Selective targeting of feedback for best effect on learning:

Core Disciplinary Assessment Criteria

Steve Draper, Psychology

Quintin Cutts, Computing Science

LTConference 22 April 2010

1

Part A:

Tacit assumptions to question

2

Tacit assumptions about providing feedback

Discussions of providing feedback to students in the literature, in staff meetings, and the NSS tend to assume unless otherwise stated that feedback:
- Is about staff judgements on finished work
- Is universally necessary to learning
- Should be evenly distributed across all courses and modules

This talk implies these are wrong, and that hugely uneven, i.e. targeted, allocation of feedback effort may be more effective.

3

Tacit assumptions about course design

Those feedback assumptions well up in a background where the great majority assume that course design:
- First decides the content to be covered
- Then the assessments by which to test the acquisition of content

- *They then argue this is "alignment". But:*
  - ❑ *Alignment is equal if you decide the assessments and then the content*
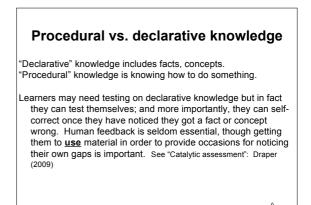 - ❑ *Biggs stated there were not 2 but 7 elements of course provision to align*

This talk implies that these assumptions too are wrong.

4

Part B:

Feedback is more important for procedural knowledge

5

Procedural vs. declarative knowledge

"Declarative" knowledge includes facts, concepts.
"Procedural" knowledge is knowing how to do something.

Learners may need testing on declarative knowledge but in fact they can test themselves; and more importantly, they can self-correct once they have noticed they got a fact or concept wrong. Human feedback is seldom essential, though getting them to **use** material in order to provide occasions for noticing their own gaps is important. See "Catalytic assessment": Draper (2009)

6

Procedural knowledge

A fact is a single item: if it's wrong, there's no puzzle where the problem is.

A procedure, however, is a long sequence of actions. If it's wrong (produces the wrong answer or effect) it is generally not clear which part of it was wrong. Diagnostic feedback is very important, though advanced learners eventually acquire sophisticated meta-procedures for self-diagnosis of failed procedures.

If you give me an essay and I just say there's something wrong with it, that is of little help. If you bake a cake, and when it comes out of the oven it is obviously bad, again you are often at a loss. Similarly if your computer program just fails, you don't know which line and which symbol is responsible.

Generally speaking, procedures have many more component parts; and learners are much more in need of both practice and helpful feedback in learning them.

7

Feedback on procedural learning

The literature also supports this, that feedback has more positive effect on learning for procedures than for declarative material.

Hattie & Timperley (2007) "The power of feedback" can be interpreted as arguing that feedback on procedural learning is more important, because (only) it leads to transfer.

8

Discussion?

Can you think of evidence of learning declarative knowledge without feedback?

Can you think of evidence of needing feedback when learning procedural knowledge?

9

Part C:

Two course designs with a tacit focus on CDC

1. Quintin: Intro computer programming

2. Vicky Gunn history course "Body & Belief"

10

Quintin's level 1 programming course

Programming students are focussed on the (procedural) task of writing programs.

A student may have to find 50 answers to finish an exercise.

They cannot possibly wait to see a tutor (once a week) to do this.

Success therefore depends upon rapidly acquiring the meta-skill of answering their own queries and problems.

Self-help is the core disciplinary skill for programming

N.B. this self-help is needed in a feedforward (not feedback) slot: while the task is being done, not afterwards.

A dozen changes to address this were introduced.

But perhaps more direct and varied teaching on this "debugging" skill might be still better?

11

Stimulating proactive self-help (2)

- Students were set pre-reading before Friday lecture — EVS then used to probe and hopefully correct understanding.

- Students given pre-lab work on Friday, to be done before entering their lab on Mon-Wed. In lab, Tutors comment as necessary.

- Feedback sheets (comment banks) encourage looking at others' problems.

- Asking tutors for lab feedback for use in Wed. lecture

- Using Mahara, students write reflectively at the end of each lab — main points summarised and presented in the Wed. lecture

- Harped on about "Read the problem immediately, sleep on it, read it again, try it, sleep on it etc. etc." [Direct instruction on self-help]

- Used Peerwyse (student generated question system)

12

Body & Belief

Redesigned by Vicky Gunn for the 2008-9 session.
Co-taught with Sarah Nicholson.

Honours option, History discipline, over two semesters, mixture of
level 3 and 4 students.
50% coursework, 50% final unseen exam
One 2-hour class per week

A major redesign was introduced (same staff), which was
evidently a huge success.

