues of inter-assessor agreement, involving four assessors and 14 young people over two days.
- The aim of this evaluation stage was to look at whether two independent assessors could reach similar conclusions about the same performance using the assessment framework.



Digging for Validity



Evaluating the assessment model (8)

- In advance of the evaluation a pre-assessment standardisation session was arranged, where assessment managers and assessors discussed the assessment documentation in relation to pieces of video evidence captured during earlier HEFA courses.
- Each pair of assessors oversaw two test pits, rotating between these throughout the two-day period of the course.
- Analysis then considered the extent of agreement between the two pairs of assessors, looking at
 - the proportion of perfect and near agreement,
 - whether any assessor variance was systematic or random,
 - whether any variance would affect any notional pass/fail thresholds,
 - whether there needed to be additional weighting or marks awarded for some constructs.



Digging for Validity



Evaluating the assessment model (8)

- Examples of completed assessment documents (to hand round)



Digging for Validity



Evaluating the assessment model (9)

- These analyses led to further refinements and the final drafting of the assessment framework which still retained the structure that was proposed prior to the evaluation phase. The final framework covers a diverse skills base, including a variety of hard and soft skills.
- The hard skills are mainly represented in the *data collection* and *written report* sections of the assessment.
 - Data collection processes involve the young people working to professional archaeological standards, using appropriate tools to extract spoil samples, mapping the extraction site, and processing recovered finds.
 - The written report assessment considers aspects of how well the participants have structured their report and presented their data, including their research, writing, and IT skills.



Digging for Validity



Evaluating the assessment model (10)

- Equally important is the recognition of the young people's soft skills or PLTS; these include aspects of the participants'
 - verbal communication,
 - structured working,
 - creative thinking,
 - reflective learning,
 - effort and persistence,
 - team working.
- These assessments involve a mixture of ongoing, on-site observations by team supervisors and, in the case of the written report, post hoc assessments by a single assessor from the HEFA team.



Digging for Validity



The new assessment model

- Application and use of the new assessment model
 - In HEFA (G&T)
 - In CAP
 - As a freestanding qualification...



Digging for Validity

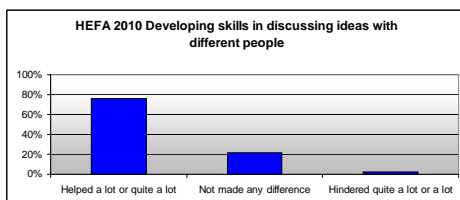


Developing the new assessment model

- Concluding thoughts
 - Useful exercise for eliciting constructs
 - Useful exercise for enhancing ability to demonstrate validity of assessment
 - Useful exercise for sharpening up assessment and reporting protocols
 - Refinements to student self-assessment feedback forms - eg

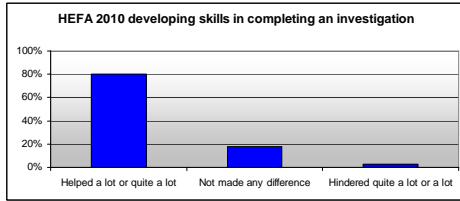


Digging for Validity



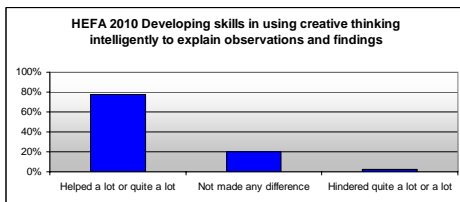


Digging for Validity



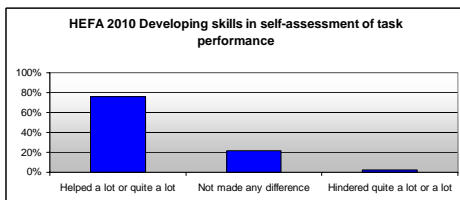


Digging for Validity



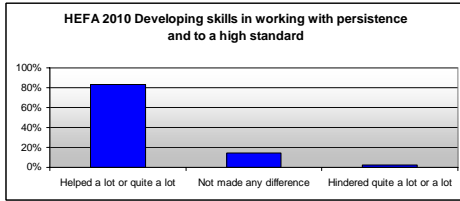


Digging for Validity



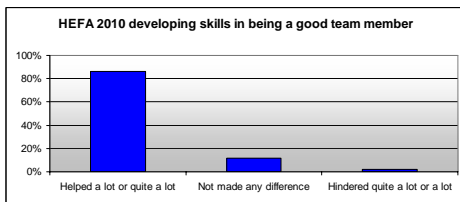


Digging for Validity



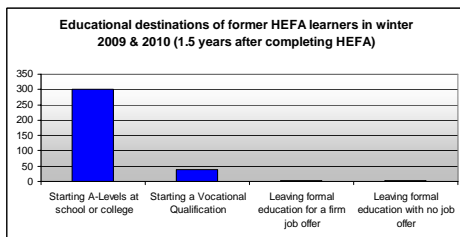


Digging for Validity





Digging for Validity





Digging for Validity



Developing the new assessment model

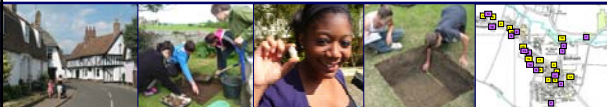
- Concluding thoughts
 - Useful exercise for eliciting constructs
 - Useful exercise for enhancing ability to demonstrate validity of assessment
 - Useful exercise for sharpening up assessment and reporting protocols
 - Refinements to student self-assessment feedback forms
 - Refinements to reports to students
 - » Provide more detail on specifics
 - » Not yet used new written report framework
 - » Some loss of personal touch – more apparent than real...
 - Very valuable as route to endorsement from OCR
 - Useful for developing CAP
 - Potential for wider application as free-standing qualification in future



**UNIVERSITY OF
CAMBRIDGE**
Discover something new

Digging for Validity

Developing new ways of assessing extra-curricular achievement



Carenza Lewis MA ScD FSA

Director, Access Cambridge Archaeology

www.arch.cam.ac.uk/access/ McDonald Institute for Archaeological Research, Downing Site,
Cambridge CB2 3ER