13

Learner evaluation: quick measures (2)

Votes for best feature of the course [some top equal votes] (plus in
parentheses: rated "very important" or "best"):

4 (8) Interesting subject matter

2 (8) Great discussion and group feeling between students

1 (8) The staff teaching it are great

2 (7) Approaches and skills I learned on this
course, I'm applying to other courses

1 (7) The methods of teaching / learning kept my attention

0 (3) I saw my work was useful to others (not artificial)

Body & Belief (4)

They did one early essay;

Then a series of formats on a single topic (each student selected
their topic): tutor feedback/forward on their preparations for a
presentation; the actual presentation to the class; a writeup of
the presentation.

Lots of feedback and feedforward
Linked tasks exercising the same core issues in different contexts;
But also, building on their development of the same (personal)
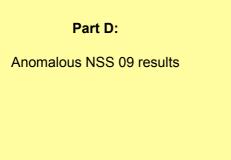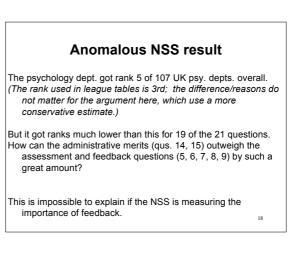topic in terms of content.

15

B&B as the tutorial component?

The staff:student ratio of 2:9 would seem wildly extravagant to a
big department (like psychology).

But as a tutorial group, this is normal: easily justified if it were
implicitly carrying the entire tutorial burden of the department.
There is no separate tutorial strand in that department.

The fact that students say that B&B has affected their approach
to study on the other honours courses (but not vice versa)
supports this interpretation.

16

Part D:

Anomalous NSS 09 results

17

Anomalous NSS result

The psychology dept. got rank 5 of 107 UK psy. depts. overall.
*(The rank used in league tables is 3rd; the difference/reasons do
not matter for the argument here, which use a more
conservative estimate.)*

But it got ranks much lower than this for 19 of the 21 questions.
How can the administrative merits (qus. 14, 15) outweigh the
assessment and feedback questions (5, 6, 7, 8, 9) by such a
great amount?

This is impossible to explain if the NSS is measuring the
importance of feedback.

18

Rank	Qu.	Qu.text
1	14	Any changes in the course or teaching have been communicated effectively.
2	15	The course is well organised and is running smoothly.
5	22	Overall, I am satisfied with the quality of the course.
8	6	Assessment arrangements and marking have been fair.
8	11	I have been able to contact staff when I needed to.
8	16	The library resources and services are good enough for my needs.
11	13	The timetable works efficiently as far as my activities are concerned.
16	1	Staff are good at explaining things.
35	5	The criteria used in marking have been clear in advance.
54	7	Feedback on my work has been prompt.
79	9	Feedback on my work has helped me clarify things I did not understand.
101	8	I have received detailed comments on my work.

Psychology honours design

For the students responding to NSS 2009, the design had been:
   Level 3:
9 modules, class exam with some formative feedback on 4
   Level 4:
6 modules, no related coursework

BUT

   Level 3:
2 CRs (critical reviews), 2 miniprojects with tutorial groups of 5-6
   Level 4:
1 CR, 1 project each with a personal tutor

20

Psychology honours design (2)

So the programme design could be redescribed as investing 100% of its tutor time in focussing on equipping the students with the ability to display critical thinking (of the kind a psychologist values). It invented a type of coursework ("critical review") that announces to students what the main point is; it requires them to produce 3 month long pieces of work focussed on it; but also marks their exams with this requirement applied.

It is the hardest thing they must learn; the most important thing; almost all our teaching investment is put into it; and the students rated us 5 out of 107 in the UK.
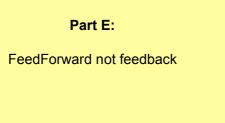
21

Psychology honours design (3)

The investment had been on:

• Feedforward not feedback

• A core procedural skill, not "psychological content"

• Understanding (operationally) the most difficult assessment criterion

• How this applies in a variety of different surface tasks

22

Part E:

FeedForward not feedback

23

Feedforward

In almost all the educational literature, in everyday understanding (of students answering the NSS), and in engineering, "feedback" means giving a measurement on finished work.

Feedforward means taking a measurement now and estimating what will be required for success in the future. This is what project supervisors do.

PhD students only ever get feedforward (apart from post-viva corrections): they cannot ever get feedback.

Educational feedforward means essentially taking the learner's plans and steps so far, and giving direction on what's good, what should be changed.

Feedforward may be somewhat harder for tutors to provide than feedback, since they cannot see the full effect of a learner's plans until the work is completed. But for the learner, it must be much more useful, and saving of time and effort, because it saves the waste of building on a mistaken step.

Feedforward (2)

Another way to look at it is to imagine what inner mental changes an ideal learner would make in response to feedback, and whether these could all be made in response to feedforward.

Consider the simple case of a learner being told they have misspelled "Wiliam" in "Black & Wiliam". They will hopefully change 3 things without further instruction:
• The spelling used in the current piece of work
• Their internal "generator" so they will spell it correctly when writing in future.
• Their internal proof reader or bug detector, so that when reading over their own or others' work, they will detect if it's misspelled.
They can do the last two just as well from feedback and feedforward, but it will save work on the first, the earlier (more "forward") the correction is communicated. If instead of spelling, it is information on the right statistics test to use, or the right structure for a major essay, huge amounts of work could be saved.

25

Feedforward (3)

Feedforward is at least as good, and in fact generally more useful, than feedback.

Courses with no feedback whatsoever, but good feedforward, should work fine.

26

Evidence from a puzzle about RPC (reciprocal peer critiquing)

Morrow (2006) found strong student attitude support for RPC's benefits, but strongest for being able to see others' work.
That's also what I find repeatedly in oral feedback.
This doesn't exactly match published theories of feedback.

Price et al (2007) found the same.

Students believe it's useful after having experienced the process; and then act on their belief by doing it voluntarily.
But it's not clear how to measure learning gains.

Not least because the gains may only be far in the future and certainly NOT on the current piece of work.

Discussion?

How much feedforward compared to feedback is there in your department? (or was there when you did your degree?)

Is feedforward better than feedback?

Any examples where one is better than the other?

28

Part F:

Core Disciplinary Procedures are what matters

29

Core disciplinary assessment criteria

We know from the feedback literature, especially Sadler 1989, that a key difficulty for students is understanding the meaning of assessment criteria. Classic ineffective feedback is "poor conclusion" or "not critical enough" because exactly what the student doesn't understand is what is not expressed there: the meaning, and its operationalisation, of "good conclusion" or "critical argument".

Not all criteria are difficult.
But the ones that are, are not just poorly communicated. They typically are the ones that lie at the heart of a discipline's tacit definition of itself. In other words, they are the most important thing a student must learn during their degree; and often, staff cannot easily explain them: they may be held as tacit knowledge.

30

Core disciplinary assessment criteria (2)

There is a real sense that the central learning aim of a history degree is to learn to write a history essay.

In psychology, to write a psychology essay.

In physics, to demonstrate analysis, reasoning and calculation like a physicist (not like an accountant, or mathematician, or logician)

31

Core disciplinary assessment criteria (3)

So on this account, the key question for each discipline is:
What is the assessment criterion that is closest to meaning:
  "Display thinking like a scholar in this discipline"?

Many disciplines in HE already have much of their assessment organised around a single standard format that exhibits this thinking style e.g. essays for most Arts and Social Science subjects (but actually, quite different essay types depending on the discipline), "problem solving" involving calculation i.e. inferential maths in most science and engineering.

The argument here is: Focus the feedback more effectively, not on the assessment format (i.e. not simply do lots of essays or whatever) but on learners grasping the core criterion.

32

Core disciplinary assessment criteria (4)

In my dept. by having major coursework focussed on the most difficult criterion ("being critical") was part of this. And by having them spend 3 months, rather than the 60 minutes of an exam essay, allowed more focus too.

But I've had promising results from a 5 minute version too: I.e. the radical change in time scale can also help focus on the criterion by changing context.
Reciprocal peer critiquing using the "criticality" criterion also may help here: I.e. exercising the same criterion as a reader-critic as well as as an author.

All of this could be done for other criteria e.g. spelling, reading recent not old literature, etc. But you wouldn't expect the same rewards. This is about selective targeting, not universal tips/approaches to feedback.

33

*Discussion?

What is the core disciplinary criterion in YOUR discipline?

34

Part G:

Conclusion: recommended actions

35

Strategy?

- Identify your core disciplinary (assessment) criteria (CDC)
- Focus most or all effort on training students on it:
    both student effort and staff effort
- Usually many assessments already do test them
- However the same focus may not be present in the feedback
- Furthermore supplementary exercises may be effective.

36

Why?

- The CDC are the procedural version of threshold concepts.
- They are the hardest things students have to learn
- They are the most important too: almost all assessment in fact uses them.
- They require a longer timescale to master (not one short module)

- <u>The reward is to see this learning transfer across modules; even across years and departments.</u>

37

Tactics?

- Exercise the/each criterion in both directions: not just as authors but as readers/critics

- Exercise the same criterion in tasks that are superficially very different. (learn what is common across contexts)

- Try radically different timescales.
  3 month, 1 hour, 5 minute versions.

38

A place to stop

For the slides, handout etc. see:

http://www.psy.gla.ac.uk/~steve/talks/ltc10c.html

39

Summary reminder on Biggs

He argues for the alignment of (Biggs, 2003, p.26):
- Learners (no two classes are the same)
- Teachers (no two act the same)
- Curriculum
- Teaching methods
- Assessment methods
- The climate we create in L&T interactions; inter-personal, ...
- Institutional climate: rules, procedures, ...

40